ETF2100/5910 Introductory Econometrics

Assignment 1, Semester 2, 2021

IMPORTANT NOTES:

- Type your answers using Microsoft Word or write your answers CLEARLY. You must submit a PDF file to Moodle. Other file formats are not accepted.
- Notation used in the assignment needs to be typed or written correctly and properly. Marks are also awarded for presentation.
- When doing calculation, keep at least 4 decimal in each step for precision. For final answer, 3 decimal point is sufficient, unless specified otherwise in the question.
- In this assignment, when you need to use t critical value, if the question does not specify, you can either find it using Eviews or use the statistical table. If the question specifies, use that method.
- This assignment is worth 15% of this unit's total mark.
- ETF2100 students must answer Question 1 and Question 2 Part 1.
- ETF5910 students must answer Question 1, Question 2 Part 1 and Question 2 Part 2.
- Total marks for ETF2100 and ETF5910 students are 40 and 50, respectively.
- Marks will be deducted for late submission on the following basis: 5 marks off for each day late, up to a maximum of 3 days. Assignments more than 3 days late will not be marked.

Question 1: For Both ETF2100 and ETF5910 (30 Marks)

You will use the data file **movie_revenue.csv** available on Moodle to answer this question. It contains data on box office revenue (i.e., revenue raised by ticket sales) and budget of movies released in the US during 2017-2019. Each observation is a movie. The variables are as follows:

- domestic_revenue: box office revenue in the US (in millions dollars)
- world_revenue : world-wide box office revenue (in millions dollars)
- budget: movie budget (in millions dollars)

We are interested in the relationship between movie's worldwide revenue (W) and its budget (B).

(a) Estimate the following linear regression model by least squares. Report the result in full (i.e., including s.e. and R-squared) and include your Eviews output in your answer. **(2 Marks)**

$$\mathrm{W}_i = \beta_1 + \beta_2 \mathrm{B}_i + e_i \tag{1}$$

(b) Interpret the estimated slope coefficient. Is this consistent with what you would expect the relationship to be? Explain briefly. **(3 Marks)**

(c) Interpret the estimated intercept. Does it make sense? Comment. **(3 Marks)**

(d) Interpret the R^2 you found in part (a). **(1 Marks)**

(e) Construct a 95% confidence interval for β_2. Interpret the confidence interval. **(3 Marks)**

(f) Predict worldwide box office revenue for a movie with a budget of 180,500,000 dollars. Give both a point predictor and a 90 % prediction interval. Use Eviews to find the exact t-critical value. Show your calculation and don't forget to interpret your answers in words. **(6 Marks)**

(g) What does the estimated variance (or standard error) of the forecast error tell us about our prediction? **(3 Marks)**

(h) How does each of the followings affect the size of the standard errors of the forecast error. Give a brief intuition (i.e., give an easy-to-understand explanation in words without the need of mathematical formula/proof).

  (i) Smaller variance of the error term **(2 Marks)**

  (ii) Larger value of $(B_0 - \bar{B})^2$, where B_0 is the value of the variable movie budget, for which we want to predict the corresponding value of movie revenue. **(2 Marks)**

(i) Test at the 1% level of significance if "an extra million dollar spent in making a movie is associated with an increase in the movie's worldwide box office revenue of MORE than three million dollars." Be sure to show all the steps used to conduct your test using the t-statistics approach. **(5 Marks)**

Question 2: Part 1 is for both ETF2100 and ETF5910; Part 2 is for ETF5910 ONLY

Consider the following simple regression model:

$$\mathrm{y}_i = \beta_1 + \beta_2 x_i + e_i \tag{2}$$

where $E(e_i \mid \mathbf{x}) = 0$, $var(e_i \mid \mathbf{x}) = \sigma^2$ and $cov(e_i, e_j \mid \mathbf{x}) = 0$ for all $i \neq j$. Suppose application of least squares rule to this equation with number of observations $N = 35$ yields $b_1 = 10, b_2 = 2$, and that we obtain $SSE = 12, R^2 = 0.7$ and

$$\widehat{cov}(b_1, b_2 \mid \mathbf{x}) = \begin{bmatrix} 4 & -0.1 \\ -0.1 & 0.25 \end{bmatrix} \tag{3}$$

Are the following statements true or false? Clearly indicate if each statement is true or false and support your choice with a brief explanation. You will get zero for that question if you only state true or false and do not give a brief explanation.

Part 1 for ETF2100 and ETF5910 (10 Marks)

(a) b_1 and b_2 are called least squares estimates because they minimize the squares $(b_1 - \beta_1)^2$ and $(b_2 - \beta_2)^2$. **(2 Marks)**

(b) The coefficient estimate b_2 is significantly different from zero at 5% significance level. **(2 Marks)**

(c) The total sum of squares (SST) is 125 **(2 Marks)**

(d) The error terms are heteroskedastic because the two variances on the diagonal of $\widehat{cov}(b_1, b_2 \mid \mathbf{x})$ are not equal to each other. **(2 Marks)**

(e) If economic theory suggests that x has a positive relationship with y, it makes sense to use a right-tail test when testing significance of β_2. **(2 Marks)**

Part 2 for ETF5910 ONLY (10 Marks)

(a) Because least squares estimators are unbiased, the estimates will always be equal to the true parameter values. **(3 Marks)**

(b) The estimated variance of the error term, $\hat{\sigma}^2$ is 0.6. **(2 Marks)**

(c) $\widehat{var}(b_1 - b_2 \mid \mathbf{x}) = 4$ **(3 Marks)**

(d) The assumption $var(e_i \mid \mathbf{x}) = \sigma^2$ implies that large errors are not more likely for some x_i than for others. **(2 Marks)**