ECO3121 INTRODUCTORY ECONOMETRICS
Problem Set #1 Solution

TA group

October 13, 2024

Question 1

1. OLS estimator $\hat{\beta}_0$ solves

$$\min_{\hat{\beta}_0} \sum_{i=1}^{n} (y_i - \hat{\beta}_0)^2$$

F.O.C with respect to $\hat{\beta}_0$ is

$$-2\sum_{i=1}^{n} y_i + 2\sum_{i=1}^{n} \hat{\beta}_0 = 0$$
$$\Rightarrow \hat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

2. OLS estimator $\hat{\beta}_1$ solves

$$\min_{\hat{\beta}_1} \sum_{i=1}^{n} (y_i - \hat{\beta}_1 x_i)^2$$

F.O.C with respect to $\hat{\beta}_1$ is

$$-2\sum_{i=1}^{n} x_i y_i + 2\sum_{i=1}^{n} \hat{\beta}_1 x_i^2 = 0$$
$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

Question 2

1. Positive sign.

2. (See Stata code) The estimated result is

$$\underset{\quad}{\widehat{wage}_i} = \underset{(2004.897)}{57405.06} + \underset{(193.9109)}{17910.03}\ education_i \tag{1}$$

$$R^2 = 0.1225$$

Yes. $\hat{\beta}_1 = 17910.03$ agrees with my prediction.

3. $\hat{\beta}_1$: one more year of education increases wages by 17910.03, holding all other factors (in μ) fixed.
   $\hat{\beta}_0$: a person with no education has a predicted 57405.06 wage, holding all other factors (in μ) fixed.

4. (See Stata code) The residuals sum to -46.60354, which is near 0.

5. Take education=10 in Eq.1, we can get $\widehat{wage} = 57405.06 + 17910.03 \times 10 = 236505.36$.

6. $R^2 = 0.1225$, which means 12.25% of the sample variation in wages is explained by education.
   For example, rural residents find it harder to afford intuition. There are fewer schools in rural areas.

7. (See Stata code) The estimated results are

$$\underset{}{\widehat{log(wage)}_i} = \underset{(0.0062188)}{11.41769} + \underset{(0.0006015)}{0.073147}\ education_i$$

   Interpretation: one more year of education increases wages by 7.3147%, holding all other factors fixed.

$$\underset{}{\widehat{wage}_i} = \underset{(3690.619)}{-49308.2} + \underset{(1655.687)}{128514.6}\ log(education)_i$$

   Interpretation: a 1% increase in years of education increases wages by 1285.146, holding all other factors fixed.

$$\underset{}{\widehat{log(wage)}_i} = \underset{(0.0113235)}{10.95083} + \underset{(0.0050799)}{0.5408492}\ log(education)_i$$

   Interpretation: a 1% increase in years of education increases wages by 0.5408492%, holding all other factors fixed.

8. SLR.4 is $E(\mu_i|education_i) = 0$, which implies that an individual's year of education is uncorrelated with any unobservable factors.

   It doesn't hold because one's education level might be correlated with his ability, intelligence, parents' education level, and so on, which are not included in the regression model.

Question 3

1. Error term μ_i refers to factors other than fertilizer that affect yields, such as water, pesticides, labor input, etc.

   The reasons are that plots differ in temperature, sunlight, planting Density, etc.

2. Yes. Because the amount of fertilizer is randomly assigned among plots.

3. (a) Yes. Because SLR.1-4 hold.
   (b) $E[\mu_i|X_i]$ is unlikely to hold, so $\hat{\beta}_1$ is a biased estimator.

(c) [1] I need data about soil quality for plots, and run a regression as follows

$$soil\ quality_i = \beta_0 + \beta_1 X_i + \mu_i \tag{2}$$

If I find $\hat{\beta}_1$ in the estimated model Eq.2 is significantly different from 0, the concern is true.

[1] *Notes*: Some students argue that we can verify the concern by running a multivariate regression where we additionally control for the soil quality, and provide a justification based on the omitted variable formula. However, two drawbacks of this approach make it unsatisfactory: First, the change in $\hat{\beta}_1$ between the two specifications doesn't tell us whether the fertilizer input is **significantly** correlated with soil quality, that is, we cannot do any statistical inference. Second, the multivariate regression approach is the way to solve the concern, rather than verify it. In the multivariate regression, we still focus on the change in $\hat{\beta}_1$, the partial effect of fertilizer on yield. However, what we want to verify is the relationship between fertilizer and soil quality, so it doesn't answer the question we require.