CHULALONGKORN UNIVERSITY
FACULTY OF ECONOMICS

2946653 Research Methods in Labour Economics
and Human Resource Management

EXERCISE 3

Hypothesis Testing

Recall the population regression function

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

Also, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Since we do not know the true value of β, we can hypothesise about the value of β and use statistical inference to test our hypothesis. Statistical estimator of β we obtain from the OLS estimation is $\widehat{\beta}$ where

$$\widehat{\beta}_j \sim \mathcal{N}(\beta_j, var(\widehat{\beta}_j)) \text{ for } j = 1, \ldots, k+1$$

We can standardise $\widehat{\beta}_j$,

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} \sim \mathcal{N}(0, 1)$$

Note that in calculating $se(\widehat{\beta}_j)$, we need to know the 'true' variance of ε which we do not. However, we can use $\widehat{\sigma}$ which we calculate from the OLS estimation. By replacing σ with $\widehat{\sigma}$, the distribution of $\widehat{\beta}_j$ becomes a student's t distribution with $N-k-1$ degrees of freedom.

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} \sim t_{N-k-1}$$

where $k+1$ is the number of the slope parameters (which are $\beta_1, \ldots, \beta_k$) plus the intercept term (β_0). We define the t statistic or t ratio of $\widehat{\beta}_j$ as follow

$$t_{\widehat{\beta}_j} = \frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)}$$

Suppose we want to test the following **null hypothesis**

$$H_0 : \beta_j = 0$$

i.e., we would like to test that X_j has no effect on the expected value of Y. In order to accept or reject H_0, we test it against the **alternative hypothesis**. The alternative hypothesis can be **one-tailed** (one-sided),

$$H_a : \beta_j > 0$$

or **two-tailed**,

$$H_a : \beta_j \neq 0$$

We then have to choose a **significance level** or the probability of rejecting H_0 when it is in fact true (**Type I error**). The **rejection rule** is that at the significance level $\alpha\%$, we will reject H_0 when

$$t_{\widehat{\beta}_j} > t_\alpha(N-k-1)$$

where $t_{\alpha}(N-k-1)$ is the **critical value** of t-statistic from the t-statistic table with $N-k-1$ degrees of freedom. Be careful about the significant level when you conduct the one-tailed and the two-tailed test!

We can also perform the overall test of significance by formulating the following hypothesis

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

against the alternative hypothesis

$$H_a : \text{ Not all slope coefficients are simultaneously zero}$$

The statistic we use to test the above null hypothesis is the F statistic,

$$F = \frac{ESS/df}{RSS/df} = \frac{ESS/k}{RSS/(N-k-1)}$$

The rejection rule is that we reject the null hypothesis at the significant level $\alpha\%$ when

$$F > F_{\alpha}(k, N-k-1)$$

where $F_{\alpha}(k, N-k-1)$ is the critical value of F at the α level of significance and k is the numerator degree of freedom and $N-k-1$ is the denominator degree of freedom.

Note that we can use the p value of t and F to perform the above hypothesis testing.

REMARK

Unless state otherwise, do not use the p value for hypothesis testing. Also, when performing the hypothesis testing, you have to draw the relevant distribution of the statistic ratio you are working on. Your diagram must contain the computed and critical values and the accpetance and rejection areas.

Question 1: Simple Regression Model (Cont.)

From Question 1 in Exercise 2

1. Test the null hypothesis for Model 1-4 that the slope coefficient is zero against the alternative hypothesis that the slope coefficient is greater than zero. State your hypothesis clearly. Specify the statistical test you use, the significant level, the degree of freedom, the associated critical value, and your decision rule. Also, provide your conclusion and comment on the result.

2. Test the null hypothesis for Model 1-4 that the slope coefficient is zero against the alternative hypothesis that the slope coefficient is not equal to zero. State your hypothesis clearly. Specify the statistical test you use, the significant level, the degree of freedom, the associated critical value, and your decision rule. Also, provide your conclusion and comment on the result.

3. From Model 5, test

$$\begin{aligned} H_0 &: \theta_j = 0 \text{ for } j = 1, \ldots, 4 \\ H_a &: \theta_j \neq 0 \text{ for } j = 1, \ldots, 4 \end{aligned}$$

   Specify the statistical test you use, the significant level, the degree of freedom, the associated critical value, and your decision rule. Also, provide your conclusion and comment on the result.

4. From Model 5, test

$$\begin{aligned} H_0 &: \theta_1 = \theta_2 = \theta_3 = \theta_4 = 0 \\ H_a &: \text{Not all slope coefficients are simultaneously zero} \end{aligned}$$

   Specify the statistical test you use, the significant level, the degree of freedom, the associated critical value, and your decision rule. Also, provide your conclusion and comment on the result.

Question 2: Exchange Rate and Imports (Cont.)

From Question 2 in Exercise 2, choose three models that have economic sense and perform the hypothesis testing that the slope coefficient is zero. You are free to choose the alternative hypothesis but you have to state it clearly. Do not forget to write down the computed statistic ratios, the critical values, the level of significance, the degree of freedom, your decision rule, and your conclusion.

Question 3: Liquor, Electgricity, GDP, and Exchange Rate (Cont.)

Using the two models you have estimated in Question 3 - Exercise 2, setup the null hypothesis for each model that the slope coefficient is zero. Test these hypotheses against the alternative hypothesis of your choice using the p value. Also, perform the overall test of significance on your multiple regression model using the p value. Specify all relevant information and state your conclusion.

Question 4: US Defence Budget

Consider the following model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \varepsilon_t$$

(Source: Gujarati 2003) where
Y_t = defence budget-outlay for year t, \$ billions
X_{1t} = GNP for year t, \$ billions
X_{2t} = US military sales/assistance in year t, \$ billions
X_{3t} = aerospace industry sales, \$ billions
X_{4t} = military conflicts involving more than 100,000 troops. This variable takes a value of 1 when 100,000 or more troops are involved but is equal to zero when that number is under 100,000. Using `dataset5.xls`:

1. Estimate the parameters of this model
2. Perform the t-test on each β. Clearly state your null and alternative hypotheses and other relevant information.
3. Comment on your results
4. Perform the overall test of significance. Clearly state your null and alternative hypotheses and other relevant information.
5. Use the p value approach to perform the t and F test. State all relevant information.