Chapter 5: Basic Statistics and Hypothesis Testing

In this chapter:

Viewing the t-value from an OLS regression (*UE* 5.2.1):

The Woody's Restaurants example is used to explain the use of t-values to test hypotheses concerning the coefficients on the independent variables in an OLS regression model. Follow these steps to open the Woody's Restaurant workfile in EViews and run the regression for the equation $Y_t = \hat{\beta}_0 + \hat{\beta}_N N_t + \hat{\beta}_P P_t + \hat{\beta}_I I_t + e_t$.

Step 1. Open the EViews workfile named *Woody3.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C N P I* in the Equation Specification: window, and click **OK.** EViews generates the following output (also printed in *UE*, Table 3.2, p. 76):

Dependent Variable: Y				
Method: Least Squares				
Date: 05/23/00 Time: 05:55				
Sample: 1 33				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic[1]	Prob.
C	102192.4	12799.83	7.983891	0.0000
N	-9074.674	2052.674	-4.420904	0.0001
P	0.354668	0.072681	4.879810	0.0000
I	1.287923	0.543294	2.370584	0.0246
R-squared	0.618154	Mean dependent var		125634.6
Adjusted R-squared	0.578653	S.D. dependent var		22404.09
S.E. of regression	14542.78	Akaike info criterion		22.12079
Sum squared resid	6.13E+09	Schwarz criterion		22.30218
Log likelihood	-360.9930	F-statistic		15.64894
Durbin-Watson stat	1.758193	Prob(F-statistic)		0.000003

All information needed for hypothesis testing using the t-test is found in the middle of the EViews equation output table (highlighted in yellow). The first column identifies the name of the variable. The second column reports the estimated coefficient ($\hat{\beta}_k$) for each variable, and the third column reports the standard error for the estimated coefficient (SE $\hat{\beta}_k$). The fourth column prints the calculated t-value given that the border value implied by the null hypothesis (β_{Ho}) is zero (i.e., the t-value in this case is ($\hat{\beta}_k$)/(SE $\hat{\beta}_k$)).

[1] The EViews program uses the term t-Statistic rather than the term t-value, which is used in *UE* and this guide.

Calculating critical t-values and applying the decision rule (*UE* 5.2.2):

The critical t-value (t_c) is the value that separates the "acceptance" region from the "rejection" region. Look up this value in *UE*, Statistical Table B-1, p.607. Its value depends on the degrees of freedom (printed in column one of Table B-1) and the level of Type I error specified (i.e., the number at the top of the column in Table B-1 for two-tailed tests or double the value for one-tailed tests). Alternately, you can follow these steps to have EViews calculate the one-tailed and two-tailed, 5% significance level critical t-values (t_c):

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C N P I* in the Equation Specification: window, and click **OK.**
Step 3. Select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.
Step 4. To create a vector object named *result* with 10 rows (to store the results of the test statistics for Woody's Restaurants regression), type the following command in the command window: *vector(10) result* and press **Enter**.[2]
Step 5. To compute the two-tailed critical t-value (t_c) for the 5% significance level and save the value in the first row of the vector object named *result*, type the following equation in the command window and press **Enter**:[3]
result(1)=@qtdist(.975,(eq01.@regobs-eq01.@ncoef)).[4]
Step 6. To compute the one-tailed critical t-value (t_c) for the 5% significance level and save the value in the second row of the vector object named *result*, type the following equation in the command window and press **Enter**:
result(2)=@qtdist(.95,(eq01.@regobs-eq01.@ncoef)).
Step 7. Double click on the Vector Object named *result* in the workfile window to view the two-tailed and one-tailed, 5% significance level critical t-value (t_c) for the Woody's Restaurants regression. In this case, the value reported in row one of the *result* vector is *2.045230* and the value reported in row two is *1.69912702653*.

Population was hypothesized to have a positive effect on the number of customers eating at Woody's Restaurants. This implies that the coefficient on *P* is expected to be positive and a one-tailed test is appropriate. Thus, the hypothesis that the coefficient is zero ($\hat{\beta}_p = 0$) is rejected at the 5% significance level because the calculated t-value (*4.879810*) is greater than the one-tailed critical t-value (5% level of significance), calculated in **Step 6** to be *1.69912702653*.

While it is important to be able to apply the decision rule by comparing the calculated t-value (reported in the EViews output) with the critical t-value just calculated, EViews makes it

---

[2] Alternately, select **Objects/New Object/Matrix-Vector-Coef** from the main menu or the workfile menu. Write *result* in the Name for Object: window, and click **OK**. Select *Vecto* in the Type window, enter *10* in the Rows window, enter *1* in the Columns window, and click **OK**.
[3] The command, *@qtdist(p,v)*, calculates the value where the cumulative density function (CDF) of the t-distribution with (*v*) degrees of freedom equals (*p*) probability, leaving *1-p%* of the t-distribution in each tail. Note that you should enter *0.975* for "*p*" for the two-tailed 5% significance level calculation and *0.95* for "*p*" for the one-tailed 5% significance level. In this case, *v* equals (*eq01.@regobs-eq01.@ncoef*), which calculates the degrees of freedom for equation *eq01*. The {*eq01.*} Part can be omitted if the calculation relates to the last regression run.
[4] If you omit the *result(1)*, EViews returns a scalar value on the status line located in the lower left of your screen.

possible to test the null hypothesis that a coefficient is zero (i.e., $\beta_{Ho} = 0$) without knowing the critical t-value. Instead, you can examine the probability (Prob.) value in the last column of the EViews OLS regression output (see the *EQ01* Estimation Output table). The Prob. value shows the probability of drawing a t-value as extreme as the one actually observed when, in fact, the coefficient value is actually zero[5]. This probability is also known as the p-value or the marginal significance level. In terms of *UE*, it represents the probability of making a Type I error if the null hypothesis, that the coefficient is zero, is rejected. Given a p-value, you can tell at a glance if the null hypothesis, that the true coefficient is zero against a two-sided alternative that it differs from zero, should be rejected or accepted. For example, a p-value lower than .05 suggests rejection of the null hypothesis, for a two-tailed test at the 5% significance level. The appropriate probability is one-half that reported by EViews, for a one-sided test. This is comparable to reading the value at the top of the Critical t-Value table that is double the significance level that you are testing for a one-tailed test (i.e., using 0.10 instead of 0.05 for 5% significance level).

Applying the Prob. value, one-tailed test to the population (*P*) variable in the Woody's Restaurants regression, the hypothesis that the coefficient is zero ($\hat{\beta}_{p} = 0$) is rejected at the 5% significance level if one-half of the Prob. value in the last column is less than or equal to 0.05. Note that the null hypothesis is also rejected at the 1% significance level.

Calculating confidence intervals (*UE* 5.2.4):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). To calculate and record the 90% confidence interval for a coefficient using EViews:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To calculate the lower value for the 90% confidence interval for the population coefficient, enter the following formula in the command window, and press **Enter**:[6]
*result(3)= eq01.@coefs(3)-(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))*eq01.@stderrs(3).*
Step 3. To calculate the upper value for the 90% confidence interval for the population coefficient, enter the following formula in the command window, and press **Enter**:
*result(4)= eq01.@coefs(3)+(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))*eq01.@stderrs(3).*
Step 4. To view the lower and upper confidence interval values, double click the vector icon named *result*. Note that the values *0.231175* and *0.478162* are printed in rows three and four respectively (the same values are reported in *UE*, p. 128).

---

[5] Under the assumption that the errors are normally distributed, or that the estimated coefficients are asymptotically normally distributed.

[6] *eq01.@coefs(i)* and *eq01.@stderrs(i)* are scalar values of the coefficient and standard error of the i^{th} variable in regression *eq01*,where i represents the coefficient number (including the constant) listed in the EViews OLS Estimation Output. Since the population (*P*) variable is listed third in the Estimation Output, *@coefs(3)* and *@stderrs(3)*, calculates the value for the population coefficient and its standard error respectively. As in section 5.2.2, *(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))* calculates the value where the cumulative density function (CDF) of the t-distribution equals *0.95* probability, leaving 5% of the t-distribution in each tail. The term *(eq01.@regobs-eq01.@ncoef)* calculates the degrees of freedom for *EQ01*, with *eq01.@regobs* calculating the number of observations used to estimate *EQ01* and *eq01.@ncoef* calculating the number of coefficients estimated, including the constant.

Performing the t-test of the simple correlation coefficient (*UE* 5.3.3):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). To use the t-test to determine whether a particular simple correlation coefficient between *Y* and *P* is significant:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To calculate the simple correlation coefficient (r) and store it in the fifth row of the result vector, type the following command in the command window, and press **Enter**: *result(5)= @cor(y,p).*
Step 3. To convert the simple correlation coefficient between *Y* and *P* into a t-value and store it in row six of the *result* vector, type the following command in the command window, and press **Enter**: *result(6) = (@cor(y,p)*((@obs(y)-2)^.5))/((1-@cor(y , p)^2)^.5).*
Step 4. To calculate the critical t-value (t_c) for the t-distribution with *@obs(y)-2* degrees of freedom and store it in the seventh row of the *result* vector, type the following command in the command window, and press **Enter**: *result(7) = @qtdist(.975,(@obs(y)-2)).*
Step 5. Double click the icon for the *result* vector to view the results in rows 5, 6, and 7.

Performing the F-test of overall significance (*UE* 5.5):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). The F-statistic is reported as (15.64894) in the *EQ01* estimation output table. Follow these steps to calculate the F-statistic from *EQ01* and save it in row seven of a vector named *result*:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To store the F-statistic from *EQ01* in row eight of the results vector, type the following command in the command window, and press **Enter**: *result(8)=eq01.@f.*
In this case, the F-statistic tests the hypothesis that all of the slope coefficients excluding the constant, in *EQ01*, are zero. Under the null hypothesis with normally distributed errors, this statistic has an F-distribution with k degrees of freedom in the numerator and n-k-1 degrees of freedom in the denominator, where n equals the number of observation and k equals the number of independent variables (k does not include the constant) in the model. The null hypothesis can be rejected if the calculated F-statistic exceeds the critical F-value at the chosen level of significance. The critical F-value can be determined from *UE*, Appendix B, Statistical Tables B-2 or B-3, depending on the level of significance chosen. Alternately,
Step 3. To have EViews calculate the 5% critical F-value for *EQ01* and store it in row nine of the results vector, type the following command in the command window, and press **Enter**: *result(9)=@qfdist(0.95,eq01.@ncoef-1,eq01.@regobs- eq01.@ncoef).*
Step 4. Double click the icon for the *result* vector to view the results in rows 8 and 9.

test statistic	result row	value
t-critical for regression - 5% level of significance (two-tailed test) =	R1	2.04523
t-critical for regression - 5% level of significance (one-tailed test) =	R2	1.699127
lower confidence interval =	R3	0.231175
upper confidence interval =	R4	0.478162
The simple correlation coefficient (r) =	R5	0.392568
t-calculated for correlation =	R6	2.376503
t-critical for correlation =	R7	2.039513
The F-statistic =	R8	15.64894
Critical value of the F-statistic - 5% level of significance =	R9	2.93403
	R10	0

Since the 5% critical F-value for *EQ01* (i.e., *2.934030*) is significantly less than the calculated F-statistic (i.e., *15.64894*), we can reject the null hypothesis that all of the slope coefficients in *EQ01* are zero.

The p-value printed just below the F-statistic in the EViews regression output, denoted Prob(F-statistic), represents the marginal significance level of the F-test. If the p-value is less than the significance level you are testing, say .05, you reject the null hypothesis that all slope coefficients are equal to zero. For *EQ01*, the p-value is 0.000003, so we reject the null hypothesis that all of the regression coefficients are zero. Note that the F-test is a joint test so that even if all the t-values are insignificant, the F-statistic can be highly significant.

Exercises:

16. EViews can be used to complete parts 1, b, c & f of exercise 16.
    a. Review the section Calculating critical t-values and applying the decision rule to learn how to use EViews to calculate the critical values for testing your hypothesis concerning the regression coefficients in this exercise.
    b. Review the section Performing the F-test of overall significance to learn how to use EViews to calculate the critical value for the F-test of the overall significance of the estimated equation.
    c. Review the section Calculating confidence intervals to learn how to use EViews to determine lower and upper confidence intervals for an estimated coefficient.
    d.
    e.
    f. Review the section Using EViews to estimate a multiple regression model of beef demand in Chapter 2, if you have trouble estimating this multiple regression model using EViews.