

BEEM 011 Applied Econometrics 1

Assignment 1 (August 2021 Reassessment)

Submission Deadline: 10th August before 12pm

Direction: Answer all the questions and submit your exercise in PDF. Make sure to include your R scripts or computer codes as an appendix to your submission (You can copy your codes from the R script and paste them on the MS Word document). Please comment on each procedure so I would know what you are doing (or intend to do). Submission is via E-BART.

Tips:

- Install and load the following packages: “AER” and “sandwich”
- Label your plots and tables properly.
- Make sure you copy and paste all the necessary figures, tables and regression results.
- Be explicit about your assumptions. For example, what is your significance level when you do conduct hypothesis tests?
- Always make use of the evidence provided by your results (e.g., estimates, measures of fit, etc.).
- Be precise and concise.

(50 points)

1. Think of the situation of rolling a die and tossing a fair coin, with head having a value of 1 and tail having a value of 2. Let M denote the sum of the number of dots on the die and the outcome of the tossed coin (so M is a number between 2 and 8.)
 - a. In a table, list all of the possible outcomes for the random variable M together with its probability distribution and cumulative probability distribution. Using R, plot the probability and cumulative probability distribution. Label it properly. (10 points)
 - b. Calculate the expected value and the standard deviation for M . Calculate the minimum, maximum and median values. (10 points)
 - c. In R, set seed = 100. Then, draw a sample of size 10 with replacement from the list of possible outcomes in M . Calculate the mean. Repeat this for the 100 times and save it as an object called “sample”. Draw a histogram of the “sample”. (10 points)
 - d. Increase the number of times you repeat the procedure in (c) from 100 to 10,000 times. Save the results in an object called “sample2”. Draw a histogram of the “sample2”. Compare this histogram to the one you drew in (c). What do you observe? What can you conclude from this exercise. (20 points).

(20 points)

2. Using R, plot the following (overlay each on the other):
 - i. Standard normal distribution (limit your x values between -4 and 4).
 - ii. t density for 1 degrees of freedom

- iii. t density for 2 degrees of freedom
 - iv. t density for 50 degrees of freedom
- Label your plot properly. (10 points)

What do you observe? What can you conclude from this exercise? (10 points)

(30 points)

3. In R, load the “AER” package in the library and load the dataset “CPS1988”) which is a cross-section data originating from the March 1988 Current Population Survey by the US Census Bureau. The dataset has 28,155 observations and includes the following variables: *wage* (wage, in dollars per week); *experience* (number of years of potential work experience); and *parttime* (factor, “Does the individual work part-time?”).
 - a. Describe the distribution of wage for (1) part-time and (2) full-time workers (those that are not part-time workers). Use summary statistics, such as the mean, median, variance, and skewness to compare their distribution. (10 points)
 - b. Run a linear regression of wage on experience only (i.e., wage is the dependent variable) for (1) part-time and (2) non-part-time workers (i.e., part-time = “no”). Present your results. Interpret the coefficients and measures of fit. Be explicit with your unit of measurement. What can you conclude from this exercise? (20)