11

Two-Sample Tests of Hypothesis

GOALS

When you have completed this chapter, you will be able to:

1 Conduct a test of a hypothesis about the difference between two independent population means.

2 Conduct a test of a hypothesis about the difference between two population proportions.

3 Conduct a test of a hypothesis about the mean difference between *paired or dependent observations*.

4 Understand the difference between *dependent and independent samples*.

Gibbs Baby Food Company wishes to compare the weight gain of infants using its brand versus its competitor's. A sample of 40 babies revealed a mean weight gain of 7.6 pounds in the first three months after birth, with a population standard deviation of the sample of 2.3 pounds. A sample of 55 babies using the competitor's brand revealed a mean increase of 8.1 pounds with a 2.9 pound population standard deviation. At the .05 significance level, can we conclude the babies using the Gibbs brand gained less weight? (See Exercise 3, Goal 1.)

Statistics in Action

The U.S. presidential election of 2000 turned out to be one of the closest in history. The news media were unable to project a winner, and the final decision, including recounts and court decisions, took more than five weeks. This was not the only election in which there was controversy. Shortly before the 1936 presidential election, the *New York Times* carried the headline: "*Digest* Poll Gives Landon 32 States: Landon Leads 4–3." However, Alfred Landon of Kansas was not elected president. In fact, Roosevelt won by more than 11 million votes and received 523 Electoral College votes. How could the headline have been so wrong?

The *Literary Digest* collected a sample of voters from lists of telephone numbers, automobile registrations, and *Digest* readers. In 1936 not many people could afford a telephone or an automobile. In addition those who read the *Digest* tended to be

(continued)

Introduction

Chapter 10 began our study of hypothesis testing. We described the nature of hypothesis testing and conducted tests of a hypothesis in which we compared the results of a single sample to a population value. That is, we selected a single random sample from a population and conducted a test of whether the proposed population value was reasonable. Recall, in Chapter 10 we selected a sample of the number of desks assembled per week at Jamestown Steel Company to determine whether there was a change in the production rate. Similarly, we sampled voters in one area of a particular state to determine whether the population proportion that would support the governor for reelection was less than .80. In both of these cases, we compared the results of a *single* sample statistic to a population parameter.

In this chapter we expand the idea of hypothesis testing to two samples. That is, we select random samples from two different populations to determine whether the population means or proportions are equal. Some questions we might want to test are:

1. Is there a difference in the mean value of residential real estate sold by male agents and female agents in south Florida?
2. Is there a difference in the mean number of defects produced on the day and the afternoon shifts at Kimble Products?
3. Is there a difference in the mean number of days absent between young workers (under 21 years of age) and older workers (more than 60 years of age) in the fast-food industry?
4. Is there a difference in the proportion of Ohio State University graduates and University of Cincinnati graduates who pass the state Certified Public Accountant Examination on their first attempt?
5. Is there an increase in the production rate if music is piped into the production area?

We begin this chapter with the case in which we select random samples from two independent populations and wish to investigate whether these populations have the same mean.

Two-Sample Tests of Hypothesis: Independent Samples

A city planner in Florida wishes to know whether there is a difference in the mean hourly wage rate of plumbers and electricians in central Florida. A financial accountant wishes to know whether the mean rate of return for high yield mutual funds is different from the mean rate of return on global mutual funds. In each of these cases there are two independent populations. In the first case, the plumbers represent one population and the electricians the other. In the second case, high-yield mutual funds are one population and global mutual funds the other.

In each of these cases, to investigate the question, we would select a random sample from each population and compute the mean of the two samples. If the two wealthier and vote Republican. Thus, the population that was sampled did not represent the population of voters. A second problem was with the nonresponses. More than 10 million people were sent surveys, and more than 2.3 million responded. However, no attempt was made to see whether those responding represented a cross-section of all the voters.

With modern computers and survey methods, samples are carefully selected and checked to ensure they are representative. What happened to the *Literary Digest?* It went out of business shortly after the 1936 election.

population means are the same, that is, the mean hourly rate is the same for the plumbers and the electricians, we would expect the *difference* between the two sample means to be zero. But what if our sample results yield a difference other than zero? Is that difference due to chance or is it because there is a real difference in the hourly earnings? A two-sample test of means will help to answer this question.

We do need to return to the results of Chapter 8. Recall that we showed that a distribution of sample means would tend to approximate the normal distribution. We need to again assume that a distribution of sample means will follow the normal distribution. It can be shown mathematically that the distribution of the differences between sample means for two normal distributions is also normal.

We can illustrate this theory in terms of the city planner in Tampa, Florida. To begin, let's assume some information that is not usually available. Suppose that the population of plumbers has a mean of \$30.00 per hour and a standard deviation of \$5.00 per hour. The population of electricians has a mean of \$29.00 and a standard deviation of \$4.50. Now, from this information it is clear that the two population means are not the same. The plumbers actually earn \$1.00 per hour more than the electricians. But we cannot expect to uncover this difference each time we sample the two populations.

Suppose we select a random sample of 40 plumbers and a random sample of 35 electricians and compute the mean of each sample. Then, we determine the difference between the sample means. It is this difference between the sample means that holds our interest. If the populations have the same mean, then we would expect the difference between the two sample means to be zero. If there is a difference between the population means, then we expect to find a difference between the sample means.

To understand the theory, we need to take several pairs of samples, compute the mean of each, determine the difference between the sample means, and study the distribution of the differences in the sample means. Because of our study of the distribution of sample means in Chapter 8, we know that the distribution of the sample means follows the normal distribution. If the two distributions of sample means follow the normal distribution, then we can reason that the distribution of their differences will also follow the normal distribution. This is the first hurdle.

The second hurdle refers to the mean of this distribution of differences. If we find the mean of this distribution is zero, that implies that there is no difference in the two populations. On the other hand, if the mean of the distribution of differences is equal to some value other than zero, either positive or negative, then we conclude that the two populations do not have the same mean.

To report some concrete results, let's return to the city planner in Tampa, Florida. Table 11–1 shows the result of selecting 20 different samples of 40 plumbers and 35 electricians, computing the mean of each sample, and finding the difference between the two sample means. In the first case the sample of 40 plumbers has a mean of \$29.80, and for the 35 electricians the mean is \$28.76. The difference between the sample means is \$1.04. This process was repeated 19 more times. Observe that in 17 of the 20 cases the mean of the plumbers is larger than the mean of the electricians.

Our final hurdle is that we need to know something about the *variability* of the distribution of differences. To put it another way, what is the standard deviation of this distribution of differences? Statistical theory shows that when we have independent populations, such as the case here, the distribution of the differences has a variance (standard deviation squared) equal to the sum of the two individual variances. This means that we can add the variances of the two sampling distributions. To put it another way, the variance of the difference in sample means $(\bar{X}_1 - \bar{X}_2)$ is equal to the sum of the variance for the plumbers and the variance for the electricians.

VARIANCE OF THE DISTRIBUTION OF DIFFERENCES IN MEANS

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \qquad [11\text{–}1]$$

TABLE 11–1 The Means of Random Samples of Plumbers and Electricians

Sample	Plumbers	Electricians	Difference
1	$29.80	$28.76	$1.04
2	30.32	29.40	0.92
3	30.57	29.94	0.63
4	30.04	28.93	1.11
5	30.09	29.78	0.31
6	30.02	28.66	1.36
7	29.60	29.13	0.47
8	29.63	29.42	0.21
9	30.17	29.29	0.88
10	30.81	29.75	1.06
11	30.09	28.05	2.04
12	29.35	29.07	0.28
13	29.42	28.79	0.63
14	29.78	29.54	0.24
15	29.60	29.60	0.00
16	30.60	30.19	0.41
17	30.79	28.65	2.14
18	29.14	29.95	−0.81
19	29.91	28.75	1.16
20	28.74	29.21	−0.47

The term $\sigma^2_{\overline{X}_1-\overline{X}_2}$ looks complex but need not be difficult to interpret. The σ^2 portion reminds us that it is a variance, and the subscript $\overline{X}_1 - \overline{X}_2$ that it is a distribution of differences in the sample means.

We can put this equation in a more usable form by taking the square root, so that we have the standard deviation of the distribution or "standard error" of the differences. Finally, we standardize the distribution of the differences. The result is the following equation.

TWO-SAMPLE TEST OF MEANS—KNOWN σ

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \qquad \textbf{[11–2]}$$

Before we present an example, let's review the assumptions necessary for using formula (11–2).

Assumptions for testing independent sample means.

1. The two samples must be unrelated, that is, independent.
2. The standard deviations for both populations must be known.

The following example shows the details of the test of hypothesis for two population means.

Example

Customers at FoodTown Super Markets have a choice when paying for their groceries. They may check out and pay using the standard cashier-assisted checkout, or they may use the new U-Scan procedure. In the standard procedure a FoodTown employee scans each item, puts it on a short conveyor where another employee puts it in a bag and then into the grocery cart. In the U-Scan procedure the customer scans each item, bags it, and places the bags in the cart themselves. The U-Scan procedure is designed to reduce the time a customer spends in the checkout line.

The U-Scan facility was recently installed at the Byrne Road Food-Town location. The store manager would like to know if the mean checkout time using the standard checkout method is longer than using the U-Scan. She gathered the following sample information. The time is measured from when the customer enters the line until their bags are in the cart. Hence the time includes both waiting in line and checking out. What is the *p*-value?

Customer Type	Sample Mean	Population Standard Deviation	Sample Size
Standard	5.50 minutes	0.40 minutes	50
U-Scan	5.30 minutes	0.30 minutes	100

Solution

We use the five-step hypothesis testing procedure to investigate the question.

Step 1: State the null hypothesis and the alternate hypothesis. The null hypothesis is that there is no difference in the mean checkout times for the two groups. In other words, the difference of 0.20 minutes between the mean checkout time for the standard method and the mean checkout time for U-Scan is due to chance. The alternate hypothesis is that the mean checkout time is longer for those using the standard method. We will let μ_s refer to the mean checkout time for the population of standard customers and μ_u the mean checkout time for the U-Scan customers. The null and alternative hypotheses are:

$$H_0: \mu_s \leq \mu_u$$
$$H_1: \mu_s > \mu_u$$

Step 2: Select the level of significance. The significance level is the probability that we reject the null hypothesis when it is actually true. This likelihood is determined prior to selecting the sample or performing any calculations. The .05 and .01 significance levels are the most common, but other values, such as .02 and .10, are also used. In theory, we may select any value between 0 and 1 for the significance level. In this case we selected the .01 significance level.

Step 3: Determine the test statistic. In Chapter 10 we used the standard normal distribution (that is z) and t as test statistics. In this case we use the z distribution as the test statistic because the standard deviations of both populations are known.

Step 4: Formulate a decision rule. The decision rule is based on the null and the alternate hypotheses (i.e., one-tailed or two-tailed test), the level of significance, and the test statistic used. We selected the .01 significance level and the z distribution as the test statistic, and we wish to determine whether the mean checkout time is longer using the standard method. We set the alternate hypothesis to indicate that the mean checkout time is longer for those using the standard method than the U-Scan method. Hence, the rejection region is in the upper tail of the standard normal distribution (a one-tailed test). To find the critical value, place .01 of the total area in the upper tail. This means that .4900 (.5000 − .0100) of the area is located between the z value of 0 and the critical value. Next, we search the body of Appendix B.1 for a value located near .4900. It is 2.33, so our decision rule is to reject H_0 if the value computed from the test statistic exceeds 2.33. Chart 11–1 depicts the decision rule.

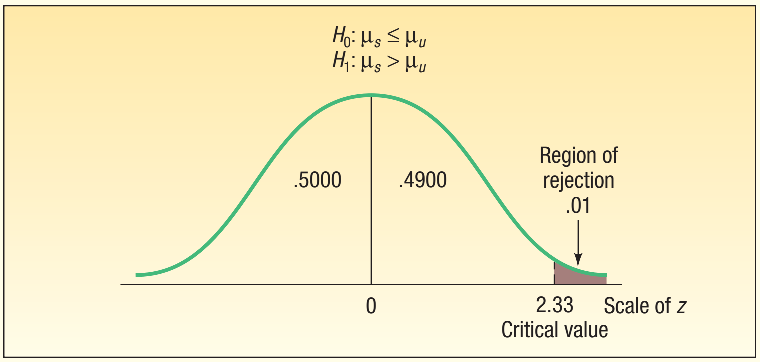


CHART 11–1 Decision Rule for One-Tailed Test at .01 Significance Level

Statistics in Action

Do you live to work or work to live? A recent poll of 802 working Americans revealed that, among those who considered their work as a career, the mean number of hours worked per day was 8.7. Among those who considered their work as a job, the mean number of hours worked per day was 7.6.

Step 5: Make the decision regarding H_0 and interpret the result. We use formula (11–2) to compute the value of the test statistic.

$$z = \frac{\bar{X}_s - \bar{X}_u}{\sqrt{\frac{\sigma_s^2}{n_s} + \frac{\sigma_u^2}{n_u}}} = \frac{5.5 - 5.3}{\sqrt{\frac{0.40^2}{50} + \frac{0.30^2}{100}}} = \frac{0.2}{0.064} = 3.13$$

The computed value of 3.13 is larger than the critical value of 2.33. Our decision is to reject the null hypothesis and accept the alternate hypothesis. The difference of .20 minutes between the mean checkout time using the standard method is too large to have occurred by chance. To put it another way, we conclude the U-Scan method is faster.

What is the p-value for the test statistic? Recall that the p-value is the probability of finding a value of the test statistic this extreme when the null hypothesis is true. To calculate the p-value we need the probability of a z value larger than 3.13. From Appendix B.1 we cannot find the probability associated with 3.13. The largest value available is 3.09. The area corresponding to 3.09 is .4990. In this case we can report that the p-value is less than .0010, found by .5000 − .4990. We conclude that there is very little likelihood that the null hypothesis is true!

In summary, the criteria for using formula (11–2) are:

1. *The samples are from independent populations*. This means, for example, that the checkout time for the U-Scan customers is unrelated to the checkout time for the other customers. For example, Mr. Smith's checkout time does not affect any other customer's checkout time.
2. *Both population standard deviations are known*. In the FoodTown example, the population standard deviation of the U-Scan times was 0.30 minutes. The standard deviation of the standard checkout times was 0.40 minutes. We use formula (11–2) to find the value of the test statistic.

Self-Review 11–1

Tom Sevits is the owner of the Appliance Patch. Recently Tom observed a difference in the dollar value of sales between the men and women he employs as sales associates. A sample of 40 days revealed the men sold a mean of $1,400 worth of appliances per day. For a sample of 50 days, the women sold a mean of $1,500 worth of appliances per day. Assume the population standard deviation for men is $200 and for women $250. At the .05 significance level can Mr. Sevits conclude that the mean amount sold per day is larger for the women?

(a) State the null hypothesis and the alternate hypothesis.
(b) What is the decision rule?
(c) What is the value of the test statistic?
(d) What is your decision regarding the null hypothesis?
(e) What is the *p*-value?
(f) Interpret the result.

Exercises

1. A sample of 40 observations is selected from one population with a population standard deviation of 5. The sample mean is 102. A sample of 50 observations is selected from a second population with a population standard deviation of 6. The sample mean is 99. Conduct the following test of hypothesis using the .04 significance level.

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

a. Is this a one-tailed or a two-tailed test?
b. State the decision rule.
c. Compute the value of the test statistic.
d. What is your decision regarding H_0?
e. What is the *p*-value?

2. A sample of 65 observations is selected from one population with a population standard deviation of 0.75. The sample mean is 2.67. A sample of 50 observations is selected from a second population with a population standard deviation of 0.66. The sample mean is 2.59. Conduct the following test of hypothesis using the .08 significance level.

H_0: $\mu_1 \leq \mu_2$
H_1: $\mu_1 > \mu_2$

a. Is this a one-tailed or a two-tailed test?
b. State the decision rule.
c. Compute the value of the test statistic.
d. What is your decision regarding H_0?
e. What is the *p*-value?

Note: Use the five-step hypothesis testing procedure to solve the following exercises.

3. Gibbs Baby Food Company wishes to compare the weight gain of infants using its brand versus its competitor's. A sample of 40 babies using the Gibbs products revealed a mean weight gain of 7.6 pounds in the first three months after birth. For the Gibbs brand the population standard deviation of the sample is 2.3 pounds. A sample of 55 babies using the competitor's brand revealed a mean increase in weight of 8.1 pounds. The population standard deviation is 2.9 pounds. At the .05 significance level, can we conclude that babies using the Gibbs brand gained less weight? Compute the *p*-value and interpret it.

4. As part of a study of corporate employees, the director of human resources for PNC, Inc., wants to compare the distance traveled to work by employees at its office in downtown Cincinnati with the distance for those in downtown Pittsburgh. A sample of 35 Cincinnati employees showed they travel a mean of 370 miles per month. A sample of 40 Pittsburgh employees showed they travel a mean of 380 miles per month. The population standard deviation for the Cincinnati and Pittsburgh employees are 30 and 26 miles, respectively. At the .05 significance level, is there a difference in the mean number of miles traveled per month between Cincinnati and Pittsburgh employees?

5. A financial analyst wants to compare the turnover rates, in percent, for shares of oil-related stocks versus other stocks, such as GE and IBM. She selected 32 oil-related stocks and 49 other stocks. The mean turnover rate of oil-related stocks is 31.4 percent and the population standard deviation 5.1 percent. For the other stocks, the mean rate was computed to be 34.9 percent and the population standard deviation 6.7 percent. Is there a significant difference in the turnover rates of the two types of stock? Use the .01 significance level.

6. Mary Jo Fitzpatrick is the vice president for Nursing Services at St. Luke's Memorial Hospital. Recently she noticed in the job postings for nurses that those that are unionized seem to offer higher wages. She decided to investigate and gathered the following information.

Group	Mean Wage	Population Standard Deviation	Sample Size
Union	\$20.75	\$2.25	40
Nonunion	\$19.80	\$1.90	45

Would it be reasonable for her to conclude that union nurses earn more? Use the .02 significance level. What is the *p*-value?

Two-Sample Tests about Proportions

In the previous section, we considered a test involving population means. However, we are often interested also in whether two sample proportions come from populations that are equal. Here are several examples.

- The vice president of human resources wishes to know whether there is a difference in the proportion of hourly employees who miss more than 5 days of work per year at the Atlanta and the Houston plants.
- General Motors is considering a new design for the Pontiac G6. The design is shown to a group of potential buyers under 30 years of age and another group over 60 years of age. Pontiac wishes to know whether there is a difference in the proportion of the two groups who like the new design.
- A consultant to the airline industry is investigating the fear of flying among adults. Specifically, the company wishes to know whether there is a difference in the proportion of men versus women who are fearful of flying.

In the above cases each sampled item or individual can be classified as a "success" or a "failure." That is, in the Pontiac G6 example each potential buyer is classified as "liking the new design" or "not liking the new design." We then compare the proportion in the under 30 group with the proportion in the over 60 group who indicated they liked the new design. Can we conclude that the differences are due to chance? In this study there is no measurement obtained, only classifying the individuals or objects. Then we assume the nominal scale of measurement.

To conduct the test, we assume each sample is large enough that the normal distribution will serve as a good approximation of the binomial distribution. The test statistic follows the standard normal distribution. We compute the value of z from the following formula:

TWO-SAMPLE TEST OF PROPORTIONS

$$z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_c(1 - p_c)}{n_1} + \dfrac{p_c(1 - p_c)}{n_2}}} \quad [11\text{–}3]$$

Formula (11–3) is formula (11–2) with the respective sample proportions replacing the sample means and $p_c(1 - p_c)$ replacing the two variances. In addition:

n_1 is the number of observations in the first sample.
n_2 is the number of observations in the second sample.
p_1 is the proportion in the first sample possessing the trait.
p_2 is the proportion in the second sample possessing the trait.
p_c is the pooled proportion possessing the trait in the combined samples. It is called the pooled estimate of the population proportion and is computed from the following formula.

POOLED PROPORTION

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} \quad [11\text{–}4]$$

where:
X_1 is the number possessing the trait in the first sample.
X_2 is the number possessing the trait in the second sample.

The following example will illustrate the two-sample test of proportions.

Example

Manelli Perfume Company recently developed a new fragrance that it plans to market under the name Heavenly. A number of market studies indicate that Heavenly has very good market potential. The Sales Department at Manelli is particularly interested in whether there is a difference in the proportions of younger and older women who would purchase Heavenly if it were marketed. There are two independent populations, a population consisting of the younger women and a population consisting of the older women. Each sampled woman will be asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase a bottle.

Solution

We will use the usual five-step hypothesis-testing procedure.

Step 1: State H_0 and H_1. In this case the null hypothesis is: "There is no difference in the proportion of young women and older women who prefer Heavenly." We designate π_1 as the proportion of young women who would purchase Heavenly and π_2 as the proportion of older women who would purchase. The alternate hypothesis is that the two proportions are not equal.

$$H_0: \pi_1 = \pi_2$$
$$H_1: \pi_1 \neq \pi_2$$

Step 2: Select the level of significance. We choose the .05 significance level in this example.

Step 3: Determine the test statistic. The test statistic follows the standard normal distribution. The value of the test statistic can be computed from formula (11–3).

Step 4: Formulate the decision rule. Recall that the alternate hypothesis from step 1 does not state a direction, so this is a two-tailed test. To determine the critical value, we divide the significance level in half and place this amount in each tail of the z distribution. Next, we subtract this amount from the total area to the right of zero. That is $.5000 - .0250 = .4750$. Finally, we search the body of the z table (Appendix B.1) for the closest value. It is 1.96. The critical values are -1.96 and $+1.96$. As before, if the computed z value falls in the region between $+1.96$ and -1.96, the null hypothesis is not rejected. If that does occur, it is assumed that any difference between the two sample proportions is due to chance variation. This information is summarized in Chart 11–2.

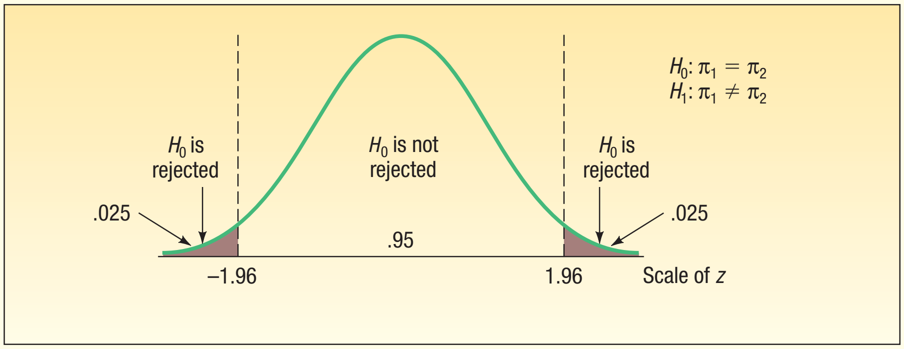


CHART 11–2 Decision Rules for Heavenly Fragrance Test, .05 Significance Level

Step 5: Select a sample and make a decision. A random sample of 100 young women revealed 19 liked the Heavenly fragrance well enough to purchase it. Similarly, a sample of 200 older women revealed 62 liked the fragrance well enough to make a purchase. We let p_1 refer to the young women and p_2 to the older women.

$$p_1 = \frac{X_1}{n_1} = \frac{19}{100} = .19 \qquad p_2 = \frac{X_2}{n_2} = \frac{62}{200} = .31$$

The research question is whether the difference of .12 in the two sample proportions is due to chance or whether there is a difference in the proportion of younger and older women who like the Heavenly fragrance.

Next, we combine or pool the sample proportions. We use formula (11–4).

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{19 + 62}{100 + 200} = \frac{81}{300} = 0.27$$

Note that the pooled proportion is closer to .31 than to .19 because more older women than younger women were sampled.

We use formula (11–3) to find the value of the test statistic.

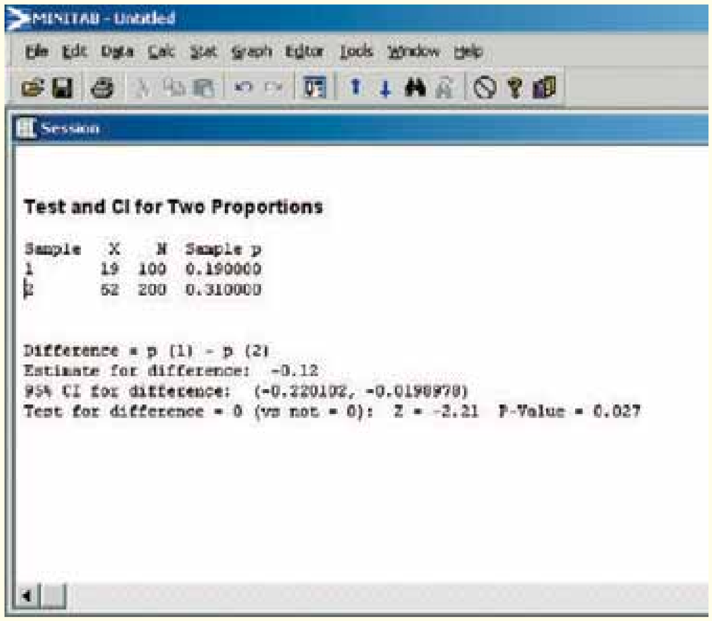
$$z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_c(1 - p_c)}{n_1} + \dfrac{p_c(1 - p_c)}{n_2}}} = \frac{.19 - .31}{\sqrt{\dfrac{.27(1 - .27)}{100} + \dfrac{.27(1 - .27)}{200}}} = -2.21$$

The computed value of −2.21 is in the area of rejection; that is, it is to the left of −1.96. Therefore, the null hypothesis is rejected at the .05 significance level. To put it another way, we reject the null hypothesis that the proportion of young women who would purchase Heavenly is equal to the proportion of older women who would purchase Heavenly. It is unlikely that the difference between the two sample proportions is due to chance. To find the p-value, go to Appendix B.1 and find the likelihood of a z value less than −2.21 or greater than 2.21. The z value corresponding to 2.21 is .4864. So the likelihood of finding the value of the test statistic to be less than −2.21 or greater than 2.21 is:

$$p\text{-value} = 2(.5000 - .4864) = 2(.0136) = .0272$$

The p-value of .0272 is less than the significance level of .05, so our decision is to reject the null hypothesis. Again, we conclude that there is a difference in the proportion of younger and older women who would purchase Heavenly.

The MINITAB system has a procedure to quickly determine the value of the test statistic and compute the *p*-value. The results follow.



Notice the MINITAB output includes the two sample proportions, the value of *z*, and the *p*-value.

Self-Review 11–2

Of 150 adults who tried a new peach-flavored Peppermint Patty, 87 rated it excellent. Of 200 children sampled, 123 rated it excellent. Using the .10 level of significance, can we conclude that there is a significant difference in the proportion of adults and the proportion of children who rate the new flavor excellent?

(a) State the null hypothesis and the alternate hypothesis.
(b) What is the probability of a Type I error?
(c) Is this a one-tailed or a two-tailed test?
(d) What is the decision rule?
(e) What is the value of the test statistic?
(f) What is your decision regarding the null hypothesis?
(g) What is the *p*-value? Explain what it means in terms of this problem.

Exercises

7. The null and alternate hypotheses are:

$$H_0: \pi_1 \leq \pi_2$$
$$H_1: \pi_1 > \pi_2$$

A sample of 100 observations from the first population indicated that X_1 is 70. A sample of 150 observations from the second population revealed X_2 to be 90. Use the .05 significance level to test the hypothesis.

a. State the decision rule.
b. Compute the pooled proportion.
c. Compute the value of the test statistic.
d. What is your decision regarding the null hypothesis?

8. The null and alternate hypotheses are:

$$H_0: \pi_1 = \pi_2$$
$$H_1: \pi_1 \neq \pi_2$$

A sample of 200 observations from the first population indicated that X_1 is 170. A sample of 150 observations from the second population revealed X_2 to be 110. Use the .05 significance level to test the hypothesis.
a. State the decision rule.
b. Compute the pooled proportion.
c. Compute the value of the test statistic.
d. What is your decision regarding the null hypothesis?

Note: Use the five-step hypothesis-testing procedure in solving the following exercises.

9. The Damon family owns a large grape vineyard in western New York along Lake Erie. The grapevines must be sprayed at the beginning of the growing season to protect against various insects and diseases. Two new insecticides have just been marketed: Pernod 5 and Action. To test their effectiveness, three long rows were selected and sprayed with Pernod 5, and three others were sprayed with Action. When the grapes ripened, 400 of the vines treated with Pernod 5 were checked for infestation. Likewise, a sample of 400 vines sprayed with Action were checked. The results are:

Insecticide	Number of Vines Checked (sample size)	Number of Infested Vines
Pernod 5	400	24
Action	400	40

At the .05 significance level, can we conclude that there is a difference in the proportion of vines infested using Pernod 5 as opposed to Action?

10. The Roper Organization conducted identical surveys 5 years apart. One question asked of women was “Are most men basically kind, gentle, and thoughtful?” The earlier survey revealed that, of the 3,000 women surveyed, 2,010 said that they were. The later revealed 1,530 of the 3,000 women surveyed thought that men were kind, gentle, and thoughtful. At the .05 level, can we conclude that women think men are less kind, gentle, and thoughtful in the later survey compared with the earlier one?

11. A nationwide sample of influential Republicans and Democrats was asked as a part of a comprehensive survey whether they favored lowering environmental standards so that high-sulfur coal could be burned in coal-fired power plants. The results were:

	Republicans	Democrats
Number sampled	1,000	800
Number in favor	200	168

At the .02 level of significance, can we conclude that there is a larger proportion of Democrats in favor of lowering the standards? Determine the *p*-value.

12. The research department at the home office of New Hampshire Insurance conducts ongoing research on the causes of automobile accidents, the characteristics of the drivers, and so on. A random sample of 400 policies written on single persons revealed 120 had at least one accident in the previous three-year period. Similarly, a sample of 600 policies written on married persons revealed that 150 had been in at least one accident. At the .05 significance level, is there a significant difference in the proportions of single and married persons having an accident during a three-year period? Determine the *p*-value.

Comparing Means with Unknown Population Standard Deviations (the Pooled *t*-test)

In the previous two sections we described conditions where the standard normal distribution, that is *z*, is used as the test statistic. In one case we were working with a variable (calculating the mean) and in the second an attribute (calculating a proportion). In the first case we wished to compare two sample means from independent populations to determine if they came from the same or equal populations. In that instance we assumed the population followed the normal probability distribution and that we knew the standard deviation of the population. In many cases, in fact in most cases, we do not know the population standard deviation. We can overcome this problem, as we did in the one sample case in the previous chapter, by substituting the sample standard deviation (s) for the population standard deviation (σ). See formula (10–2) on page 345.

This section describes another method for comparing the sample means of two independent populations to determine if the sampled populations could reasonably have the same mean. The method described does *not* require that we know the standard deviations of the populations. This gives us a great deal more flexibility when investigating the difference in sample means. There are two major differences in this test and the previous test described earlier in this chapter.

1. We assume the sampled populations have equal but unknown standard deviations. Because of this assumption we combine or "pool" the sample standard deviations.
2. We use the t distribution as the test statistic.

The formula for computing the value of the test statistic t is similar to (11–2), but an additional calculation is necessary. The two sample standard deviations are pooled to form a single estimate of the unknown population standard deviation. In essence we compute a weighted mean of the two sample standard deviations and use this value as an estimate of the unknown population standard deviation. The weights are the degrees of freedom that each sample provides. Why do we need to pool the sample standard deviations? Because we assume that the two populations have equal standard deviations, the best estimate we can make of that value is to combine or pool all the sample information we have about the value of the population standard deviation.

The following formula is used to pool the sample standard deviations. Notice that two factors are involved: the number of observations in each sample and the sample standard deviations themselves.

POOLED VARIANCE

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad \textbf{[11–5]}$$

where:

s_1^2 is the variance (standard deviation squared) of the first sample.

s_2^2 is the variance of the second sample.

The value of t is computed from the following equation.

TWO-SAMPLE TEST OF MEANS—UNKNOWN σ'S

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \textbf{[11–6]}$$

where:

$\bar{X}_1$ is the mean of the first sample.
$\bar{X}_2$ is the mean of the second sample.
n_1 is the number of observations in the first sample.
n_2 is the number of observations in the second sample.
s_p^2 is the pooled estimate of the population variance.

The number of degrees of freedom in the test is the total number of items sampled minus the total number of samples. Because there are two samples, there are $n_1 + n_2 - 2$ degrees of freedom.

To summarize, there are three requirements or assumptions for the test.

1. The sampled populations follow the normal distribution.
2. The sampled populations are independent.
3. The standard deviations of the two populations are equal.

The following example/solution explains the details of the test.

Example

Owens Lawn Care, Inc., manufactures and assembles lawnmowers that are shipped to dealers throughout the United States and Canada. Two different procedures have been proposed for mounting the engine on the frame of the lawnmower. The question is: Is there a difference in the mean time to mount the engines on the frames of the lawnmowers? The first procedure was developed by longtime Owens employee Herb Welles (designated as procedure 1), and the other procedure was developed by Owens Vice President of Engineering William Atkins (designated as procedure 2). To evaluate the two methods, it was decided to conduct a time and motion study. A sample of five employees was timed using the Welles method and six using the Atkins method. The results, in minutes, are shown below. Is there a difference in the mean mounting times? Use the .10 significance level.

Welles (minutes)	Atkins (minutes)
2	3
4	7
9	5
3	8
2	4
	3

Solution

Following the five steps to test a hypothesis, the null hypothesis states that there is no difference in mean mounting times between the two procedures. The alternate hypothesis indicates that there is a difference.

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

The required assumptions are:

1. The observations in the Welles sample are *independent* of the observations in the Atkins sample.
2. The two populations follow the normal distribution.
3. The two populations have equal standard deviations.

Is there a difference between the mean assembly times using the Welles and the Atkins methods? The degrees of freedom are equal to the total number of items sampled minus the number of samples. In this case that is $n_1 + n_2 - 2$. Five assemblers used the Welles method and six the Atkins method. Thus, there are 9 degrees of freedom, found by $5 + 6 - 2$. The critical values of t, from Appendix B.2 for $df = 9$, a two-tailed test, and the .10 significance level, are -1.833 and 1.833. The decision rule is portrayed graphically in Chart 11–3. We do not reject the null hypothesis if the computed value of t falls between -1.833 and 1.833.

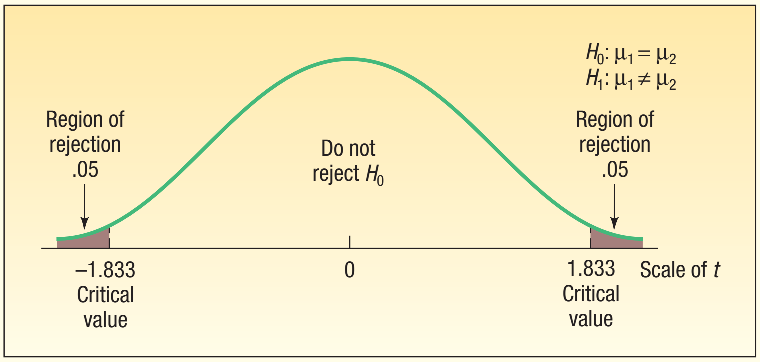


CHART 11–3 Regions of Rejection, Two-Tailed Test, $df = 9$, and .10 Significance Level

We use three steps to compute the value of t.

Step 1: Calculate the sample standard deviations. See the details below.

Welles Method		Atkins Method	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2 - 4)^2 = 4$	3	$(3 - 5)^2 = 4$
4	$(4 - 4)^2 = 0$	7	$(7 - 5)^2 = 4$
9	$(9 - 4)^2 = 25$	5	$(5 - 5)^2 = 0$
3	$(3 - 4)^2 = 1$	8	$(8 - 5)^2 = 9$
2	$(2 - 4)^2 = 4$	4	$(4 - 5)^2 = 1$
20	34	3	$(3 - 5)^2 = 4$
		30	22

$$\bar{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{20}{5} = 4 \qquad \bar{X}_2 = \frac{\Sigma X_2}{n_2} = \frac{30}{6} = 5$$

$$s_1 = \sqrt{\frac{\Sigma(X_1 - \bar{X}_1)^2}{n_1 - 1}} = \sqrt{\frac{34}{5 - 1}} = 2.9155 \qquad s_2 = \sqrt{\frac{\Sigma(X_2 - \bar{X}_2)^2}{n_2 - 1}} = \sqrt{\frac{22}{6 - 1}} = 2.0976$$

Step 2: Pool the Sample Variances. We use formula (11–5) to pool the sample variances (standard deviations squared).

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(5 - 1)(2.9155)^2 + (6 - 1)(2.0976)^2}{5 + 6 - 2} = 6.2222$$

Step 3: Determine the value of *t*. The mean mounting time for the Welles method is 4.00 minutes, found by $\bar{X}_1 = 20/5$. The mean mounting time for the Atkins method is 5.00 minutes, found by $\bar{X}_2 = 30/6$. We use formula (11–6) to calculate the value of t.

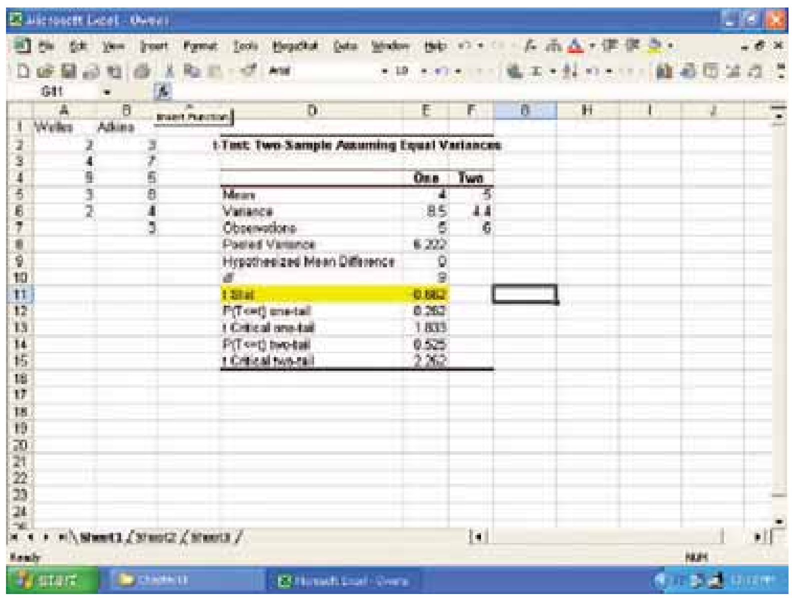
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{4.00 - 5.00}{\sqrt{6.2222\left(\frac{1}{5} + \frac{1}{6}\right)}} = -0.662$$

The decision is not to reject the null hypothesis, because −0.662 falls in the region between −1.833 and 1.833. We conclude that there is no difference in the mean times to mount the engine on the frame using the two methods.

We can also estimate the *p*-value using Appendix B.2. Locate the row with 9 degrees of freedom, and use the two-tailed test column. Find the *t* value, without regard to the sign, which is closest to our computed value of 0.662. It is 1.383, corresponding to a significance level of .20. Thus, even had we used the 20 percent significance level, we would not have rejected the null hypothesis of equal means. We can report that the *p*-value is greater than .20.

Excel has a procedure called "t-Test: Two Sample Assuming Equal Variances" that will perform the calculations of formulas (11–5) and (11–6) as well as find the sample means and sample variances. The data are input in the first two columns of the Excel spreadsheet. They are labeled "Welles" and "Atkins." The output follows. The value of *t*, called the "t Stat," is −0.662, and the two-tailed *p*-value is .525. As we would expect, the *p*-value is larger than the significance level of .10. The conclusion is not to reject the null hypothesis.

Excel



Self-Review 11–3

The production manager at Bellevue Steel, a manufacturer of wheelchairs, wants to compare the number of defective wheelchairs produced on the day shift with the number on the afternoon shift. A sample of the production from 6 day shifts and 8 afternoon shifts revealed the following number of defects.

Day	5	8	7	6	9	7		
Afternoon	8	10	7	11	9	12	14	9

At the .05 significance level, is there a difference in the mean number of defects per shift?

(a) State the null hypothesis and the alternate hypothesis.
(b) What is the decision rule?
(c) What is the value of the test statistic?
(d) What is your decision regarding the null hypothesis?
(e) What is the *p*-value?
(f) Interpret the result.
(g) What are the assumptions necessary for this test?

Exercises

For Exercises 13 and 14: (a) state the decision rule, (b) compute the pooled estimate of the population variance, (c) compute the test statistic, (d) state your decision about the null hypothesis, and (e) estimate the *p*-value.

13. The null and alternate hypotheses are:

H_0: $\mu_1 = \mu_2$.
H_1: $\mu_1 \neq \mu_2$

A random sample of 10 observations from one population revealed a sample mean of 23 and a sample deviation of 4. A random sample of 8 observations from another population revealed a sample mean of 26 and a sample standard deviation of 5. At the .05 significance level, is there a difference between the population means?

14. The null and alternate hypotheses are:

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

A random sample of 15 observations from the first population revealed a sample mean of 350 and a sample standard deviation of 12. A random sample of 17 observations from the second population revealed a sample mean of 342 and a sample standard deviation of 15. At the .10 significance level, is there a difference in the population means?

Note: Use the five-step hypothesis testing procedure for the following exercises.

15. A sample of scores on an examination given in Statistics 201 are:

Men	72	69	98	66	85	76	79	80	77
Women	81	67	90	78	81	80	76		

At the .01 significance level, is the mean grade of the women higher than that of the men?

16. A recent study compared the time spent together by single- and dual-earner couples. According to the records kept by the wives during the study, the mean amount of time spent together watching television among the single-earner couples was 61 minutes per day, with a standard deviation of 15.5 minutes. For the dual-earner couples, the mean number of minutes spent watching television was 48.4 minutes, with a standard deviation of 18.1 minutes. At the .01 significance level, can we conclude that the single-earner couples on average spend more time watching television together? There were 15 single-earner and 12 dual-earner couples studied.

17. Ms. Lisa Monnin is the budget director for Nexus Media, Inc. She would like to compare the daily travel expenses for the sales staff and the audit staff. She collected the following sample information.

Sales ($)	131	135	146	165	136	142	
Audit ($)	130	102	129	143	149	120	139

At the .10 significance level, can she conclude that the mean daily expenses are greater for the sales staff than the audit staff? What is the *p*-value?

18. The Tampa Bay (Florida) Area Chamber of Commerce wanted to know whether the mean weekly salary of nurses was larger than that of school teachers. To investigate, they collected the following information on the amounts earned last week by a sample of school teachers and nurses.

School Teachers ($)	845	826	827	875	784	809	802	820	829	830	842	832
Nurses ($)	841	890	821	771	850	859	825	829				

Is it reasonable to conclude that the mean weekly salary of nurses is higher? Use the .01 significance level. What is the *p*-value?

Comparing Population Means with Unequal Standard Deviations

In the previous sections it was necessary to assume that the populations had equal standard deviations. To put it another way, we did not know the population standard deviations but we assumed they were equal. In many cases this is a reasonable assumption, but what if it is not? In the next chapter we present a formal method for testing this equal variance assumption.

If it is not reasonable to assume the population standard deviations are equal, then we use a statistic very much like formula [11–2]. The sample standard deviations, s_1 and s_2, are used in place of the respective population standard deviations. In addition, the degrees of freedom are adjusted downward by a rather complex approximation formula. The effect is to reduce the number of degrees of freedom in the test, which will require a larger value of the test statistic to reject the null hypothesis.

The formula for the t statistic is:

TEST STATISTIC FOR NO DIFFERENCE IN MEANS, UNEQUAL VARIANCES

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \qquad \textbf{[11–7]}$$

The degrees of freedom statistic is found by:

DEGREES OF FREEDOM FOR UNEQUAL VARIANCE TEST

$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \qquad \textbf{[11–8]}$$

where n_1 and n_2 are the respective sample sizes and s_1 and s_2 are the respective sample standard deviations. If necessary, this fraction is rounded down to an integer value. An example will explain the details.

Example

Personnel in a consumer testing laboratory are evaluating the absorbency of paper towels. They wish to compare a set of store brand towels to a similar group of name brand ones. For each brand they dip a ply of the paper into a tub of fluid, allow the paper to drain back into the vat for two minutes, and then evaluate the amount of liquid the paper has taken up from the vat. A random sample of 9 store brand paper towels absorbed the following amounts of liquid in milliliters.

8	8	3	1	9	7	5	5	12

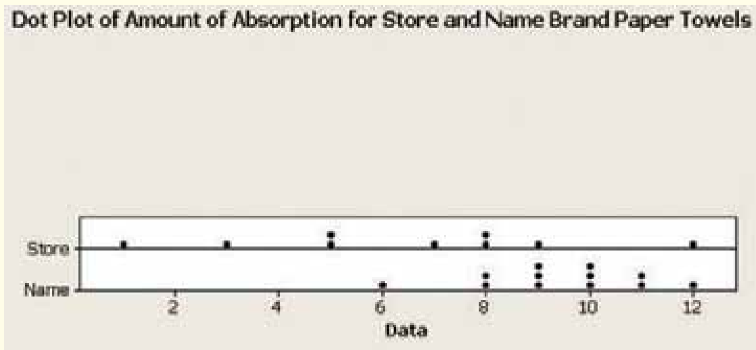
An independent random sample of 12 name brand towels absorbed the following amounts of liquid in milliliters:

12	11	10	6	8	9	9	10	11	9	8	10

Use the .10 significance level and test if there is a difference in the mean amount of liquid absorbed by the two types of paper towels.

Solution

To begin let's assume that the amounts of liquid absorbed follow the normal probability distribution for both the store brand and the name brand towels. We do not know either of the population standard deviations so we are going to use the *t* distribution as the test statistic. The assumption of equal population standard deviations does not appear reasonable. The amount of absorption in the store brand ranges from 1 ml to 12 ml. For the name brand, the amount of absorption ranges from 6 ml to 12 ml. That is, there is more variation in the amount of absorption in the store brand than in the name brand. We observe the difference in the variation in the following dot plot provided by MINITAB. The software commands to create a MINITAB dot plot are given on page 129.



So we decide to use the *t* distribution and assume that the population standard deviations are not the same.

In the five-step hypothesis testing procedure the first step is to state the null hypothesis and the alternate hypothesis. The null hypothesis is that there is no difference in the mean amount of liquid absorbed between the two types of paper towels. The alternate hypothesis is that there is a difference.

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

The significance level is .10 and the test statistic follows the *t* distribution. Because we do not wish to assume equal population standard deviations, we adjust the degrees of freedom using formula (11–8). To do so we need to find the sample standard deviations. We can use the MINITAB system to quickly find these results. We will also find the mean absorption rate, which we will use shortly. The respective samples sizes are $n_1 = 9$ and $n_2 = 12$ and the respective standard deviations are 3.32 ml and 1.621 ml.

Descriptive Statistics: Store, Name

Variable	N	Mean	StDev
Store	9	6.44	3.32
Name	12	9.417	1.621

Inserting this information into formula (11–8):

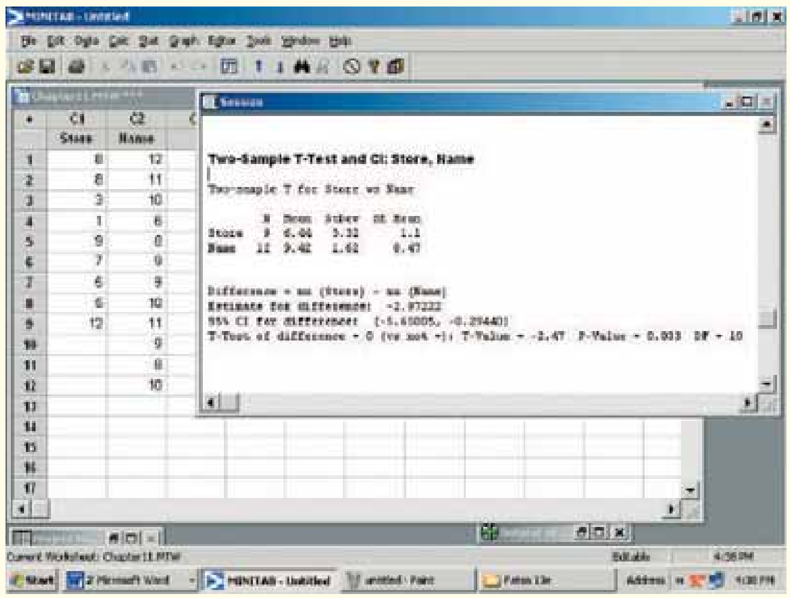
$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{[(3.32^2/9) + (1.621^2/12)]^2}{\dfrac{(3.32^2/9)^2}{9 - 1} + \dfrac{(1.621^2/12)^2}{12 - 1}} = \frac{1.4436^2}{.1875 + .0043} = 10.88$$

The usual practice is to round down to the integer, so we use 10 degrees of freedom. From Appendix B.2 with 10 degrees of freedom, a two-tailed test, and the .10 significance level, the critical *t* values are −1.812 and 1.812. Our decision rule is to reject the null hypothesis if the computed value of *t* is less than −1.812 or greater than 1.812.

To find the value of the test statistic we use formula (11–7). Recall from the MINITAB output above the mean amount of absorption for the store paper towels is 6.44 ml and 9.417 ml for the brand.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{6.44 - 9.417}{\sqrt{\dfrac{3.32^2}{9} + \dfrac{1.621^2}{12}}} = -2.478$$

The computed value of t is less than the lower critical value, so our decision is to reject the null hypothesis. We conclude that the mean absorption rate for the two towels is not the same. The MINITAB output for this example follows.



Minitab

Self-Review 11–4

It is often useful for companies to know who their customers are and how they became customers. A credit card company is interested in whether the owner of the card applied for the card on their own or was contacted by a telemarketer. The company obtained the following sample information regarding end-of-the-month balances for the two groups.

Source	Mean	Standard Deviation	Sample Size
Applied	$1,568	$356	10
Contacted	1,967	857	8

Is it reasonable to conclude the mean balance is larger for the credit card holders that were contacted by telemarketers than for those who applied on their own for the card? Assume the population standard deviations are not the same. Use the .05 significance level.

(a) State the null hypothesis and the alternate hypothesis.
(b) How many degrees of freedom are there?
(c) What is the decision rule?
(d) What is the value of the test statistic?
(e) What is your decision regarding the null hypothesis?
(f) Interpret the result.

Exercises

For exercises 19 and 20 assume the sample populations do not have equal standard deviations and use the .05 significance level: (a) determine the number of degrees of freedom, (b) state the decision rule, (c) compute the value of the test statistic, and (d) state your decision about the null hypothesis.

19. The null and alternate hypotheses are:

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

A random sample of 15 items from the first population showed a mean of 50 and a standard deviation of 5. A sample of 12 items for the second population showed a mean of 46 and a standard deviation of 15.

20. The null and alternate hypotheses are:

H_0: $\mu_1 \leq \mu_2$
H_1: $\mu_1 > \mu_2$

A random sample of 20 items from the first population showed a mean of 100 and a standard deviation of 15. A sample of 16 items for the second population showed a mean of 94 and a standard deviation of 8. Use the .05 significant level.

21. A recent article in *The Wall Street Journal* compared the cost of adopting children from China with that of Russia. For a sample of 16 adoptions from China the mean cost was $11,045 with a standard deviation of $835. For a sample of 18 adoptions from Russia the mean cost was $12,840 with a standard deviation of $1,545. Can we conclude the mean cost is larger for adopting children from Russia? Assume the two population standard deviations are not the same. Use the .05 significance level.

22. Suppose you are an expert on the fashion industry and wish to gather information to compare the amount earned per month by models featuring Liz Claiborne attire with those of Calvin Klein. The following is the amount ($000) earned per month by a sample of Claiborne models:

$5.0	$4.5	$3.4	$3.4	$6.0	$3.3	$4.5	$4.6	$3.5	$5.2
4.8	4.4	4.6	3.6	5.0					

The following is the amount ($000) earned by a sample of Klein models.

$3.1	$3.7	$3.6	$4.0	$3.8	$3.8	$5.9	$4.9	$3.6	$3.6
2.3	4.0								

Is it reasonable to conclude that Claiborne models earn more? Use the .05 significance level and assume the population standard deviations are not the same.

Two-Sample Tests of Hypothesis: Dependent Samples

On page 381, we tested the difference between the means from two independent samples. We compared the mean time required to mount an engine using the Welles method to the time to mount the engine using the Atkins method. The samples were *independent,* meaning that the sample of assembly times using the Welles method was in no way related to the sample of assembly times using the Atkins method.

There are situations, however, in which the samples are not independent. To put it another way, the samples are *dependent* or related. As an example, Nickel Savings and Loan employs two firms, Schadek Appraisals and Bowyer Real Estate, to appraise the value of the real estate properties on which it makes loans. It is important that these two firms be similar in their appraisal values. To review the consistency of the two appraisal firms, Nickel Savings randomly selects 10 homes and has both Schadek Appraisals and Bowyer Real Estate appraise the value of the

selected homes. For each home, there will be a pair of appraisal values. That is, for each home there will be an appraised value from both Schadek Appraisals and Bowyer Real Estate. The appraised values depend on, or are related to, the home selected. This is also referred to as a **paired sample.**

For hypothesis testing, we are interested in the distribution of the *differences* in the appraised value of each home. Hence, there is only one sample. To put it more formally, we are investigating whether the mean of the distribution of differences in the appraised values is 0. The sample is made up of the *differences* between the appraised values determined by Schadek Appraisals and the values from Bowyer Real Estate. If the two appraisal firms are reporting similar estimates, then sometimes Schadek Appraisals will be the higher value and sometimes Bowyer Real Estate will have the higher value. However, the mean of the distribution of differences will be 0. On the other hand, if one of the firms consistently reports the larger appraisal values, then the mean of the distribution of the differences will not be 0.

We will use the symbol μ_d to indicate the population mean of the distribution of differences. We assume the distribution of the population of differences follows the normal distribution. The test statistic follows the t distribution and we calculate its value from the following formula:

PAIRED *t* TEST

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad [11\text{–}9]$$

There are $n - 1$ degrees of freedom and

$\bar{d}$ is the mean of the difference between the paired or related observations.
s_d is the standard deviation of the differences between the paired or related observations.
n is the number of paired observations.

The standard deviation of the differences is computed by the familiar formula for the standard deviation, except d is substituted for X. The formula is:

$$s_d = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}}$$

The following example illustrates this test.

Example

Recall that Nickel Savings and Loan wishes to compare the two companies it uses to appraise the value of residential homes. Nickel Savings selected a sample of 10 residential properties and scheduled both firms for an appraisal. The results, reported in $000, are:

Home	Schadek	Bowyer
1	235	228
2	210	205
3	231	219
4	242	240
5	205	198
6	230	223
7	231	227
8	210	215
9	225	222
10	249	245

At the .05 significance level, can we conclude there is a difference in the mean appraised values of the homes?

Solution

The first step is to state the null and the alternate hypotheses. In this case a two-tailed alternative is appropriate because we are interested in determining whether there is a *difference* in the appraised values. We are not interested in showing whether one particular firm appraises property at a higher value than the other. The question is whether the sample differences in the appraised values could have come from a population with a mean of 0. If the population mean of the differences is 0, then we conclude that there is no difference in the appraised values. The null and alternate hypotheses are:

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

There are 10 homes appraised by both firms, so $n = 10$, and $df = n - 1 = 10 - 1 = 9$. We have a two-tailed test, and the significance level is .05. To determine the critical value, go to Appendix B.2, move across the row with 9 degrees of freedom to the column for a two-tailed test and the .05 significance level. The value at the intersection is 2.262. This value appears in the box in Table 11–2. The decision rule is to reject the null hypothesis if the computed value of t is less than -2.262 or greater than 2.262. Here are the computational details.

Home	Schadek	Bowyer	Difference, d	$(d - \bar{d})$	$(d - \bar{d})^2$
1	235	228	7	2.4	5.76
2	210	205	5	0.4	0.16
3	231	219	12	7.4	54.76
4	242	240	2	−2.6	6.76
5	205	198	7	2.4	5.76
6	230	223	7	2.4	5.76
7	231	227	4	−0.6	0.36
8	210	215	−5	−9.6	92.16
9	225	222	3	−1.6	2.56
10	249	245	4	−0.6	0.36
			46	0	174.40

$$\bar{d} = \frac{\Sigma d}{n} = \frac{46}{10} = 4.60$$

$$s_d = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}} = \sqrt{\frac{174.4}{10 - 1}} = 4.402$$

Using formula (11–9), the value of the test statistic is 3.305, found by

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{4.6}{4.402/\sqrt{10}} = 3.305$$

Because the computed t falls in the rejection region, the null hypothesis is rejected. The population distribution of differences does not have a mean of 0. We conclude that there is a difference in the mean appraised values of the homes. The largest difference of $12,000 is for Home 3. Perhaps that would be an appropriate place to begin a more detailed review.

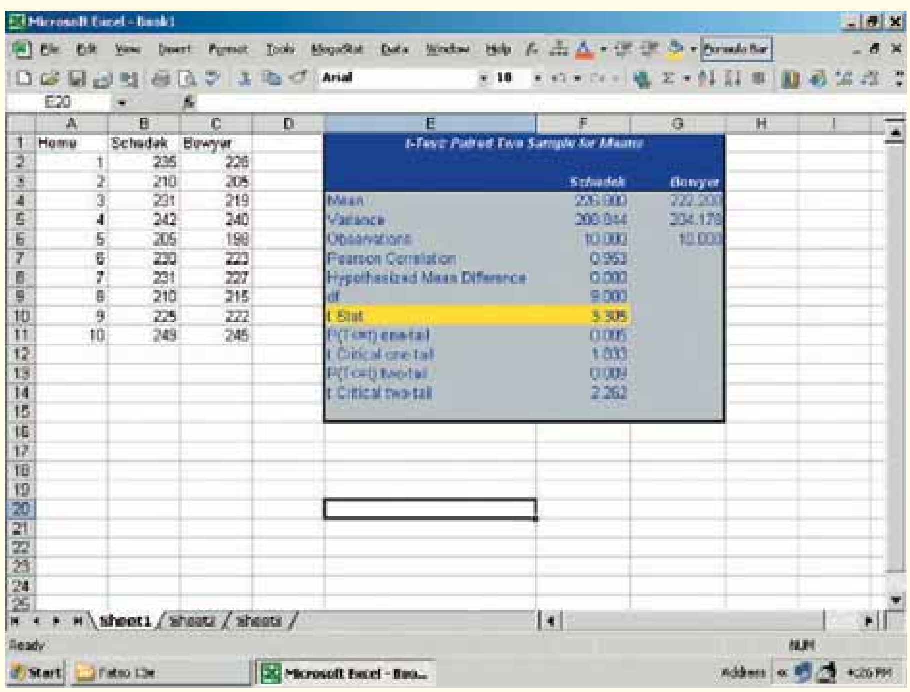
To find the p-value, we use Appendix B.2 and the section for a two-tailed test. Move along the row with 9 degrees of freedom and find the values of t that are closest to our calculated value. For a .01 significance level, the value of t is 3.250. The computed value is larger than this value, but smaller than the value of 4.781 corresponding to the .001 significance level. Hence, the p-value is less than .01. This information is highlighted in Table 11–2.

TABLE 11–2 A Portion of the *t* Distribution from Appendix B.2

	Confidence Intervals					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test					
df	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test					
	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587

Excel has a procedure called "t-Test: Paired Two-Sample for Means" that will perform the calculations of formula (11–9). The output from this procedure is given below.

The computed value of *t* is 3.305, and the two-tailed *p*-value is .009. Because the *p*-value is less than .05, we reject the hypothesis that the mean of the distribution of the differences between the appraised values is zero. In fact, this *p*-value is between .01 and .001. There is a small likelihood that the null hypothesis is true.



Excel

Comparing Dependent and Independent Samples

Beginning students are often confused by the difference between tests for independent samples [formula (11–6)] and tests for dependent samples [formula (11–9)]. How do we tell the difference between dependent and independent samples? There are two types of dependent samples: (1) those characterized by a measurement, an intervention of some type, and then another measurement; and (2) a matching or pairing of the observations. To explain further:

1. The first type of dependent sample is characterized by a measurement followed by an intervention of some kind and then another measurement. This could be called a "before" and "after" study. Two examples will help to clarify. Suppose we want to show that, by placing speakers in the production area and playing soothing music, we are able to increase production. We begin by selecting a sample of workers and measuring their output under the current conditions. The speakers are then installed in the production area, and we again measure the output of the same workers. There are two measurements, before placing the speakers in the production area and after. The intervention is placing speakers in the production area.

   A second example involves an educational firm that offers courses designed to increase test scores and reading ability. Suppose the firm wants to offer a course that will help high school juniors increase their SAT scores. To begin, each student takes the SAT in the junior year in high school. During the summer between the junior and senior year, they participate in the course that gives them tips on taking tests. Finally, during the fall of their senior year in high school, they retake the SAT. Again, the procedure is characterized by a measurement (taking the SAT as a junior), an intervention (the summer workshops), and another measurement (taking the SAT during their senior year).
2. The second type of dependent sample is characterized by matching or pairing observations. Nickel Savings in the previous example is a dependent sample of this type. It selected a property for appraisal and then had two appraisals on the same property. As a second example, suppose an industrial psychologist wishes to study the intellectual similarities of newly married couples. She selects a sample of newlyweds. Next, she administers a standard intelligence test to both the man and woman to determine the difference in the scores. Notice the matching that occurred: comparing the scores that are paired or matched by marriage.

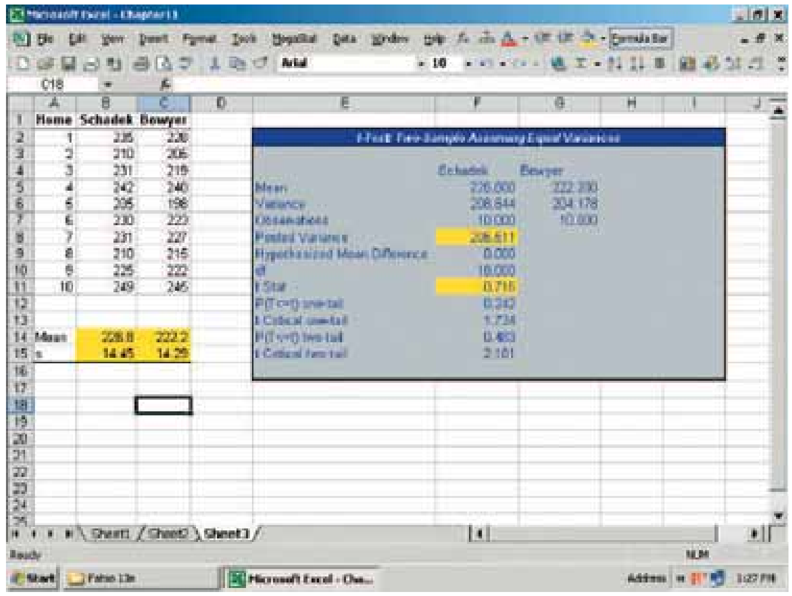
Why do we prefer dependent samples to independent samples? By using dependent samples, we are able to reduce the variation in the sampling distribution. To illustrate, we will use the Nickel Savings and Loan example just completed. Suppose we assume that we have two independent samples of real estate property for appraisal and conduct the following test of hypothesis, using formula (11–6). The null and alternate hypotheses are:

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

There are now two independent samples of 10 each. So the number of degrees of freedom is $10 + 10 - 2 = 18$. From Appendix B.2, for the .05 significance level, H_0 is rejected if t is less than -2.101 or greater than 2.101.

We use the same Excel commands as on page 95 in Chapter 3 to find the mean and the standard deviation of the two independent samples. We use the Excel commands on page 404 of this chapter to find the pooled variance and the value of the "t Stat." These values are highlighted in yellow.

The mean of the appraised value of the 10 properties by Schadek is $226,800, and the standard deviation is $14,500. For Bowyer Real Estate the mean appraised



value is \$222,200, and the standard deviation is \$14,290. To make the calculations easier, we use \$000 instead of \$. The value of the pooled estimate of the variance from formula (11–5) is

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(10 - 1)(14.45^2) + (10 - 1)(14.29)^2}{10 + 10 - 2} = 206.50$$

From formula (11–6), t is 0.716.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{226.8 - 222.2}{\sqrt{206.50\left(\frac{1}{10} + \frac{1}{10}\right)}} = \frac{4.6}{6.4265} = 0.716$$

The computed t (0.716) is less than 2.101, so the null hypothesis is not rejected. We cannot show that there is a difference in the mean appraisal value. That is not the same conclusion that we got before! Why does this happen? The numerator is the same in the paired observations test (4.6). However, the denominator is smaller. In the paired test the denominator is 1.3920 (see the calculations on page 390). In the case of the independent samples, the denominator is 6.4265. There is more variation or uncertainty. This accounts for the difference in the t values and the difference in the statistical decisions. The denominator measures the standard error of the statistic. When the samples are *not* paired, two kinds of variation are present: differences between the two appraisal firms and the difference in the value of the real estate. Properties numbered 4 and 10 have relatively high values, whereas number 5 is relatively low. These data show how different the values of the property are, but we are really interested in the difference between the two appraisal firms.

The trick is to pair the values to reduce the variation among the properties. The paired test uses only the difference between the two appraisal firms for the same property. Thus, the paired or dependent statistic focuses on the variation between Schadek Appraisals and Bowyer Real Estate. Thus, its standard error is always smaller. That, in turn, leads to a larger test statistic and a greater chance of rejecting the null hypothesis. So whenever possible you should pair the data.

There is a bit of bad news here. In the paired observations test, the degrees of freedom are half of what they are if the samples are not paired. For the real estate example, the degrees of freedom drop from 18 to 9 when the observations are paired. However, in most cases, this is a small price to pay for a better test.

Self-Review 11–5

Advertisements by Sylph Fitness Center claim that completing its course will result in losing weight. A random sample of eight recent participants showed the following weights before and after completing the course. At the .01 significance level, can we conclude the students lost weight?

Name	Before	After
Hunter	155	154
Cashman	228	207
Mervine	141	147
Massa	162	157
Creola	211	196
Peterson	164	150
Redding	184	170
Poust	172	165

(a) State the null hypothesis and the alternate hypothesis.
(b) What is the critical value of t?
(c) What is the computed value of t?
(d) Interpret the result. What is the p-value?
(e) What assumption needs to be made about the distribution of the differences?

Exercises

23. The null and alternate hypotheses are:

H_0: $\mu_d \leq 0$
H_1: $\mu_d > 0$

The following sample information shows the number of defective units produced on the day shift and the afternoon shift for a sample of four days last month.

	Day			
	1	**2**	**3**	**4**
Day shift	10	12	15	19
Afternoon shift	8	9	12	15

At the .05 significance level, can we conclude there are more defects produced on the afternoon shift?

24. The null and alternate hypotheses are:

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

The following paired observations show the number of traffic citations given for speeding by Officer Dhondt and Officer Meredith of the South Carolina Highway Patrol for the last five months.

	Day				
	May	**June**	**July**	**August**	**September**
Officer Dhondt	30	22	25	19	26
Officer Meredith	26	19	20	15	19

At the .05 significance level, is there a difference in the mean number of citations given by the two officers?

Note: Use the five-step hypothesis testing procedure to solve the following exercises.

25. The management of Discount Furniture, a chain of discount furniture stores in the Northeast, designed an incentive plan for salespeople. To evaluate this innovative plan, 12 salespeople were selected at random, and their weekly incomes before and after the plan were recorded.

Salesperson	Before	After
Sid Mahone	$320	$340
Carol Quick	290	285
Tom Jackson	421	475
Andy Jones	510	510
Jean Sloan	210	210
Jack Walker	402	500
Peg Mancuso	625	631
Anita Loma	560	560
John Cuso	360	365
Carl Utz	431	431
A. S. Kushner	506	525
Fern Lawton	505	619

Was there a significant increase in the typical salesperson's weekly income due to the innovative incentive plan? Use the .05 significance level. Estimate the *p*-value, and interpret it.

26. The federal government recently granted funds for a special program designed to reduce crime in high-crime areas. A study of the results of the program in eight high-crime areas of Miami, Florida, yielded the following results.

	Number of Crimes by Area							
	A	B	C	D	E	F	G	H
Before	14	7	4	5	17	12	8	9
After	2	7	3	6	8	13	3	5

Has there been a decrease in the number of crimes since the inauguration of the program? Use the .01 significance level. Estimate the *p*-value.

Chapter Summary

I. In comparing two population means we wish to know whether they could be equal.

- **A.** We are investigating whether the distribution of the difference between the means could have a mean of 0.
- **B.** The test statistic follows the standard normal distribution if the population standard deviations are known.
  - **1.** No assumption about the shape of either population is required.
  - **2.** The samples are from independent populations.
  - **3.** The formula to compute the value of z is

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \qquad \textbf{[11–2]}$$

II. We can also test whether two samples came from populations with an equal proportion of successes.
   A. The two sample proportions are pooled using the following formula:

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} \qquad \textbf{[11–4]}$$

   B. We compute the value of the test statistic from the following formula:

$$z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_c(1 - p_c)}{n_1} + \dfrac{p_c(1 - p_c)}{n_2}}} \qquad \textbf{[11–3]}$$

III. The test statistic to compare two means is the t distribution if the population standard deviations are not known.
   A. Both populations must follow the normal distribution.
   B. The populations must have equal standard deviations.
   C. The samples are independent.
   D. Finding the value of t requires two steps.
      1. The first step is to pool the standard deviations according to the following formula:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad \textbf{[11–5]}$$

      2. The value of t is computed from the following formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \qquad \textbf{[11–6]}$$

      3. The degrees of freedom for the test are $n_1 + n_2 - 2$.

IV. If we cannot assume the population standard deviations are equal, we
   A. Use the t distribution as the test statistic but adjust the degrees of freedom using the following formula.

$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \qquad \textbf{[11–8]}$$

   B. The value of the test statistic is computed from the following formula.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \qquad \textbf{[11–7]}$$

V. For dependent samples, we assume the distribution of the paired differences between the populations has a mean of 0.
   A. We first compute the mean and the standard deviation of the sample differences.
   B. The value of the test statistic is computed from the following formula:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad \textbf{[11–9]}$$

Pronunciation Key

SYMBOL	MEANING	PRONUNCIATION
p_c	Pooled proportion	*p sub c*
s_p^2	Pooled sample variance	*s sub p squared*
$\bar{X}_1$	Mean of the first sample	*X bar sub 1*
$\bar{X}_2$	Mean of the second sample	*X bar sub 2*
$\bar{d}$	Mean of the difference between dependent observations	*d bar*
s_d	Standard deviation of the difference between dependent observations	*s sub d*

Chapter Exercises

27. A recent study focused on the number of times men and women who live alone buy take-out dinner in a month. The information is summarized below.

Statistic	Men	Women
Sample mean	24.51	22.69
Population standard deviation	4.48	3.86
Sample size	35	40

At the .01 significance level, is there a difference in the mean number of times men and women order take-out dinners in a month? What is the *p*-value?

28. Clark Heter is an industrial engineer at Lyons Products. He would like to determine whether there are more units produced on the night shift than on the day shift. Assume the population standard deviation for the number of units produced on the day shift is 21 and is 28 on the night shift. A sample of 54 day-shift workers showed that the mean number of units produced was 345. A sample of 60 night-shift workers showed that the mean number of units produced was 351. At the .05 significance level, is the number of units produced on the night shift larger?

29. Fry Brothers Heating and Air Conditioning, Inc., employs Larry Clark and George Murnen to make service calls to repair furnaces and air conditioning units in homes. Tom Fry, the owner, would like to know whether there is a difference in the mean number of service calls they make per day. Assume the population standard deviation for Larry Clark is 1.05 calls per day and 1.23 calls per day for George Murnen. A random sample of 40 days last year showed that Larry Clark made an average of 4.77 calls per day. For a sample of 50 days George Murnen made an average of 5.02 calls per day. At the .05 significance level, is there a difference in the mean number of calls per day between the two employees? What is the *p*-value?

30. A coffee manufacturer is interested in whether the mean daily consumption of regular-coffee drinkers is less than that of decaffeinated-coffee drinkers. Assume the population standard deviation for those drinking regular coffee is 1.20 cups per day and 1.36 cups per day for those drinking decaffeinated coffee. A random sample of 50 regular-coffee drinkers showed a mean of 4.35 cups per day. A sample of 40 decaffeinated-coffee drinkers showed a mean of 5.84 cups per day. Use the .01 significance level. Compute the *p*-value.

31. A cell phone company offers two plans to its subscribers. At the time new subscribers sign up, they are asked to provide some demographic information. The mean yearly income for a sample of 40 subscribers to Plan A is $57,000 with a standard deviation of $9,200. This distribution is positively skewed; the actual coefficient of skewness is 2.11. For a sample of 30 subscribers to Plan B the mean income is $61,000 with a standard deviation of $7,100. The distribution of Plan B subscribers is also positively skewed, but not as severely. The coefficient of skewness is 1.54. At the .05 significance level, is it reasonable to conclude the mean income of those selecting Plan B is larger? What is the *p*-value? Do the coefficients of skewness affect the results of the hypothesis test? Why?

32. A computer manufacturer offers a help line that purchasers can call for help 24 hours a day, 7 days a week. Clearing these calls for help in a timely fashion is important to the company's image. After telling the caller that resolution of the problem is important the caller is asked whether the issue is software or hardware related. The mean time it takes a technician to resolve a software issue is 18 minutes with a standard deviation of 4.2 minutes. This information was obtained from a sample of 35 monitored calls. For a study of 45 hardware issues, the mean time for the technician to resolve the problem was 15.5 minutes with a standard deviation of 3.9 minutes. This information was also obtained from monitored calls. At the .05 significance level does it take longer to resolve software issues? What is the *p*-value?

33. Suppose the manufacturer of Advil, a common headache remedy, recently developed a new formulation of the drug that is claimed to be more effective. To evaluate the new drug, a sample of 200 current users is asked to try it. After a one-month trial, 180 indicated the new drug was more effective in relieving a headache. At the same time a sample of 300 current Advil users is given the current drug but told it is the new formulation. From this group, 261 said it was an improvement. At the .05 significance level can we conclude that the new drug is more effective?

34. Each month the National Association of Purchasing Managers publishes the NAPM index. One of the questions asked on the survey to purchasing agents is: Do you think the economy is expanding? Last month, of the 300 responses, 160 answered yes to the question. This month, 170 of the 290 responses indicated they felt the economy was expanding. At the .05 significance level, can we conclude that a larger proportion of the agents believe the economy is expanding this month?

35. As part of a recent survey among dual-wage-earner couples, an industrial psychologist found that 990 men out of the 1,500 surveyed believed the division of household duties was fair. A sample of 1,600 women found 970 believed the division of household duties was fair. At the .01 significance level, is it reasonable to conclude that the proportion of men who believe the division of household duties is fair is larger? What is the *p*-value?

36. There are two major Internet providers in the Colorado Springs, Colorado, area, one called HTC and the other Mountain Communications. We want to investigate whether there is a difference in the proportion of times a customer is able to access the Internet. During a one-week period, 500 calls were placed at random times throughout the day and night to HTC. A connection was made to the Internet on 450 occasions. A similar one-week study with Mountain Communications showed the Internet to be available on 352 of 400 trials. At the .01 significance level, is there a difference in the percent of time that access to the Internet is successful?

37. In a poll recently conducted at Iowa State University, 68 out of 98 male students and 45 out of 85 female students expressed "at least some support" for implementing an "exit strategy" from Iraq. Test at the .05 significance level the null hypothesis that the population proportions are equal against the two-tailed alternative.

38. A study was conducted to determine if there was a difference in the humor content in British and American trade magazine advertisements. In an independent random sample of 270 American trade magazine advertisements, 56 were humorous. An independent random sample of 203 British trade magazines contained 52 humorous ads. Does this data provide evidence at the .05 significance level that there is a difference in the proportion of humorous ads in British versus American trade magazines?

39. Harriet's Shoe Emporium operates stores in enclosed shopping malls as well as outlet malls. There are more than 1,000 stores throughout the United States and Canada. Harriet, the CEO, wishes to determine whether the number of shoes sold per week at mall stores is more than the number sold at outlet malls. She selects a sample of 22 mall and 25 outlet stores and determines the number of shoes sold in each sample store last week. The sample information is reported below.

	Sample Mean	Sample Standard Deviation	Sample Size
Mall	1,078	633	22
Outlet	908.2	369.8	25

Harriet believes there is much more variation in the number of shoes sold in the mall stores than in the outlet stores. Hence, she is not willing to assume equal population standard deviations. At the .02 significance level is it reasonable to conclude that the mean number of shoes sold at mall stores is larger than at outlet stores?

40. The manufacturers of DVD players want to test whether a small price reduction is enough to increase sales of their products. Randomly chosen data on 15 weekly sales totals at department stores in a Houston, Texas, region before the price reduction show a sample mean of $6,598 and a sample standard deviation of $844. A random sample of 12 weekly sales totals after the small price reduction gives a sample mean of $6,870 and a sample standard deviation of $669. Assume the population standard deviations are not the same. Is there evidence using a significance level of 0.05 that the small price reduction is enough to increase sales of DVD players?

41. One of the music industry's most pressing questions is: Can paid download stores contend nose-to-nose with free peer-to-peer download services? Data gathered over the last 12 months show Apple's iTunes was used by an average of 1.65 million households with a sample standard deviation of 0.56 million family units. Over the same 12 months WinMX (a no-cost P2P download service) was used by an average of 2.2 million families with a sample standard deviation of 0.30 million. Assume the population standard deviations are not the same. Using a significance level of 0.05, test the hypothesis of no difference in the mean number of households picking either variety of service to download songs.

42. Businesses, particularly those in the food preparation industry such as General Mills, Kellogg, and Betty Crocker regularly use coupons as a brand allegiance builder to stimulate their retailing. There is uneasiness that the users of paper coupons are different from the users of e-coupons (coupons disseminated by means of the Internet). One survey recorded the age of each person who redeemed a coupon along with the type (either electronic or paper). The sample of 35 e-coupon users had a mean age of 33.6 years with a standard deviation of 10.9, while a similar sample of 25 traditional paper-coupon clippers had a mean age of 39.5 with a standard deviation of 4.8. Assume the population standard deviations are not the same. Using a significance level of 0.01, test the hypothesis of no difference in the mean ages of the two groups of coupon clients.

43. The owner of Bun 'N' Run Hamburgers wishes to compare the sales per day at two locations. The mean number sold for 10 randomly selected days at the Northside site was 83.55, and the standard deviation was 10.50. For a random sample of 12 days at the Southside location, the mean number sold was 78.80 and the standard deviation was 14.25. At the .05 significance level, is there a difference in the mean number of hamburgers sold at the two locations? What is the *p*-value?

44. The Engineering Department at Sims Software, Inc., recently developed two chemical solutions designed to increase the usable life of computer disks. A sample of disks treated with the first solution lasted 86, 78, 66, 83, 84, 81, 84, 109, 65, and 102 hours. Those treated with the second solution lasted 91, 71, 75, 76, 87, 79, 73, 76, 79, 78, 87, 90, 76, and 72 hours. Assume the population standard deviations are not the same. At the .10 significance level, can we conclude that there is a difference in the length of time the two types of treatment lasted?

45. The Willow Run Outlet Mall has two Haggar Outlet Stores, one located on Peach Street and the other on Plum Street. The two stores are laid out differently, but both store managers claim their layout maximizes the amounts customers will purchase on impulse. A sample of 10 customers at the Peach Street store revealed they spent the following amounts more than planned: $17.58, $19.73, $12.61, $17.79, $16.22, $15.82, $15.40, $15.86, $11.82, and $15.85. A sample of 14 customers at the Plum Street store revealed they spent the following amounts more than they planned: $18.19, $20.22, $17.38, $17.96, $23.92, $15.87, $16.47, $15.96, $16.79, $16.74, $21.40, $20.57, $19.79, and $14.83. At the .01 significance level, is there a difference in the mean amounts purchased on impulse at the two stores?

46. Grand Strand Family Medical Center is specifically set up to treat minor medical emergencies for visitors to the Myrtle Beach area. There are two facilities, one in the Little River Area and the other in Murrells Inlet. The Quality Assurance Department wishes to compare the mean waiting time for patients at the two locations. Samples of the waiting times, reported in minutes, follow:

Location	Waiting Time											
Little River	31.73	28.77	29.53	22.08	29.47	18.60	32.94	25.18	29.82	26.49		
Murrells Inlet	22.93	23.92	26.92	27.20	26.44	25.62	30.61	29.44	23.09	23.10	26.69	22.31

Assume the population standard deviations are not the same. At the .05 significance level, is there a difference in the mean waiting time?

47. Commercial Bank and Trust Company is studying the use of its automatic teller machines (ATMs). Of particular interest is whether young adults (under 25 years) use the machines more than senior citizens. To investigate further, samples of customers under 25 years of age and customers over 60 years of age were selected. The number of ATM transactions last month was determined for each selected individual, and the results are shown below. At the .01 significance level, can bank management conclude that younger customers use the ATMs more?

Under 25	10	10	11	15	7	11	10	9			
Over 60	4	8	7	7	4	5	1	7	4	10	5

48. Two boats, the *Prada* (Italy) and the *Oracle* (U.S.A.), are competing for a spot in the upcoming *America's Cup* race. They race over a part of the course several times. Below are the sample times in minutes. Assume the population standard deviations are not the same. At the .05 significance level, can we conclude that there is a difference in their mean times?

Boat	Time (minutes)											
Prada (Italy)	12.9	12.5	11.0	13.3	11.2	11.4	11.6	12.3	14.2	11.3		
Oracle (U.S.A.)	14.1	14.1	14.2	17.4	15.8	16.7	16.1	13.3	13.4	13.6	10.8	19.0

49. The manufacturer of an MP3 player wanted to know whether a 10 percent reduction in price is enough to increase the sales of its product. To investigate, the owner randomly selected eight outlets and sold the MP3 player at the reduced price. At seven randomly selected outlets, the MP3 player was sold at the regular price. Reported below is the number of units sold last month at the sampled outlets. At the .01 significance level, can the manufacturer conclude that the price reduction resulted in an increase in sales?

Regular price	138	121	88	115	141	125	96	
Reduced price	128	134	152	135	114	106	112	120

50. A number of minor automobile accidents occur at various high-risk intersections in Teton County despite traffic lights. The Traffic Department claims that a modification in the type of light will reduce these accidents. The county commissioners have agreed to a proposed experiment. Eight intersections were chosen at random, and the lights at those intersections were modified. The numbers of minor accidents during a six-month period before and after the modifications were:

	Number of Accidents							
	A	B	C	D	E	F	G	H
Before modification	5	7	6	4	8	9	8	10
After modification	3	7	7	0	4	6	8	2

At the .01 significance level is it reasonable to conclude that the modification reduced the number of traffic accidents?

51. Lester Hollar is vice president for human resources for a large manufacturing company. In recent years he has noticed an increase in absenteeism that he thinks is related to the general health of the employees. Four years ago, in an attempt to improve the situation, he began a fitness program in which employees exercise during their lunch hour. To evaluate the program, he selected a random sample of eight participants and found the number of days each was absent in the six months before the exercise program began and in the last six months. Below are the results. At the .05 significance level, can he conclude that the number of absences has declined? Estimate the *p*-value.

Employee	Before	After
1	6	5
2	6	2
3	7	1
4	7	3
5	4	3
6	3	6
7	5	3
8	6	7

52. The president of the American Insurance Institute wants to compare the yearly costs of auto insurance offered by two leading companies. He selects a sample of 15 families, some with only a single insured driver, others with several teenage drivers, and pays each family a stipend to contact the two companies and ask for a price quote. To make the data comparable, certain features, such as the deductible amount and limits of liability, are standardized. The sample information is reported below. At the .10 significance level, can we conclude that there is a difference in the amounts quoted?

Family	Progressive Car Insurance	GEICO Mutual Insurance
Becker	$2,090	$1,610
Berry	1,683	1,247
Cobb	1,402	2,327
Debuck	1,830	1,367
DuBrul	930	1,461
Eckroate	697	1,789
German	1,741	1,621
Glasson	1,129	1,914
King	1,018	1,956
Kucic	1,881	1,772
Meredith	1,571	1,375
Obeid	874	1,527
Price	1,579	1,767
Phillips	1,577	1,636
Tresize	860	1,188

53. Fairfield Homes is developing two parcels near Pigeon Fork, Tennessee. In order to test different advertising approaches, it uses different media to reach potential buyers. The mean annual family income for 15 people making inquiries at the first development is $150,000, with a standard deviation of $40,000. A corresponding sample of 25 people at the second development had a mean of $180,000, with a standard deviation of $30,000. Assume the population standard deviations are the same. At the .05 significance level, can Fairfield conclude that the population means are different?

54. The following data resulted from a taste test of two different chocolate bars. The first number is a rating of the taste, which could range from 0 to 5, with a 5 indicating the person liked the taste. The second number indicates whether a "secret ingredient" was present. If the ingredient was present a code of 1 was used and a 0 otherwise. Assume the population standard deviations are the same. At the .05 significance level, do these data show a difference in the taste ratings?

Rating	With/Without	Rating	With/Without
3	1	1	1
1	1	4	0
0	0	4	0
2	1	2	1
3	1	3	0
1	1	4	0

55. An investigation of the effectiveness of an antibacterial soap in reducing operating room contamination resulted in the accompanying table. The new soap was tested in a sample of eight operating rooms in the greater Seattle area during the last year.

	Operating Room							
	A	B	C	D	E	F	G	H
Before	6.6	6.5	9.0	10.3	11.2	8.1	6.3	11.6
After	6.8	2.4	7.4	8.5	8.1	6.1	3.4	2.0

At the 0.05 significance level, can we conclude the contamination measurements are lower after use of the new soap?

56. The following data on annual rates of return were collected from five stocks listed on the New York Stock Exchange ("the big board") and five stocks listed on NASDAQ. Assume the population standard deviations are the same. At the .10 significance level, can we conclude that the annual rates of return are higher on the big board?

NYSE	NASDAQ
17.16	15.80
17.08	16.28
15.51	16.21
8.43	17.97
25.15	7.77

57. The city of Laguna Beach operates two public parking lots. The one on Ocean Drive can accommodate up to 125 cars and the one on Rio Rancho can accommodate up to 130 cars. City planners are considering both increasing the size of the lots and changing the fee structure. To begin, the Planning Office would like some information on the number of cars in the lots at various times of the day. A junior planner officer is assigned the task of visiting the two lots at random times of the day and evening and counting the number of cars in the lots. The study lasted over a period of one month. Below is the number of cars in the lots for 25 visits of the Ocean Drive lot and 28 visits of the Rio Rancho lot. Assume the population standard deviations are equal.

Ocean Drive												
89	115	93	79	113	77	51	75	118	105	106	91	54
63	121	53	81	115	67	53	69	95	121	88	64	
Rio Rancho												
128	110	81	126	82	114	93	40	94	45	84	71	74
92	66	69	100	114	113	107	62	77	80	107	90	129
105	124											

Is it reasonable to conclude that there is a difference in the mean number of cars in the two lots? Use the .05 significance level.

58. The amount of income spent on housing is an important component of the cost of living. The total costs of housing for homeowners might include mortgage payments, property taxes, and utility costs (water, heat, electricity). An economist selected a sample of 20 homeowners in New England and then calculated these total housing costs as a percent of monthly income, five years ago and now. The information is reported below. Is it reasonable to conclude the percent is less now than five years ago?

Homeowner	Five Years Ago	Now	Homeowner	Five Years Ago	Now
1	17%	10%	11	35%	32%
2	20	39	12	16	32
3	29	37	13	23	21
4	43	27	14	33	12
5	36	12	15	44	40
6	43	41	16	44	42
7	45	24	17	28	22
8	19	26	18	29	19
9	49	28	19	39	35
10	49	26	20	22	12

exercises.com

59. Listed below are several prominent companies and their stock prices in August 2005. Go to the Web and look up today's price. There are many sources to find stock prices, such as Yahoo and CNNFI. The Yahoo address is http://finance.yahoo.com. Enter the symbol identification to find the current price. At the .05 significance level, can we conclude that the prices have changed?

Company	Symbol	Price
Coca-Cola	KO	$43.99
Walt Disney	DIS	25.56
Eastman Kodak	EK	26.56
Ford Motor Company	F	10.92
General Motors	GM	37.01
Goodyear Tire	GT	17.41
IBM	IBM	83.74
McDonald's	MCD	31.24
The McGraw-Hill Companies	MHP	46.46
Oracle	ORCL	13.58
Johnson & Johnson	JNJ	64.62
General Electric	GE	34.40
Home Depot	HD	42.80

60. The *USA Today* website (http://www.usatoday.com/sports/baseball/salaries/default.aspx) reports information on individual player salaries. Go to the site and find the individual salaries for your favorite team in the American League and in the National League. Compute the mean and the standard deviation for each. Is it reasonable to conclude that there is a difference in the salaries of the two teams?

Data Set Exercises

61. Refer to the Real Estate data, which report information on the homes sold in Denver, Colorado, last year.
- **a.** At the .05 significance level, can we conclude that there is a difference in the mean selling price of homes with a pool and homes without a pool?
- **b.** At the .05 significance level, can we conclude that there is a difference in the mean selling price of homes with an attached garage and homes without an attached garage?
- **c.** At the .05 significance level, can we conclude that there is a difference in the mean selling price of homes in Township 1 and Township 2?
- **d.** Find the median selling price of the homes. Divide the homes into two groups, those that sold for more than (or equal to) the median price and those that sold for less. Is there a difference in the proportion of homes with a pool for those that sold at or above the median price versus those that sold for less than the median price? Use the .05 significance level.

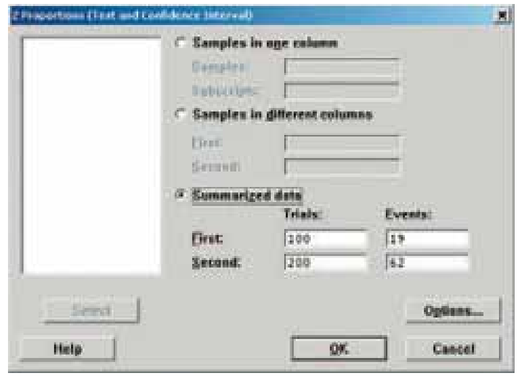
62. Refer to the Baseball 2005 data, which report information on the 30 Major League Baseball teams for the 2005 season.
- **a.** At the .05 significance level, can we conclude that there is a difference in the mean salary of teams in the American League versus teams in the National League?
- **b.** At the .05 significance level, can we conclude that there is a difference in the mean home attendance of teams in the American League versus teams in the National League?
- **c.** Compute the mean and the standard deviation of the number of wins for the 10 teams with the highest salaries. Do the same for the 10 teams with the lowest salaries. At the .05 significance level is there a difference in the mean number of wins for the two groups?

63. Refer to the Wage data, which report information on annual wages for a sample of 100 workers. Also included are variables relating to industry, years of education, and gender for each worker.
- **a.** Conduct a test of hypothesis to determine if there is a difference in the mean annual wages of Southern residents versus non-Southern residents.
- **b.** Conduct a test of hypothesis to determine if there is a difference in the mean annual wages of white and non-white wage earners.
- **c.** Conduct a test of hypothesis to determine if there is a difference in the mean annual wages of Hispanic and non-Hispanic wage earners.
- **d.** Conduct a test of hypothesis to determine if there is a difference in the mean annual wages of female and male wage earners.
- **e.** Conduct a test of hypothesis to determine if there is a difference in the mean annual wages of married and nonmarried wage earners.

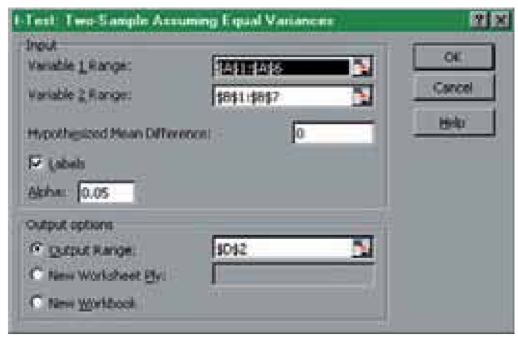
64. Refer to the CIA data, which report demographic and economic information on 46 countries. Conduct a test of hypothesis to determine whether the mean percent of the population over 65 years of age in G-20 countries is different from those that are not G-20 members.

Software Commands

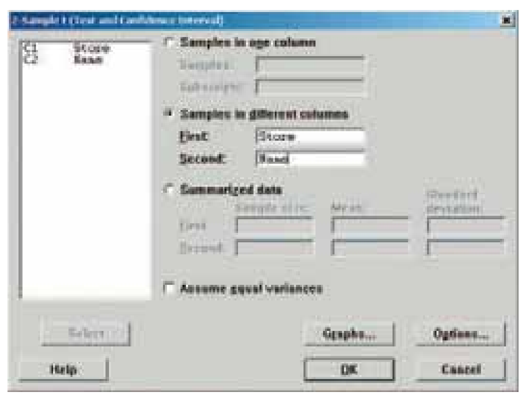
1. The MINITAB commands for the two-sample test of proportions on page 378 are:
   a. From the toolbar, select **Stat, Basic Statistics,** and then **2 Proportions.**
   b. In the next dialog box select **Summarized data,** in the row labeled **First** enter *100* for **Trials** and *19* for **Events.** In the row labeled **Second** put *200* for **Trials** and *62* for **Events,** then click **OK.**



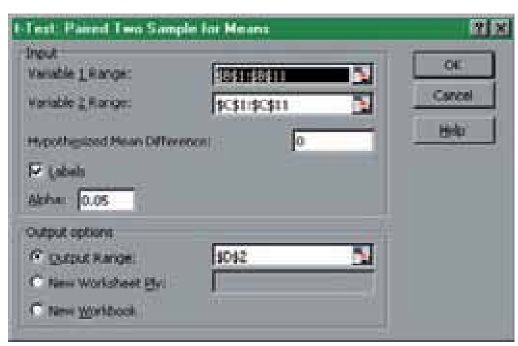
2. The Excel commands for the two-sample *t*-test on page 383 are:
   a. Enter the data into columns A and B (or any other columns) in the spreadsheet. Use the first row of each column to enter the variable name.
   b. From the menu bar select **Tools** and **Data Analysis.** Select **t-Test: Two-Sample Assuming Equal Variances,** then click **OK.**
   c. In the dialog box indicate that the range of **Variable 1** is from *A1* to *A6* and **Variable 2** from *B1* to *B7,* the **Hypothesized Mean Difference** is 0, click **Labels, Alpha** is *0.05,* and the **Output Range** is *D2*. Click **OK.**



3. The MINITAB commands for the two-sample test of proportions on page 387 are:
   a. Put the amount absorbed by the Store brand in *C1* and the amount absorbed by the Name brand paper towel in *C2*.
   b. From the toolbar select **Stat, Basic Statistics,** and then **2-Sample** and click.
   c. In the next dialog box select **Samples in different columns,** select C1 Store for the **First** column and C2 Name of the **Second** and click OK.



4. The Excel commands for the paired *t*-test on page 391 are:
   a. Enter the data into columns B and C (or any other two columns) in the spreadsheet, with the variable names in the first row.
   b. From the menu bar select **Tools** and **Data Analysis.** Select **t-Test: Paired Two Sample for Means,** then click **OK.**
   c. In the dialog box indicate that the range of **Variable 1** is from *B1* to *B11* and **Variable 2** from *C1* to *C11,* the **Hypothesized Mean Difference** is *0,* click **Labels, Alpha** is *.05,* and the **Output Range** is *D2*. Click **OK.**



Chapter 11 Answers to Self-Review

11–1 **a.** H_0: $\mu_W \leq \mu_M$
H_1: $\mu_W > \mu_M$
The subscript W refers to the women and M to the men.

b. Reject H_0 if $z > 1.65$

c. $z = \dfrac{\$1{,}500 - \$1{,}400}{\sqrt{\dfrac{(\$250)^2}{50} + \dfrac{(\$200)^2}{40}}} = 2.11$

d. Reject the null hypothesis

e. p-value $= .5000 - .4826 = .0174$

f. The mean amount sold per day is larger for women.

11–2 **a.** H_0: $\pi_1 = \pi_2$
H_1: $\pi_1 \neq \pi_2$

b. .10

c. Two-tailed

d. Reject H_0 if z is less than -1.65 or greater than 1.65.

e. $p_c = \dfrac{87 + 123}{150 + 200} = \dfrac{210}{350} = .60$

$p_1 = \dfrac{87}{150} = .58 \qquad p_2 = \dfrac{123}{200} = .615$

$$z = \frac{.58 - .615}{\sqrt{\dfrac{.60(.40)}{150} + \dfrac{.60(.40)}{200}}} = -0.66$$

f. Do not reject H_0.

g. p-value $= 2(.5000 - .2454) = .5092$
There is no difference in the proportion of adults and children that liked the proposed flavor.

11–3 **a.** H_0: $\mu_d = \mu_a$
H_1: $\mu_d \neq \mu_a$

b. $df = 6 + 8 - 2 = 12$
Reject H_0 if t is less than -2.179 or t is greater than 2.179.

c. $\bar{X}_1 = \dfrac{42}{6} = 7.00 \quad s_1 = \sqrt{\dfrac{10}{6 - 1}} = 1.4142$

$\bar{X}_2 = \dfrac{80}{8} = 10.00 \quad s_2 = \sqrt{\dfrac{36}{8 - 1}} = 2.2678$

$$s_p^2 = \frac{(6 - 1)(1.4142)^2 + (8 - 1)(2.2678)^2}{6 + 8 - 2}$$
$$= 3.8333$$

$$t = \frac{7.00 - 10.00}{\sqrt{3.8333\left(\dfrac{1}{6} + \dfrac{1}{8}\right)}} = -2.837$$

d. Reject H_0 because -2.837 is less than the critical value.

e. The p-value is less than .02.

f. The mean number of defects is not the same on the two shifts.

g. Independent populations, populations follow the normal distribution, populations have equal standard deviations.

11–4 **a.** H_0: $\mu_c \geq \mu_a \qquad H_1$: $\mu_c < \mu_a$

b. $df = \dfrac{[(356^2/10) + (857^2/8)]^2}{\dfrac{(356^2/10)^2}{10 - 1} + \dfrac{(857^2/8)^2}{8 - 1}} = 8.93$
so $df = 8$

c. Reject H_0 if $t < -1.860$

d. $t = \dfrac{\$1{,}568 - \$1{,}967}{\sqrt{\dfrac{356^2}{10} + \dfrac{857^2}{8}}} = \dfrac{-399.00}{323.23} = -1.234$

e. Do not reject H_0.

f. There is no difference in the mean account balance of those who applied for their card or were contacted by a telemarketer.

11–5 **a.** H_0: $\mu_d \leq 0$, H_1: $\mu_d > 0$.

b. Reject H_0 if $t > 2.998$

c.

Name	Before	After	d	$(d - \bar{d})$	$(d - d)^2$
Hunter	155	154	1	−7.875	62.0156
Cashman	228	207	21	12.125	147.0156
Mervine	141	147	−6	−14.875	221.2656
Massa	162	157	5	−3.875	15.0156
Creola	211	196	15	6.125	37.5156
Peterson	164	150	14	5.125	26.2656
Redding	184	170	14	5.125	26.2656
Poust	172	165	7	−1.875	3.5156
			71		538.8750

$$\bar{d} = \frac{71}{8} = 8.875$$

$$s_d = \sqrt{\frac{538.875}{8 - 1}} = 8.774$$

$$t = \frac{8.875}{8.774/\sqrt{8}} = 2.861$$

d. Do not reject H_0. We cannot conclude that the students lost weight. The p-value is less than .025 but larger than .01.

e. The distribution of the differences must follow a normal distribution.