Exercise Session - Problem Set 3

The Simple Linear Regression Model

Problem 1 (Wooldridge, Example 2.2, p. 61)

The following table contains the ACT (achievement examination for college admissions in the US) scores and the GPA (grade point average) for 8 college students. Grade point average is based on a four–point scale and has been rounded to one digit after the decimal.

Table 1: GPA and ACT

Student	GPA	ACT
1	2.8	21
2	3.4	24
3	3.0	26
4	3.5	27
5	3.6	29
6	3.0	25
7	2.7	25
8	3.7	30

a) Estimate the relationship between GPA and ACT using OLS, i.e., obtain the OLS estimates of the equation

$$\widehat{GPA} = \widehat{\beta}_0 + \widehat{\beta}_1 ACT. \tag{1}$$

Comment on the direction of the relationship. Does the intercept have a useful interpretation here? How much higher is GPA predicted to be, if the ACT score is increased by 5 points?

b) What is the predicted value of GPA when ACT = 20?

c) How much of the variation in GPA for these 8 students is explained by ACT? Explain.

Problem 2 (Wooldridge, Example 2.1, p. 61)

In the simple linear regression model $y = \beta_0 + \beta_1 x + u$, suppose $\mathrm{E}(u) = \alpha_0 \neq 0$. Show that the model can always be rewritten with the same slope, but a new intercept and error, where the new error has zero mean.

Problem 3

In the simple linear model, what happens to the OLS intercept and slope estimates if each observation of the dependent (independent) variable is multiplied by a constant c? (Economically, this amounts to a change of the unit of measurement of the variables in question.)

Problem 4 (Wooldridge, Problem 2.3, page 62)

Let **kids** denote the number of children ever born to a woman, and let **edu** denote years of education for the woman. A simple model relating fertility to years of education is

$$kids = \beta_0 + \beta_1 \cdot edu + u, \tag{2}$$

where u is the unobserved error.

a) What kind of factors are contained in u? Are these likely to be correlated with level of education?

b) Will a simple regression analysis uncover the ceteris paribus effect of education on fertility? Explain.

Problem 5

Consider a random sample of size n, i.e., $\{(x_i, y_i), i = 1, \ldots, n\}$, from the simple linear model $y = \beta_0 + \beta_1 x + u$ with the assumptions $\mathrm{E}(u|x) = 0$ and $\mathrm{Var}(u|x) = \sigma^2$.

i) Find $\mathrm{Var}(\widehat{\beta}_0)$, $\mathrm{Var}(\widehat{\beta}_1)$, and $\mathrm{Cov}(\widehat{\beta}_0, \widehat{\beta}_1)$.

ii) Show that

$$\widehat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \widehat{u}_i^2 \tag{3}$$

is an unbiased estimator of σ^2. You may, for example, first show

$$\mathrm{Var}(\widehat{y}_i) = \frac{\sigma^2}{n} \left[1 + \frac{(x_i - \bar{x})^2}{s_x^2} \right], \quad s_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2, \tag{4}$$

and then use $\widehat{u}_i = y_i - \widehat{y}_i = \epsilon_i - [\widehat{y}_i - \mathrm{E}(\widehat{y}_i)]$, where $\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$.