INTERNATIONAL
EDITION

Statistics, Data Analysis, and Decision Modeling

FIFTH EDITION

James R. Evans

ALWAYS LEARNING

PEARSON

Useful Statistical Functions in Excel 2010	*Description*
AVERAGE(*data range*)	Computes the average value (arithmetic mean) of a set of data.
BINOM.DIST(*number_s, trials, probability_s, cumulative*)	Returns the individual term binomial distribution.
BINOM.INV(*trials, probability_s, alpha)*	Returns the smallest value for which the cumulative binomial distribution is greater than or equal to a criterion value.
CHISQ.DIST(*x, deg_freedom, cumulative*)	Returns the left-tailed probability of the chi-square distribution.
CHISQ.DIST.RT(*x, deg_freedom, cumulative*)	Returns the right-tailed probability of the chi-square distribution.
CHISQ.TEST(*actual_range, expected_range*)	Returns the test for independence; the value of the chi-square distribution and the appropriate degrees of freedom.
CONFIDENCE.NORM(*alpha, standard_dev, size*)	Returns the confidence interval for a population mean using a normal distribution.
CONFIDENCE.T(*alpha, standard_dev, size*)	Returns the confidence interval for a population mean using a *t*-distribution.
CORREL(*arrayl, array2*)	Computes the correlation coefficient between two data sets.
EXPON.DIST(*x, lambda, cumulative*)	Returns the exponential distribution.
F.DIST(*x. deg_freedom1, deg_freedom2, cumulative*)	Returns the left-tailed *F*-probability distribution value.
F.DIST.RT(*x. deg_freedom1, deg_freedom2, cumulative*)	Returns the left-tailed *F*-probability distribution value.
FORECAST(*x, known_y's, known_x's*)	Calculates a future value along a linear trend.
GROWTH(*known_y's, known_x's, new_x's, constant)*	Calculates predicted exponential growth.
LINEST*(known_y's, known_x's, new_x's, constant, stats*)	Returns an array that describes a straight line that best fits the data.
LOGNORM.DIST(*x, mean, standard_deviation*)	Returns the cumulative lognormal distribution of *x*, where ln (*x*) is normally distributed with parameters mean and standard deviation.
MEDIAN(*data range*)	Computes the median (middle value) of a set of data.
MODE.MULT(*data range*)	Computes the modes (most frequently occurring values) of a set of data.
MODE.SNGL(*data range*)	Computes the mode of a set of data.
NORM.DIST(*x, mean, standard_dev, cumulative*)	Returns the normal cumulative distribution for the specified mean and standard deviation.
NORM.INV(*probability, mean, standard_dev*)	Returns the inverse of the cumulative normal distribution.
NORM.S.DIST(*z*)	Returns the standard normal cumulative distribution (mean = 0, standard deviation = 1).
NORM.S.INV(*probability*)	Returns the inverse of the standard normal distribution.
PERCENTILE.EXC(*array, k*)	Computes the *k*th percentile of data in a range, exclusive.
PERCENTILE.INC(*array, k*)	Computes the *k*th percentile of data in a range, inclusive.
POISSON.DIST(*x, mean, cumulative*)	Returns the Poisson distribution.
QUARTILE(*array, quart*)	Computes the quartile of a distribution.
SKEW(*data range*)	Computes the skewness, a measure of the degree to which a distribution is not symmetric around its mean.
STANDARDIZE(*x, mean, standard_deviation*)	Returns a normalized value for a distribution characterized by a mean and standard deviation.
STDEV.S(*data range*)	Computes the standard deviation of a set of data, assumed to be a sample.
STDEV.P(*data range*)	Computes the standard deviation of a set of data, assumed to be an entire population.
TREND(*known_y's, known_x's, new_x's, constant*)	Returns values along a linear trend line.
T.DIST(*x, deg_freedom, cumulative*)	Returns the left-tailed *t*-distribution value.
T.DIST.2T(*x, deg_freedom*)	Returns the two-tailed *t*-distribution value.
T.DIST.RT(*x, deg_freedom*)	Returns the right-tailed *t*-distribution.
T.INV(*probability, deg_freedom*)	Returns the left-tailed inverse of the *t*-distribution.
T.INV.2T(*probability, deg_freedom*)	Returns the two-tailed inverse of the *t*-distribution.
T.TEST(*arrayl, array2, tails, type*)	Returns the probability associated with a *t*-test.
VAR.S(*data range*)	Computes the variance of a set of data, assumed to be a sample.
VAR.P(*data range*)	Computes the variance of a set of data, assumed to be an entire population.
Z.TEST(*array, x, sigma*)	Returns the two-tailed *p*-value of a *z*-test.

3. The null hypothesis is actually true, but the hypothesis test incorrectly rejects it (called **Type I error**).
4. The null hypothesis is actually false, but the hypothesis test incorrectly fails to reject it (called **Type II error**).

The probability of making a Type I error, P(Rejecting $H_0 | H_0$ is true), is generally denoted by α and is called the **level of significance** of the test. This probability is essentially the risk that you can afford to take in making the incorrect conclusion that the alternative hypothesis is true when in fact the null hypothesis is true. The **confidence coefficient** is $1 - \alpha$, which is the probability of *correctly failing to reject* the null hypothesis, or P(Not rejecting $H_0 | H_0$ is true). For a confidence coefficient of 0.95, we mean that we expect 95 out of 100 samples to support the null hypothesis rather than the alternate hypothesis. Commonly used levels for α are 0.10, 0.05, and 0.01, resulting in confidence levels of 0.90, 0.95, and 0.99, respectively.

The probability of a Type II error, P(Not rejecting $H_0 | H_0$ is false), is denoted by β. Unlike α, this cannot be specified in advance but depends on the true value of the (unknown) population parameter. To see this, consider the hypotheses in the customer survey example:

$$H_0\text{: } \textit{mean response time} \geq 25 \textit{ minutes}$$
$$H_1\text{: } \textit{mean response time} < 25 \textit{ minutes}$$

If the true mean response from which the sample is drawn is, say, 15 minutes, we would expect to have a much smaller probability of incorrectly concluding that the null hypothesis is true than when the true mean response is 24 minutes, for example. In the first case, the sample mean would very likely be much less than 25, leading us to reject H_0. If the true mean is 24, however, even though the true mean response time is less than 25 minutes, we would have a much higher probability of failing to reject H_0 because a higher likelihood exists that the sample mean would be greater than 25 due to sampling error. Thus, the farther away the true mean response time is from the hypothesized value, the smaller is β. Generally, as α decreases, β increases, so the decision maker must consider the trade-offs of these risks.

The value $1 - \beta$ is called the **power of the test** and represents the probability of *correctly rejecting* the null hypothesis when it is indeed false, or P(Rejecting $H_0 | H_0$ is false). We would like the power of the test to be high to allow us to make a valid conclusion. The power of the test is sensitive to the sample size; small sample sizes generally result in a low value of $1 - \beta$. The power of the test can be increased by taking larger samples, which enable us to detect small differences between the sample statistics and population parameters with more accuracy. However, a larger sample size incurs higher costs, giving more meaning to the adage, "There is no such thing as a free lunch." Table 5.2 summarizes this discussion.

Decision Rules

The decision to reject or fail to reject a null hypothesis is based on computing a test statistic from sample data that is a function of the population parameter of interest and comparing it to a critical value from the hypothesized sampling distribution of the test

TABLE 5.2 Error Types in Hypothesis Testing

	Test Rejects H_0	Test Fails to Reject H_0
Alternative hypothesis (H_1) is true	Correct	Type II error (β)
Null hypothesis (H_0) is true	Type I error (α)	Correct

statistic. The sampling distribution is usually the normal distribution, t-distribution, or some other well-known distribution. The sampling distribution is divided into two parts, a *rejection region* and a *nonrejection region*. If the null hypothesis is false, it is more likely that the test statistic will fall into the rejection region. If it does, we reject the null hypothesis; otherwise, we fail to reject it. The rejection region is chosen so that the probability of the test statistic falling into it if H_0 is true is the probability of a Type I error, α.

The rejection region generally occurs in the tails of the sampling distribution of the test statistic and depends on the structure of the hypothesis test (Table 5.1). If the null hypothesis is structured as "$=$" and the alternative hypothesis as "$\neq$," then we would reject H_0 if the test statistic is *either* significantly high or low. In this case, the rejection region will occur in *both* the upper and lower tails of the distribution [see Figure 5.1(a)]. This is called a **two-tailed test of hypothesis**. Because the probability that the test statistic falls into the rejection region, given that H_0 is true, the combined area of both tails must be α. Usually, each tail has an area of $\alpha/2$.

The other types of hypothesis, which specifies a direction of relationship (where H_0 is either "$\geq$" or " $\leq$," are called **one-tailed tests of hypothesis.** In this case, the rejection region occurs only in one tail of the distribution [see Figure 5.1(b)].Determining the correct tail of the distribution to use as the rejection region for a one-tailed test is easy. If H_1 is stated as "$<$, " the rejection region is in the lower tail; if H_1 is stated as "$>$, " the rejection region is in the upper tail (just think of the inequality as an arrow pointing to the proper tail direction!).

The rejection region is defined by a *critical value* (see Figure 5.1), which is the value that divides the rejection region from the rest of the distribution. Two-tailed tests have both upper and lower critical values, while one-tailed tests have either a lower or upper

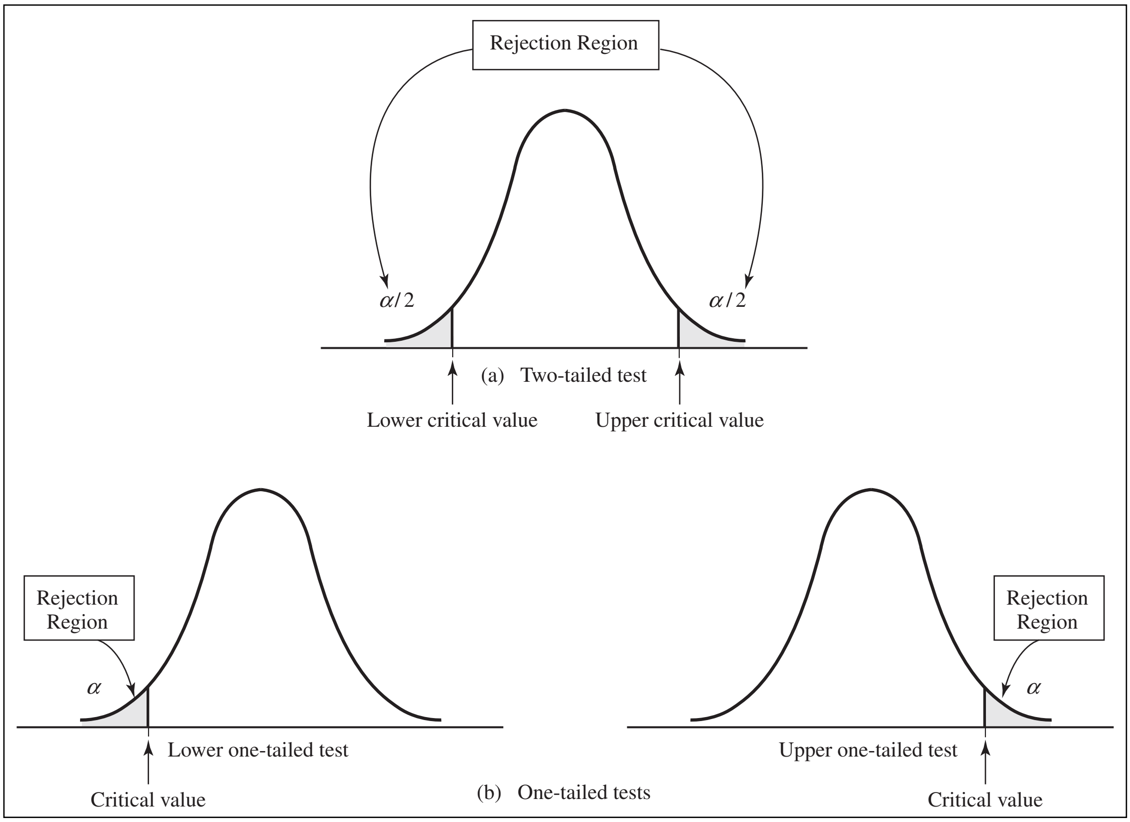


FIGURE 5.1 Illustration of Rejection Regions in Hypothesis Testing

TABLE 5.3 Common Types of Hypothesis Tests

Type of Test	Excel/*PHStat* Procedure
One-sample test for the mean, σ known	*PHStat:* One-Sample Test—*Z*-test for the Mean, Sigma Known
One-sample test for the mean, σ unknown	*PHStat:* One-Sample Test—*t*-test for the Mean, Sigma Unknown
One-sample test for the proportion	*PHStat:* One-Sample Test—*Z*-test for the Proportion
One-sample test for the variance	*PHStat:* One-Sample Test—Chi-Square Test for the Variance
Two-sample test for means, σ^2 known	Excel *Z*-test: Two-Sample for Means *PHStat:* Two-Sample Tests—*Z*-test for Differences in Two Means
Two-sample test for means, σ^2 unknown, unequal	Excel *t*-test: Two-Sample Assuming Unequal Variances *PHStat:* Two-Sample Tests—Separate-Variance *t*-test
Two-sample test for means, σ^2 unknown, assumed equal	Excel *t*-test: Two-Sample Assuming Equal Variances *PHStat:* Two Sample Tests—Pooled Variance *t*-test
Paired two-sample test for means	Excel *t*-test: Paired Two-Sample for Means
Two-sample test for proportions	*PHStat:* Two-Sample Tests—*Z*-test for Differences in Two Proportions
Two-sample test for equality of variances	Excel *F*-test Two-Sample for Variances *PHStat:* Two-Sample Tests—*F*-Test for Differences in Two Variances

TABLE 5.4 Additional Excel Support for Hypothesis Testing and Statistical Inference

Excel 2010 Function	Description
CHISQ.TEST(*actual_range, expected_range*)	Returns the test for independence, the value of the chi-square distribution, and the appropriate degrees of freedom
T.TEST(*array1, array2, tails, type*)	Returns the probability associated with a *t*-test
T.DIST(*x, deg_freedom, cumulative*)	Returns the left-tailed *t*-distribution
T.DIST.2T(*x, deg_freedom*)	Returns the two-tailed *t*-distribution
T.DIST.RT(*x, deg_freedom*)	Returns the right-tailed *t*-distribution
Z.TEST(*array, x, sigma*)	Returns the two-tailed *p*-value of a *Z*-test
F.TEST(*array1,array2*)	Returns the result of an *F*-test, the two-tailed probability that the variances in Array1 and Array2 are not significantly different
Analysis Toolpak Tools	**Description**
ANOVA: Single Factor	Tests hypothesis that means of two or more samples measured on one factor are equal
***PHStat* Add-In**	**Description**
One Sample Tests	Hypothesis tests for the mean, proportion, and variance
Two Sample Tests	Hypothesis tests for differences in means, proportions, and variances
Multiple-Sample Tests/One-Way ANOVA	Tests hypothesis that means of two or more samples measured on one factor are equal
Multiple-Sample Tests/Chi-Square Test	Performs chi-square test of independence

critical value. For standard normal and *t*-distributions, which have a mean of 0, lower-tail critical values are negative; upper-tail critical values are positive.

Critical values make it easy to determine whether or not the test statistic falls in the rejection region of the proper sampling distribution. For example, for an upper one-tailed test, if the test statistic is greater than the critical value, the decision would be to reject the null hypothesis. Similarly, for a lower one-tailed test, if the test statistic is less than the critical value, we would reject the null hypothesis. For a two-tailed test, if the test statistic is *either* greater than the upper critical value or less than the lower critical value, the decision would be to reject the null hypothesis.

Spreadsheet Support for Hypothesis Testing

Both Excel and *PHStat* have numerous tools for conducting hypothesis tests. In some cases, only one procedure is available; in others, both Excel and *PHStat* provide tools. Table 5.3 summarizes some of the more common types of hypothesis tests and Excel and *PHStat* tools for conducting these tests. Each test can be applied to different forms of the hypotheses (that is, when the null hypothesis is stated as “≥, ” “≤, ” or “= ”). The challenge is to identify the proper test statistic and decision rule and to understand the information provided in the output. Table 5.4 provides a summary of additional spreadsheet support for hypothesis testing. We will illustrate these tests through examples in the next several sections. Many other types of hypothesis tests exist, some of which are described in the appendix to this chapter; however, spreadsheet-based procedures are not available.

ONE-SAMPLE HYPOTHESIS TESTS

In this section, we will discuss several hypothesis tests for means, proportions, and variances involving a single sample.

One-Sample Tests for Means

We will first consider one-sample tests for means. The appropriate sampling distribution and test statistic depends on whether the population standard deviation is known or unknown. If the population standard deviation is known, then the sampling distribution of the mean is normal; if not, we use a *t*-distribution, which was introduced in Chapter 4 when discussing confidence intervals for the mean with an unknown population standard deviation. In most practical applications, the population standard deviation will not be known but is estimated from the sample, so we will only illustrate this case.

For the *Customer Support Survey* data, we will test the hypotheses:

$$H_0\text{: mean response time} \geq 25 \text{ minutes}$$

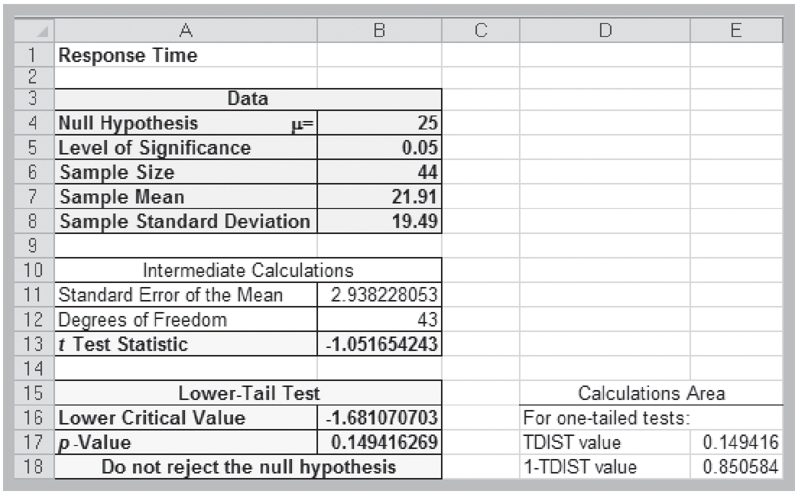
$$H_1\text{: mean response time} < 25 \text{ minutes}$$

with a level of significance of 0.05. This is a lower-tailed, one-sample test for the mean with an unknown standard deviation.

We will use the *PHStat* procedure for a one-sample test for the mean with an unknown population standard deviation (see Appendix 5.2A, “One-Sample Test for the Mean, Sigma Unknown”). From the 44 observations in the Excel file, we computed the sample mean to be 21.91 and the sample standard deviation as 19.49. Figure 5.2 shows the output provided by *PHStat*. The *Data* portion of the output simply summarizes the hypothesis we are testing, level of significance specified, and sample statistics. The *t-Test Statistic* is calculated using the formula:

Spreadsheet Note

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \tag{5.1}$$



	A	B	C	D	E
1	**Response Time**				
2					
3	**Data**				
4	**Null Hypothesis** μ=	**25**			
5	**Level of Significance**	**0.05**			
6	**Sample Size**	**44**			
7	**Sample Mean**	**21.91**			
8	**Sample Standard Deviation**	**19.49**			
9					
10	Intermediate Calculations				
11	Standard Error of the Mean	2.938228053			
12	Degrees of Freedom	43			
13	***t*** **Test Statistic**	**-1.051654243**			
14					
15	Lower-Tail Test			Calculations Area	
16	**Lower Critical Value**	**-1.681070703**		For one-tailed tests:	
17	***p*-Value**	**0.149416269**		TDIST value	0.149416
18	**Do not reject the null hypothesis**			1-TDIST value	0.850584

FIGURE 5.2 Results for One-Sample Test for the Mean, Sigma Unknown

where μ_0 is the hypothesized value and $s/\sqrt{n}$ is the standard error of the sampling distribution of the mean. Applied to this example, we have:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{21.91 - 25}{19.49/\sqrt{44}} = \frac{-3.09}{20938} = -1.05$$

Observe that the numerator is the distance between the sample mean (21.91) and the hypothesized value (25). By dividing by the standard error, the value of t represents the number of standard errors the sample mean is from the hypothesized value. In this case, the sample mean is 1.05 standard errors below the hypothesized value of 25.

This notion provides the fundamental basis for the hypothesis test—if the sample mean is "too far" away from the hypothesized value, then the null hypothesis should be rejected. The decision is based on the level of significance, α. For a one-tailed test, the critical value is the number of standard errors away from the hypothesized value for which the probability of exceeding the critical value is α. If $\alpha = 0.05$, for example, then we are saying that there is only a 5% chance that a sample mean will be that far away from the hypothesized value purely because of sampling error, and that the small likelihood of this occurring suggests that the true population mean is different from what was hypothesized.

The *t-Test Statistic* has a t-distribution with $n - 1$ degrees of freedom. If the level of significance is 0.05, then the critical value for a lower-tail test is the value of the t-distribution with $n - 1$ degrees of freedom that provides a lower tail area of 0.05; that is, $t_{\alpha, n-1}$. We may find t-values in Table A.2 in the appendix at the end of the book or by using the Excel function T.INV($\alpha, n - 1$). Hence, to find the critical value, we find $t_{0.05,43}$ = T.INV(0.05,43) = −1.68. (The T.INV function returns the left-tail critical value; for the right-tail critical value for an upper-tail test, use T.INV($1 - \alpha, n - 1$).)

By comparing the *t-Test Statistic* with the *Lower Critical Value,* we see that the test statistic does not fall below the critical value (that is, $-1.05 > -1.68$) and is not in the rejection region. Therefore, we cannot reject H_0 and cannot conclude that the mean response time is less than 25 minutes. Figure 5.3 illustrates the conclusion we reached. Even though the sample mean is less than 25, we cannot conclude that the mean response time is less than 25 minutes because of the large amount of sampling error.

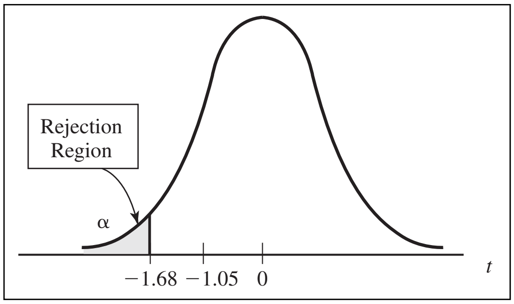


FIGURE 5.3 *t*-Test for Mean Response

Using *p*-Values

In the *PHStat* output in Figure 5.2, we see something called a *p-value.* An alternative approach to comparing a test statistic to a critical value in hypothesis testing is to find the probability of obtaining a test statistic value equal to or more extreme than that obtained from the sample data when the null hypothesis is true. This probability is commonly called a **p-value**, or **observed significance level**. For example, the *t-test Statistic* for the hypothesis test in the response time example is −1.05. If the true mean is really 25, then what is the probability of obtaining a test statistic of −1.05 or less (the area to the left of −1.05 in Figure 5.3)? Equivalently, what is the probability that a sample mean from a population with a mean of 25 will be at least 1.05 standard errors below 25? We can calculate this using the Excel function T.DIST(–1.05,43,TRUE) = 0.149416. In other words, there is about a 15% chance that the test statistic would be −1.05 or smaller if the null hypothesis were true. This is a fairly high probability, so it would be difficult to conclude that the true mean is less than 25, and we could attribute the fact that the test statistic is less than the hypothesized value to sampling error alone and not reject the null hypothesis. In general, compare the *p*-value to the chosen level of significance; whenever $p < \alpha$, reject the null hypothesis. *p*-values make it easy to draw conclusions about hypothesis tests.

Next, we illustrate a two-tailed test using the data for the respondents in the Excel file *Vacation Survey.* Suppose that the sponsor of the survey wanted to target individuals who were approximately 35 years old. Thus, we wish to test whether the average age of respondents is equal to 35. The hypothesis to test is:

$$H_0\text{: mean age} = 35$$
$$H_1\text{: mean age} \neq 35$$

The sample mean is computed to be 39.12, and the sample standard deviation is 7.53. Figure 5.4 shows the results using the *PHStat t-Test for the Mean, Sigma Unknown*.

The test statistic is the same as in the previous example:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = (39.12 - 35)/(7.53/\sqrt{33}) = 3.143$$

In this case, the sample mean is 3.143 standard errors above the hypothesized mean of 35. However, because this is a two-tailed test, the rejection region and decision rule are different. For a level of significance α, we reject H_0 if the *t-Test Statistic* falls either below the negative critical value, $-t_{\alpha/2,n-1}$, or above the positive critical value, $t_{\alpha/2,n-1}$. Using the Excel function T.INV(.025,32) to calculate $t_{.025,32,}$, we obtain −2.0369. Thus, the critical values are ±2.0369. Because the *t-Test Statistic* does not fall between these values, we must reject the null hypothesis that the average age is 35 (see Figure 5.5).