Homework 8

Multiple Regression Model - OLS Asymptotic (10 points)

Instruction:

- This HW must be done in Rmarkdown!
- Please submit both the .rmd and the Microsoft word files. (Do not submit a PDF or any other image files as the TAs are going to give you feedback in your word document)
- Name your files as: HW8-groupnumber-name
- All the HW assignments are individual work. However, I highly encourage you to discuss it with your group members.
- Late homework assignments will not be accepted under any circumstances.

Problems

Question 1 Suppose that the model

$$pctstck = \beta_0 + \beta_1 funds + \beta_2 risktol + u$$

satisfies the first four Gauss-Markov assumptions, where pctstck is the percentage of a worker's pension invested in the stock market, funds is the number of mutual funds that the worker can choose from, and risktol is some measure of risk tolerance (larger risktol means the person has a higher tolerance for risk). If funds and risktol are positively correlated, what is the inconsistency in $\tilde{\beta}_1$, the slope coefficient in the simple regression of pctstck on funds? i.e. dropping $risktol$.

Hint: you need to say if there is positive or negative asymptotic bias in $\tilde{\beta}_1$

Computer Exercises

Question 2 Use the data in WAGE1 for this exercise.

(i) Estimate the equation

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u.$$

Save the residuals and plot a histogram of residuals.

(ii) Repeat part (i), but with log(wage) as the dependent variable.

(iii) Would you say that Assumption MLR.6 is closer to being satisfied for the level-level model or the log-level model?

Question 3 Use the data in ECONMATH to answer this question.

(i) Logically, what are the smallest and largest values that can be taken on by the variable score? What are the smallest and largest values in the sample?

(ii) Consider the linear model

$$score = \beta_0 + \beta_1 colgpa + \beta_2 actmth + \beta_3 acteng + u.$$

Why cannot Assumption MLR.6 hold for the error term u? What consequences does this have for using the usual t statistic to test $H_0 : \beta_3 = 0$?

Hint: if the dependent variable is not normally distributed, then error terms u cannot be normally distributed either!

(iii) Estimate the model from part (ii) and obtain the t statistic and associated p-value for testing $H_0 : \beta_3 = 0$. How would you defend your findings to someone who makes the following statement: "You cannot trust that p-value because clearly the error term in the equation cannot have a normal distribution."