Exercise Session - Problem Set 6

Multiple Regression I

Problem 1 (Wooldridge, Example 3.6 of Chapter 3)

In a study relating college grade point average to time spent in various activities, you distribute a survey to several students. The students are asked how many hours they spend each week in four activities: studying, sleeping, working, and leisure. Any activity is put into one of the four categories, so that for each student the sum of hours in the four activities must be 168.

(a) In the model

$$GPA = \beta_0 + \beta_1 \cdot study + \beta_2 \cdot sleep + \beta_3 \cdot work + \beta_4 \cdot leisure + u \tag{1}$$

does it make sense to hold **sleep**, **work** and **leisure** fixed, while changing **study**?

(b) Does this model in (1) violate any of the Gauß-Markov assumptions?

(c) How could you reformulate the model so that its parameters have a useful interpretation and it satisfies the above-mentioned assumption?

Problem 2 (Wooldridge, Example 3.8 of Chapter 3)

Which of the following can cause OLS estimators to be biased?

(i) Heteroskedasticity? (This term means: Absence of homoscedasticity).

(ii) Omitting an important variable?

(iii) Including irrelevant variables in the model?

(iv) A sample correlation coefficient of 0.95 between two independent variables both included in the model?

Problem 3 (Wooldridge, Example 3.5 of Chapter 3)

Consider the multiple regression model with three independent variables under the Gauß-Markov assumptions,

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + u \tag{2}$$

You are interested in estimating the sum of the parameters on x_1 and x_2, $\beta_{1+2} = \beta_1 + \beta_2$.

(i) Show that $\hat{\beta}_{1+2} = \hat{\beta}_1 + \hat{\beta}_2$ is an unbiased estimator of β_{1+2}.

(ii) Explain how you would calculate the variance of $\hat{\beta}_{1+2}$.

Problem 4 (Wooldridge, Example 3.9 of Chapter 3)

Suppose you are interested in estimating the ceteris paribus relationship between y and x_1. For this purpose, you can collect data on two control variables, x_2 and x_3. Let $\tilde{\beta}_1$ be the simple regression estimate from y on x_1, and let $\hat{\beta}_1$ be the multiple regression estimate from y on x_1, x_2, and x_3.

(a) If x_1 is highly correlated with x_2 and x_3 in the sample, and x_2 and x_3 have large partial effects on y, would you expect $\tilde{\beta}_1$ and $\hat{\beta}_1$ to be similar or very different?

(b) If x_1 is almost uncorrelated with x_2 and x_3, but x_2 and x_3 are highly correlated, will $\tilde{\beta}_1$ and $\hat{\beta}_1$ tend to be similar or very different?

(c) If x_1 is highly correlated with x_2 and x_3, and x_2 and x_3 have small partial effects on y, would you expect the standard error of $\tilde{\beta}_1$ or the standard error of $\hat{\beta}_1$ to be smaller?

(d) If x_1 is almost uncorrelated with x_2 and x_3, x_2 and x_3 have large partial effects on y, and x_2 and x_3 are highly correlated, would you expect the standard error of $\tilde{\beta}_1$ or the standard error of $\hat{\beta}_1$ to be smaller?