EC 320 Problem Set 3

Winter 2022

INSTRUCTIONS:

There are three questions in total. Please answer them all and show the steps of how you derived your answer to receive the full credit

1. Simple linear regression (15 points)

Suppose that the underlying population model is

$$Y_i = \beta_0 + \beta_1 X_i + u_i.$$

Let's suppose we use ordinary least squares (OLS) to estimate β_0 and β_1 with 20 observations, and find the estimates $\hat{\beta}_1 = 2$ and s.e.$(\beta_1) = 1$. You are considering two different hypothesis tests.

The first is a two-tailed test where

$$\begin{aligned} \mathrm{H}_0 &: \beta_1 = 0 \\ \mathrm{H}_a &: \beta_1 \neq 0. \end{aligned}$$

The second is a one-tailed test where

$$\begin{aligned} \mathrm{H}_0 &: \beta_1 = 0 \\ \mathrm{H}_a &: \beta_1 > 0. \end{aligned}$$

Suppose we set the significance level at 5%, i.e. $\alpha = 0.05$.

Below is some information about critical values from the t-distribution, where $t_x(df)$ gives t-score below which x% of the data falls when the degrees of freedom is equal to df. For example $t_0.975(100)$ is t-score below which 97.5% of the data falls when the degrees of freedom is 100.

$$\begin{aligned} t_{0.975}(20) &= 2.086 \\ t_{0.95}(20) &= 1.725 \\ t_{0.975}(18) &= 2.101 \\ t_{0.95}(18) &= 1.734 \end{aligned}$$

(a) Calculate t-statistics of the first hypothesis test.

(b) Run the first hypothesis test at 5% significance level and state your conclusion.

(c) Calculate t-statistics of the second hypothesis test.

(d) Run the second hypothesis test at 5% significance level and state your conclusion.

(e) Construct a 95% confidence interval of β_1.

2. Classical assumptions (10 points)

(a) Provide me a quick scatter plot that violates sample variation assumption.

(b) Provide me a quick scatter plot that violates homoskedasticity assumption.

(c) Explain in your own words what Gauss-Markov Theorem is. Make sure that you reference the classical assumptions.

3. Multiple linear regression (15 points)

Suppose that the underlying population model as well as the regression model is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

Using OLS, it is reported that

$$\hat{\beta}_0 = -2.62 \ , \ \text{s.e.}(\hat{\beta}_0) = 3.39$$
$$\hat{\beta}_1 = 0.98 \ , \ \text{s.e.}(\hat{\beta}_1) = 3.20$$
$$\hat{\beta}_2 = -1.26 \ , \ \text{s.e.}(\hat{\beta}_2) = 0.62$$

Information about critical values from the t-distribution is given below:

$$t_{0.995}(997) = 2.581$$
$$t_{0.99}(998) = 2.33$$
$$t_{0.975}(997) = 1.962$$
$$t_{0.95}(998) = 1.646$$

Suppose we conduct two-sided hypothesis test with null hypothesis set as $\beta_2 = 0$. The number of observations is 1000.

(a) Is $\hat{\beta}_2$ statistically significant at 5% significance level?

(b) Is $\hat{\beta}_2$ statistically significant at 1% significance level?

(c) Interpret $\hat{\beta}_1$ as in terms of units of Y and X_1.

(d) Suppose instead that we estimate the following regression model using OLS

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i,$$

and find $\hat{\beta}_0 = -5.09$ and $\hat{\beta}_1 = -5.32$. Find the omitted variable bias.

(e) Continuing on (d), what can you infer about the sign of the covariance between X_1 and X_2?