What Is Hypothesis Testing?

- Hypothesis testing is used in a variety of settings
  - The **Food and Drug Administration** (FDA), for example, tests **new products** before allowing their sale
    - If the sample of people exposed to the new product shows some side effect significantly more frequently than would be expected to occur by chance, the FDA is likely to withhold approval of marketing that product
  - Similarly, **economists** have been **statistically testing** various **relationships**, for example that between consumption and income
- Note here that while we **cannot prove** a given hypothesis (for example the existence of a given relationship), we often can reject a given hypothesis (again, for example, **rejecting** the existence of a given relationship)

Classical Null and Alternative Hypotheses

- The researcher first states the hypotheses to be tested
- Here, we distinguish between the **null** and the **alternative** hypothesis:
  - **Null hypothesis (“H_0”):** the outcome that the researcher does **not** expect (almost always includes an equality sign)
  - **Alternative hypothesis (“H_A”):** the outcome the researcher **does** expect
- Example:

H_0: $\beta \leq 0$ (the values you do **not** expect)

H_A: $\beta > 0$ (the values you **do** expect)

Type I and Type II Errors

- Two types of errors possible in hypothesis testing:
  - **Type I: Rejecting** a **true null** hypothesis
  - **Type II: Not rejecting** a **false null** hypothesis
- Example: Suppose we have the following null and alternative hypotheses:

  H_0: $\beta \leq 0$

  H_A: $\beta > 0$

  - Even if the **true** β really is not positive, in any one sample we might still observe an **estimate** of β that is sufficiently positive to lead to the **rejection** of the null hypothesis
- This can be illustrated by Figure 5.1

Figure 5.1 Rejecting a True Null Hypothesis Is a Type I Error

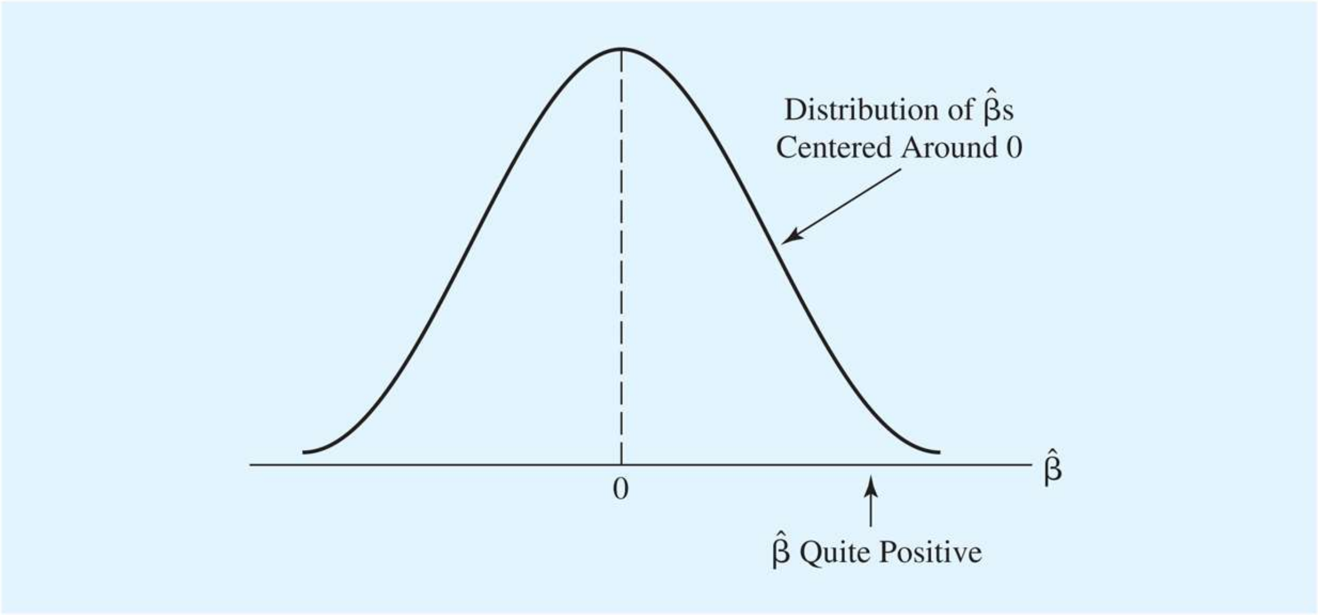

Figure 5.1 Rejecting a True Null Hypothesis Is a Type I Error

If $\beta = 0$, but you observe a $\hat{\beta}$ that is very positive, you might reject a true null hypothesis, H_0: $\beta \leq 0$, and conclude incorrectly that the alternative hypothesis H_A: $\beta > 0$ is true.

Type I and Type II Errors (cont.)

- Alternatively, it's possible to obtain an estimate of β that is **close enough** to zero (or negative) to be considered "not significantly positive"
- Such a result may lead the researcher to "accept" the null hypothesis that $\beta \leq 0$ when in truth $\beta > 0$
- This is a **Type II Error**; we have failed to reject a false null hypothesis!
- This can be illustrated by Figure 5.2

Figure 5.2 Failure to Reject a False Null Hypothesis Is a Type II Error

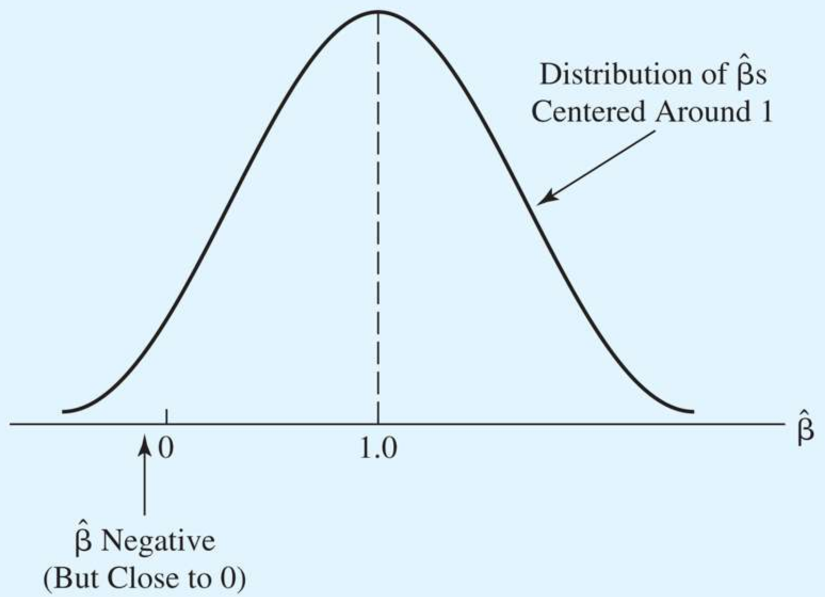

Figure 5.2 Failure to Reject a False Null Hypothesis Is a Type II Error

If $\beta = 1$, but you observe a $\hat{\beta}$ that is negative but close to zero, you might fail to reject a false null hypothesis, H_0: $\beta \leq 0$, and incorrectly ignore the fact that the alternative hypothesis, H_A: $\beta > 0$, is true.

Decision Rules of Hypothesis Testing

- To test a hypothesis, we calculate a **sample statistic** that determines when the null hypothesis can be rejected depending on the magnitude of that sample statistic relative to a preselected **critical value** (which is found in a statistical table)
- This procedure is referred to as a **decision rule**
- The decision rule is **formulated before** regression estimates are obtained
- The **range** of possible values of the estimates is divided into two regions, an “**acceptance**” (really, **non-rejection**) region and a **rejection region**
- The critical value effectively separates the “acceptance”/non-rejection region from the rejection region when testing a null hypothesis
- Graphs of these “acceptance” and rejection regions are given in Figures 5.3 and 5.4

Figure 5.3 "Acceptance" and Rejection Regions for a One-Sided Test of β

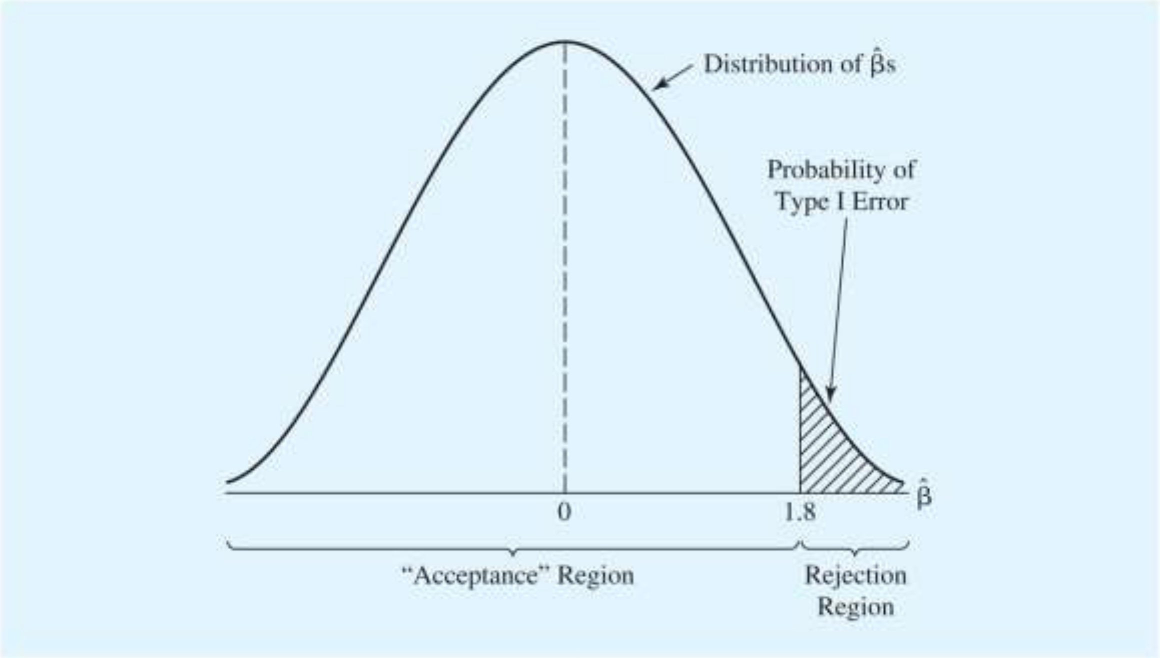

Figure 5.3 "Acceptance" and Rejection Regions for a One-Sided Test of β

For a one-sided test of H_0: $\beta \leq 0$ vs. H_A: $\beta > 0$, the critical value divides the distribution of $\hat{\beta}$ (centered around zero on the assumption that H_0 is true) into "acceptance" and rejection regions.

Figure 5.4 "Acceptance" and Rejection Regions for a Two-Sided Test of β

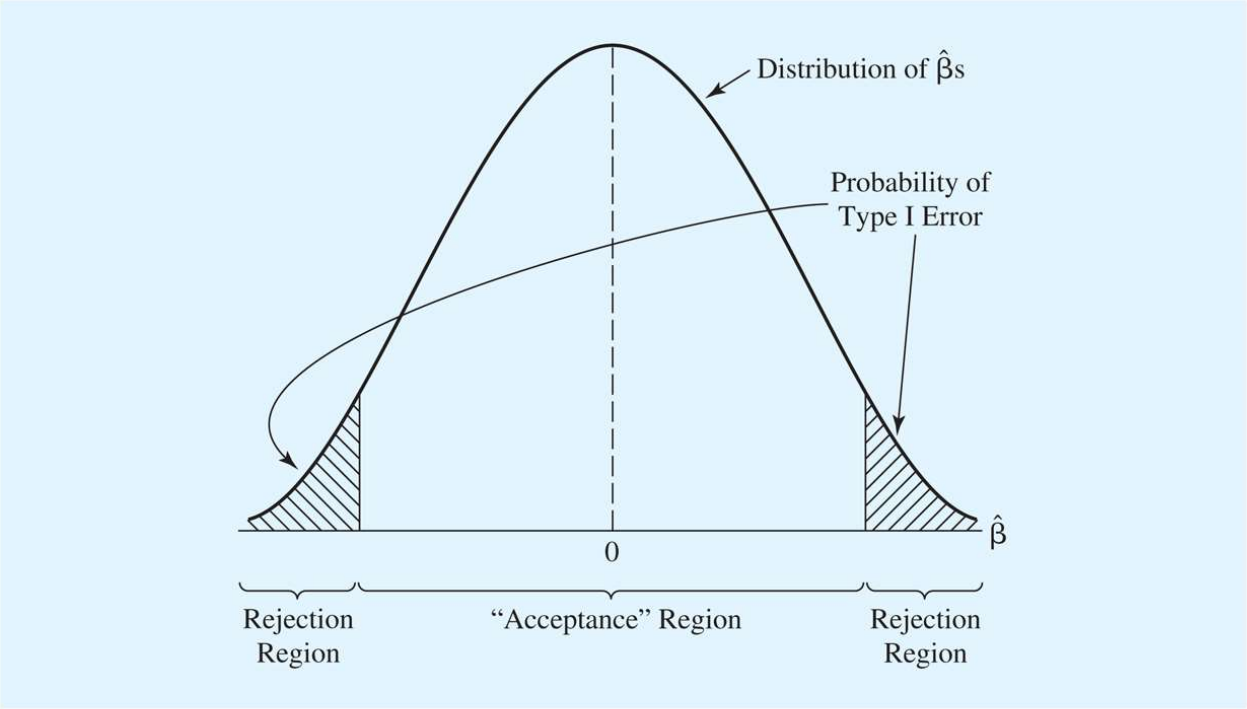

Figure 5.4 "Acceptance" and Rejection Regions for a Two-Sided Test of β

For a two-sided test of H_0: $\beta = 0$ vs. H_A: $\beta \neq 0$, we divided the distribution of $\hat{\beta}$ into an "acceptance" region and *two* rejection regions.

The *t*-Test

- The **t-test** is the test that econometricians usually use to test hypotheses about **individual** regression slope coefficients
  - Tests of more than one coefficient at a time (**joint** hypotheses) are typically done with the **F-test**, presented in Section 5.6
- The appropriate test to use when the stochastic error term is **normally distributed** and when the **variance** of that distribution must be **estimated**
  - Since these usually are the case, the use of the t-test for hypothesis testing has become **standard practice** in econometrics

The *t*-Statistic

- For a typical multiple regression equation:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i \qquad (5.1)$$

  we can calculate t-values for each of the estimated coefficients

  – Usually these are only calculated for the **slope coefficients**, though (see Section 7.1)

- Specifically, the t-statistic for the *k*th coefficient is:

$$t_k = \frac{(\hat{\beta}_k - \beta_{H_0})}{SE(\hat{\beta}_k)} \qquad (k = 1, 2, \ldots, K) \qquad (5.2)$$

The Critical *t*-Value and the *t*-Test Decision Rule

- To decide whether to reject or not to reject a null hypothesis based on a calculated *t*-value, we use a **critical *t*-value**
- A **critical *t*-value** is the value that distinguishes the “acceptance” region from the rejection region
- The critical *t*-value, t_c, is selected from a ***t*-table** (see Statistical Table B-1 in the back of the book) depending on:
  - whether the test is one-sided or two-sided,
  - the level of Type I Error specified and
  - the degrees of freedom (defined as the number of observations minus the number of coefficients estimated (including the constant) or N – K – 1)

The Critical *t*-Value and the *t*-Test Decision Rule (cont.)

- The rule to apply when testing a single regression coefficient ends up being that you should:

Reject H_0 if $|t_k| > t_c$ and if t_k also has the sign implied by H_A

Do not reject H_0 otherwise

The Critical *t*-Value and the *t*-Test Decision Rule (cont.)

- Note that this decision rule works both for calculated *t*-values and critical *t*-values for **one-sided** hypotheses around zero (or another hypothesized value, S):

$H_0: \beta_k \leq 0$	$H_0: \beta_k \leq S$
$H_A: \beta_k > 0$	$H_A: \beta_k > S$
$H_0: \beta_k \geq 0$	$H_0: \beta_k \geq S$
$H_A: \beta_k < 0$	$H_A: \beta_k < S$

The Critical t-Value and the *t*-Test Decision Rule (cont.)

- As well as for **two-sided** hypotheses around zero (or another hypothesized value, S):

$H_0: \beta_k = 0$	$H_0: \beta_k = S$
$H_A: \beta_k \neq 0$	$H_A: \beta_k \neq S$

- From Statistical Table B-1 the critical *t*-value for a **one-tailed** test at a given level of significance is exactly **equal** to the critical *t*-value for a **two-tailed** test at **twice** the level of significance of the **one-tailed** test—as also illustrated by Figure 5.5

Figure 5.5 One-Sided and Two-Sided *t*-Tests

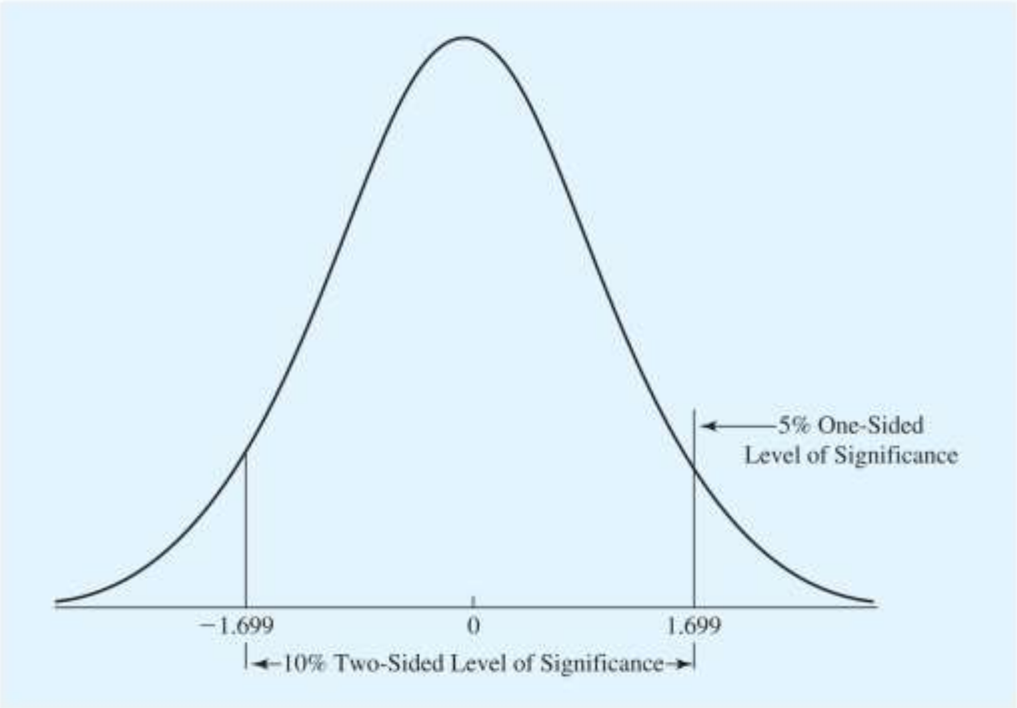

Figure 5.5 One-Sided and Two-Sided *t*-Tests

The t_c for a one-sided test at a given level of significance is equal exactly to the t_c for a two-sided test with twice the level of significance of the one-sided test. For example, $t_c = 1.699$ for a 10-percent two-sided *and* for a 5-percent one-sided test (for 29 degrees of freedom).

Choosing a Level of Significance

- The **level of significance** must be chosen **before** a critical value can be found, using Statistical Table B
- The **level of significance** indicates the probability of observing an estimated t-value **greater** than the **critical** t-value if the **null hypothesis** were **correct**
- It also measures the **amount** of **Type I Error** implied by a particular critical t-value
- Which level of significance is chosen?
  - 5 percent is recommended, **unless** you know something **unusual** about the **relative costs** of making Type I and Type II Errors

Confidence Intervals

- A confidence interval is a range that contains the true value of an item a specified percentage of the time
- It is calculated using the estimated regression coefficient, the two-sided critical t-value and the standard error of the estimated coefficient as follows:

$$\text{Confidence interval} = \hat{\beta} \pm t_c \cdot SE(\hat{\beta}) \quad (5.5)$$

- What's the relationship between confidence intervals and two-sided hypothesis testing?
- If a hypothesized value fall within the confidence interval, then we cannot reject the null hypothesis

p-Values

- This is an alternative to the t-test
- A p-value, or marginal significance level, is the **probability** of observing a t-score **that size or larger** (in absolute value) if the **null hypothesis** were **true**
- **Graphically**, it's two times the **area** under the curve of the t-distribution between the absolute value of the actual t-score and infinity.
- **In theory**, we could find this by combing through pages and pages of statistical tables
- But we don't have to, since we have **EViews** and **Stata**: these (and other) statistical software packages automatically give the p-values as part of the standard output!
- In light of all this, the **p-value decision rule** therefore is:

  Reject H_0 if p-value$_K$ < the level of significance and if $\hat{\beta}_K$ has the sign implied by H_A

Examples of *t*-Tests: One-Sided

- The **most common use** of the **one-sided** t-test is to determine whether a regression coefficient is significantly different from zero (in the direction predicted by theory!)
- This involves **four steps:**

1. Set up the null and alternative hypothesis
2. Choose a level of significance and therefore a critical t-value
3. Run the regression and obtain an estimated t-value (or t-score)
4. Apply the decision rule by comparing calculated t-value with the critical t-value in order to reject or not reject the null hypothesis

- Let's look at each step in more detail for a **specific example:**

Examples of *t*-Tests: One-Sided (cont.)

- Consider the following simple model of the aggregate retail sales of new cars:

$$Y = f(\overset{+}{X}_1, \overset{-}{X}_2, \overset{-}{X}_3) + \epsilon \quad (5.6)$$

  Where:

  Y = sales of new cars

  X1 = real disposable income

  X2 = average retail price of a new car adjusted by the consumer price index

  X3 = number of sports utility vehicles sold

- The four steps for this example then are as follows:

Step 1: Set up the null and alternative hypotheses

- From equation 5.6, the one-sided hypotheses are set up as:

1. H_0: $\beta_1 \leq 0$
   H_A: $\beta_1 > 0$

2. H_0: $\beta_2 \geq 0$
   H_A: $\beta_2 < 0$

3. H_0: $\beta_3 \geq 0$
   H_A: $\beta_3 < 0$

- Remember that a **t-test** typically is not run on the estimate of the **constant term** β_0

Step 2: Choose a level of significance and therefore a critical t-value

- Assume that you have considered the various costs involved in making Type I and Type II Errors and have chosen 5 percent as the level of significance
- There are 10 observations in the data set, and so there are 10 – 3 – 1 = 6 degrees of freedom
- At a 5-percent level of significance, the critical t-value, t_c, can be found in Statistical Table B-1 to be 1.943

Step 3: Run the regression and obtain an estimated t-value

- Use the data (annual from 2000 to 2009) to run the regression on your OLS computer package
- Again, most statistical software packages automatically report the t-values
- Assume that in this case the t-values were 2.1, 5.6, and –0.1 for β_1, β_2, and β_3, respectively

Step 4: Apply the t–test decision rule

- As stated in Section 5.2, the decision rule for the t-test is to:

  Reject H_0 if $|t_k| > t_c$ and if t_k also has the sign implied by H_A

- In this example, this amounts to the following three conditions:

  For β_1: Reject H_0 if $|2.1| > 1.943$ and if 2.1 is positive.

  For β_2: Reject H_0 if $|5.6| > 1.943$ and if 5.6 is positive.

  For β_3: Reject H_0 if $|-0.1| > 1.943$ and if –0.1 is positive.

- Figure 5.6 illustrates all three of these outcomes

Figure 5.6a One-Sided *t*-Tests of the Coefficients of the New Car Sales Model

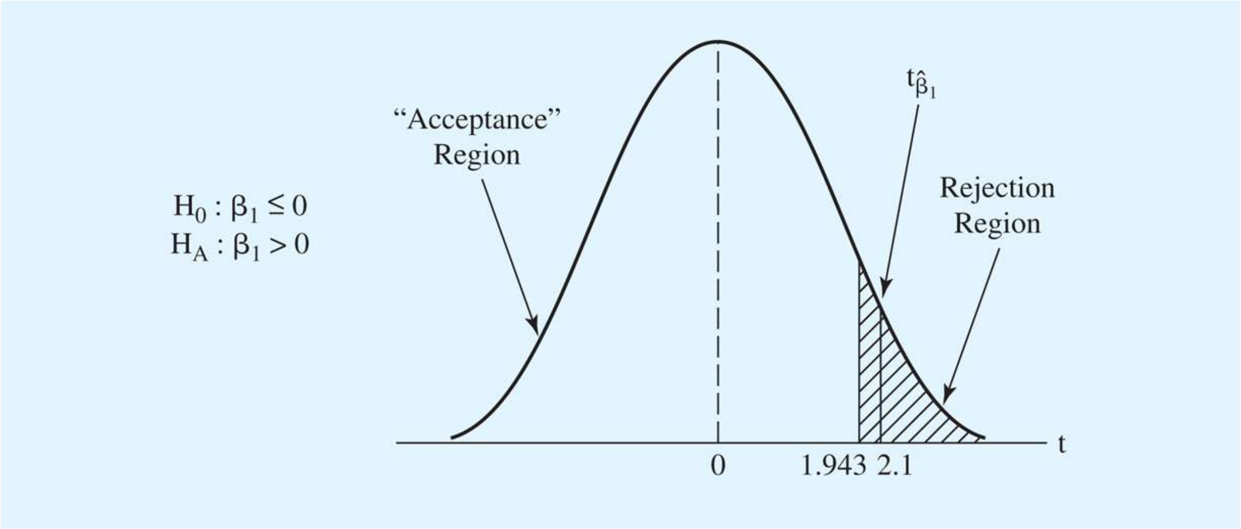

Figure 5.6b One-Sided *t*-Tests of the Coefficients of the New Car Sales Model

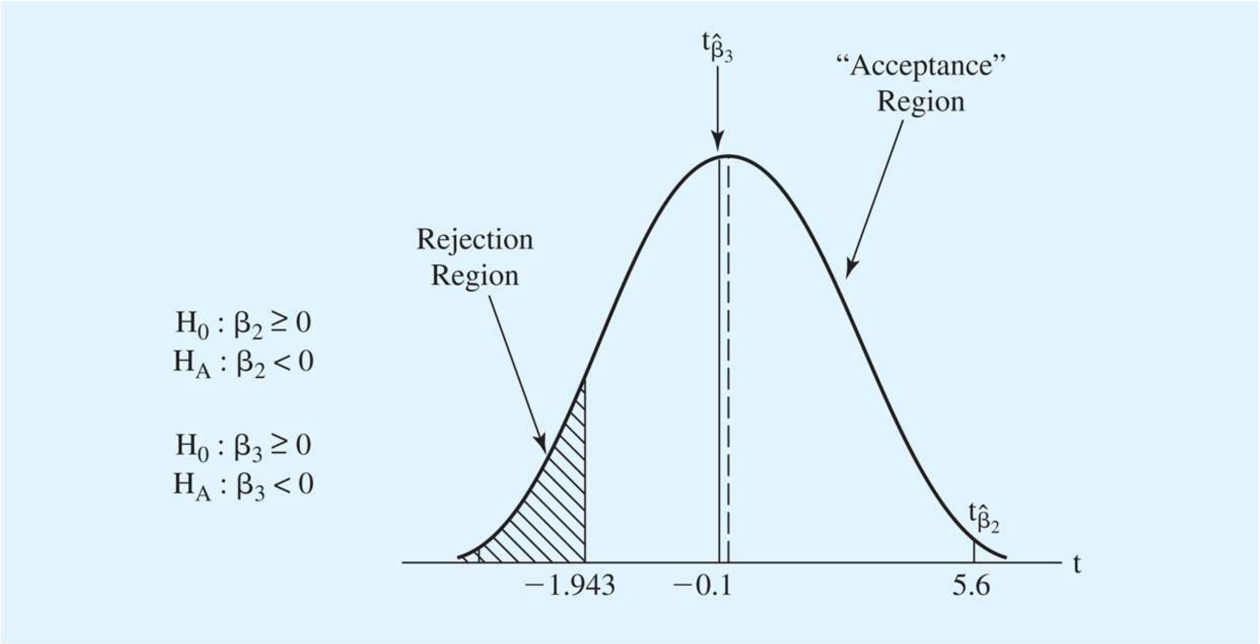

Figure 5.6 One-Sided *t*-Tests of the Coefficients of the New Car Sales Model

Given the estimates in Equation 5.7 and the critical *t*-value of 1.943 for a 5-percent level of significance, one-sided, 6 degrees of freedom *t*-test, we can reject the null hypothesis for $\hat{\beta}_1$, but not for $\hat{\beta}_2$ or $\hat{\beta}_3$.

Examples of *t*-Tests: Two-Sided

- The two-sided test is used when the hypotheses should be rejected if estimated coefficients are significantly different from zero, or a specific nonzero value, **in either direction**
- So, there are two cases:
  1. Two-sided tests of whether an estimated coefficient is significantly different from zero, and
  2. Two-sided tests of whether an estimated coefficient is significantly different from a specific nonzero value
- Let's take an example to illustrate the first of these (the second case is merely a generalized case of this, see the textbook for details), using the **Woody's restaurant example** in Chapter 3:

Examples of *t*-Tests: Two-Sided (cont.)

- Again, in the Woody's restaurant equation of Section 3.2, the impace of the average income of an area on the expected number of Woody's customer's in that area is ambiguous:
- A high-income neighborhood might have more total customers going out to dinner (positive sign), but those customers might decide to eat at a more formal restaurant that Woody's (negative sign)
- The appropriate (two-sided) t-test therefore is:

Figure 5.7 Two-Sided *t*-Test of the Coefficient of Income in the Woody's Model

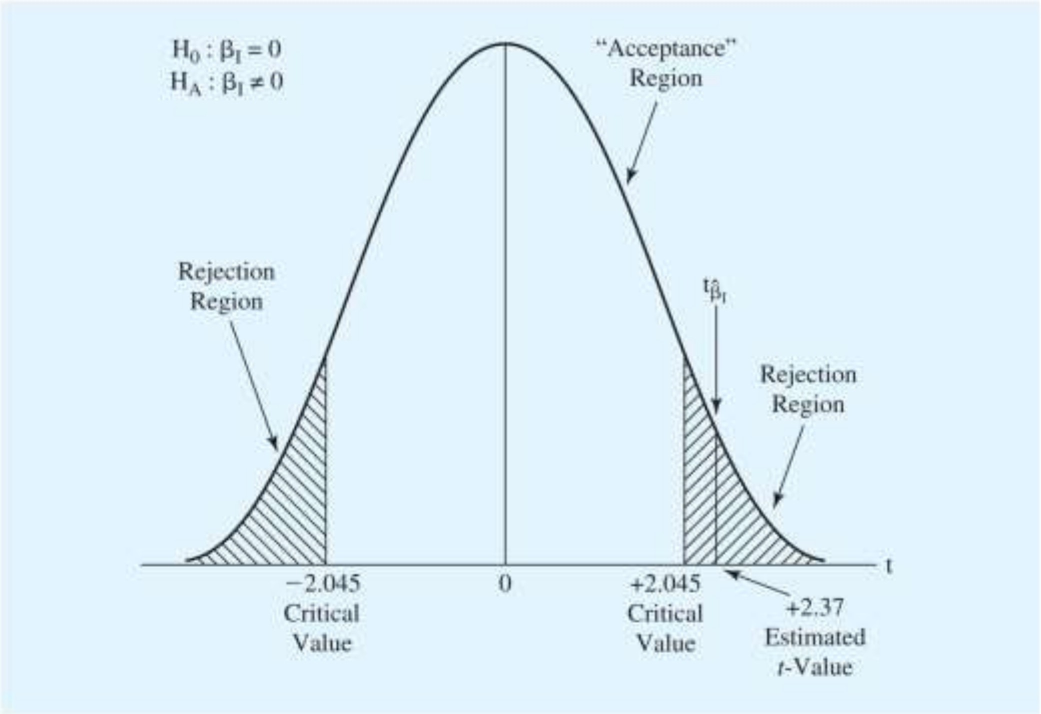

Figure 5.7 Two-Sided *t*-Test of the Coefficient of Income in the Woody's Model

Given the estimates of Equation 5.4 and the critical *t*-values of ±2.045 for a 5-percent level of significance, two-sided, 29 degrees of freedom *t*-test, we can reject the null hypothesis that $\beta_I = 0$.

Examples of *t*-Tests: Two-Sided (cont.)

- The four steps are the **same as** in the **one-sided case**:
  1. Set up the null and alternative hypothesis

     H_0: $\beta_k = 0$
     H_A: $\beta_k \neq 0$
  2. Choose a level of significance and therefore a critical t-value

     Keep the level at significance at 5 percent but this now must be distributed between two rejection regions for 29 degrees of freedom hence the correct critical t-value is 2.045 (found in Statistical Table B-1 for 29 degrees of freedom and a 5-percent, two-sided test)
  3. Run the regression and obtain an estimated t-value:

     The t-value remains at 2.37 (from Equation 5.4)
  4. Apply the decision rule:
     For the two-sided case, this simplifies to:
     Reject H_0 if $|2.37| > 2.045$; so, reject H_0