**Hypothesis Testing for Significance in Multiple Regression**

The **Regression Model**, , gives the relationship between and the in the population. From the sample, regression calculates the **Estimated Regression Equation (ERE)**, , which gives the relationship between and the in the sample.

**Hypothesis Tests About the Population Slopes**

Each population slope gives the relationship between its corresponding and . The following table summarizes the possible relationships between and

|  |  |
| --- | --- |
| **Slope Coefficient** | **Relationship between and** |
|  | Positive linear relationship |
|  | Negative linear relationship |
|  | No relationship |

In order to conclude that there is a relationship between a given and we need to confirm that the population slope coefficient on that does not equal zero. In **Chapter 14: Handout #2**, we learned that we can use a two-tailed test to confirm whether the population slope is different from zero. In multiple regression, we use that same procedure on *each* slope coefficient in turn.

So, to confirm whether there is a relationship between and , the hypotheses would be:

If is rejected and is accepted, then there is a relationship between and In other words, the relationship between and is statistically significant.

Then, to confirm whether there is a relationship between and , the hypotheses would be:

If is rejected and is accepted, then there is a relationship between and In other words, the relationship between and is statistically significant.

.

.

.

Finally, to confirm whether there is a relationship between and , the hypotheses would be:

If is rejected and is accepted, then there is a relationship between and In other words, the relationship between and is statistically significant.

Each slope coefficient must be tested separately. Luckily for us, Excel calculates the test statistic and two-tailed p-value for each slope coefficient and reports them in the regression output. We just need to compare the p-value to in each case. If the p-value , then we reject and conclude there is a statistically significant relationship between the particular and .

Remember: when the null hypothesis is rejected in these hypothesis tests, the conclusion is that the relationship between the given and is *statistically significant*.

**Hypothesis Test for the Overall Significance of the Regression Model**

The test reported in the ANOVA table in the regression output tests the overall significance of the regression model. It tests whether the model *considered as a whole* is significant. Whereas the tests discussed above test for the significance of the relationship between each individual and separately, the test shows whether is related to the , considered together.

Technically speaking, the test compares two different models for predicting the regression model, and the model in which the mean is used to predict (called the intercept-only model). If you can reject the null hypothesis in the test, then you know that using the regression to predict at each is an improvement over using the mean of to predict at each , and that the improvement is statistically significant.

1. **Formulating the Hypotheses:**

There is only one form of hypotheses, but the number of in it depends on how many IVs you have in your regression model.

|  |
| --- |
| **Hypotheses about the Overall Significance of the Regression Model** |
|  |
|  |
| Answers the question: |
| Whether the population slope coefficients are jointly different from zero.  That is to say, whether the regression model considered as a whole is statistically significant. |

1. **The Test Statistic:**

The test statistic is:

where

with (numerator) degrees of freedom

with (denominator) degrees of freedom

1. **Deciding whether or not to Reject :**

ANOVAs are always one-tailed, upper tail tests.

To calculate the p-value by Excel function, you would use =F.DIST.RT(

|  |  |
| --- | --- |
|  | **Always a One-Tailed, Upper Tail Test:** |
| **p-value approach:** | The upper-tailed is the upper tail probability of  If the then reject and accept  If the then do not reject is unsupported.  **NOTE: this is reported by Excel in the Regression Output as *Significance F* in the ANOVA table!!** |
| **Critical Value Approach:** | If , then reject and accept .  If , then do not reject is unsupported. |
| NOTES:   1. is the Test Statistic 2. is an upper tail Critical Value 3. The numerator degrees of freedom are , and the denominator degrees of freedom are | |

1. **Interpreting the test:**

|  |  |
| --- | --- |
| **When you:** | **The Interpretation is:** |
| **Reject** | At the significance level, we can conclude that overall regression model is statistically significant.  A significant relationship is present between and considered together.  Our regression model will do a better job predicting than using the mean to predict |
| **Do not reject** | At the significance level, we cannot conclude that the overall regression model is statistically significant. |