# A9W8c The use of t-test to test whether the predictor variable is a significant contributor

Recall that the true relationship between X and Y (Y = β0 + β1X) is being estimated from the sample relationship (). To determine the existence of a significant relationship between X and Y variables we will need a t-test to check whether 1 is equal to zero. This is really a test to determine if the regression model is usable.

If the slope (β1) is significantly different from zero, then we can use the regression equation to predict the dependent variable for any value of the independent variable. If the slope (β 1) is zero, then the independent variable has no prediction value. In this case for every value of the independent variable, the dependent variable would be the same. Therefore, when this is the situation we would not use the equation to make predictions.

In order to test the significance of the relationship between y and x, we test the null hypothesis

H0: β1 = 0

no linear relationship

This implies that there is no change in the value of the variable y as the variable x increases in size.

The alternative hypothesis states that the value of the y variable changes as the value of the x variable increases in size.

H1: β1  0

linear relationship exists and since we believe that the relationship is not zero (2 tail test)

For simple linear regression, which has one independent variable, the F-test is equivalent to the t-test (see textbook Section 10.1.7). In this hypothesis test we are assessing the possibility that β1 = 0. In order to test this hypothesis, we will calculate a measure of the difference between the value of the population slope (β1) and the sample slope (b1). The value of b1 will change as we collect different samples, and this would create a sampling distribution for the b1 term. It can be shown that if the regression assumptions hold, then the population of all possible values of the term b1 will be normally distributed with mean of β1 and with a standard deviation given by equation (W10.1).

$σ\_{b\_{1}}=\frac{σ}{\sqrt{SSX}}$ (W10.1)

Where SSX is the sum of squares for x, or, $SSX=\sum\_{i=1}^{n}(x\_{i}-\overbar{x})^{2}$.

Equation (W10.1) can be re-written as equation (W10.2), if we note that the standard error of the estimate Sxy is a point estimate of σ, and sb1 is a point estimate of σb1.

$s\_{b\_{1}}=\frac{S\_{xy}}{\sqrt{SSX}}=\frac{SEE}{\sqrt{(x-\overbar{x})^{2}}}$ (W10.2)

Where, SEE is the standard error of the estimate given by Excel function STEYX(), also given in textbook equation (10.11).

It can be shown that the relationship between b1, β1, and tcal, is given by equation (W10.3) which follows a t distribution with the number of degrees of freedom df = n – 2.

$t\_{cal}=\frac{b\_{1}-β\_{1}}{s\_{b\_{1}}}$ (W10.3)

**Example W10.1**

Re-consider the textbook Example 10.1 data set and test the significance of the predictor variable (x). Figures W10.1 and W10.2 illustrate the Excel solution to undertake the required hypothesis Student’s t test.



Figure W10.1

**Excel solution**

x: Cells B5:B54 Values

y: Cells C5:C54 Values

b0 = Cell C57 Formula: =INTERCEPT (C5:C54,B5:B54)

b1 = Cell C58 Formula: =SLOPE (C5:C54,B5:B54)

ŷ= Cells D5 Formula: =$C$57+$C$58\*B5

 Copy formula down D6:D54

(x – xbar)^2= Cell F5 Formula: =(B5-$K$16)^2

 Copy formula down F6:F54

****

Figure W10.2

**Excel solution**

Significance level = Cell K12 Value

SEE = Cell K15 Formula:=STEYX (C5:C54,B5:B54)

Average x = Cell K16 Formula:=AVERAGE (B5:B54)

SSX = Cell K17 Formula:=SUM (F5:F54)

Sb1 = Cell K18 Formula:=K15/SQRT(K17)

t = Cell K19 Formula:=C58/K18

n = Cell K21 Formula:=COUNTA(A5:A54)

k = Cell K22 Value

df = Cell K23 Formula:=K21-(K22+1)

Upper tcri = Cell K24 Formula:=T.INV.2T(K12,K23)

Lower tcri = Cell K25 Formula:=-K24

Two tail p-value = Cell K26 Formula:=T.DIST.2T(K19,K23)

**Step 1 - State hypothesis**

H0: β1 = 0

no linear relationship

H1: β1  0

linear relationship exists and since we believe that the relationship is not zero (2 tail test)

**Step 2 - Select test**

Select the test – we know that this is the t-test for testing if the predictor variable is a significant contributor.

**Step 3 - Set the level of significance** (α = 0.05) (see Cell K12)

**Step 4 - Extract relevant statistic**

Calculate the test statistic, tcalc

If H0 is true then β1 = 0 and equation (W10.3) simplifies to equation (W10.4).

$t\_{cal}=\frac{b\_{1}}{s\_{b\_{1}}}$ (W10.4)

The test statistic follows a t distribution with n – 2 degrees of freedom. From Excel, tcal = 10.50 with 48 degrees of freedom.

Critical t value, tcri

We can now test to see if this sample t value would result in accepting or rejecting H0. From Excel we see that the critical t value = ± 2.01 at a 5% significance level (0.05). At this stage, we need to remember that the hypothesis test implies no relationship for H0 can be rejected.

**Step 5 - Make decision**

Since tcal > tcri (10.50 > 2.01), then the test statistic lies in the rejection zone for H0. Therefore, reject H0. Alternatively, also the p-value < α (5.0007E-14 < 0.05), which confirms we need to reject H0.

We concluded that evidence exist that a significant relationship exists between the two variables.

Note that a similar approach can be used to test if the constant term b0 is a significant contributor to the value of y. This requires the following t-test statistic $t\_{cal}=\frac{b\_{0}}{s\_{b\_{0}}}$ to be calculated and compared with the critical t value.