ECO3151D

ASSIGNMENT 1 **Solution**

Dr. M. Rafiquzzaman Winter 2015

---

1. In order to derive the OLS estimator $\hat{\beta}_0$ and $\hat{\beta}_1$, please go through class notes and/ or chapter 2 of the textbook.

To show that (1) and (2) are equivalent, it is sufficient to show that the numerators of (1) and (2) are same, that is,

$$\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) = \sum_{i=1}^{n}(X_i - \bar{X})Y_i$$

$$\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) = \sum_{i=1}^{n}\left[(X_i - \bar{X})Y_i - (X_i - \bar{X})\bar{Y}\right] = \sum_{i=1}^{n}(X_i - \bar{X})Y_i - \bar{Y}\sum_{i=1}^{n}(X_i - \bar{X}) = \sum_{i=1}^{n}(X_i - \bar{X})Y_i$$

, because $\sum_{i=1}^{n}(X_i - \bar{X}) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\bar{X} = n\bar{X} - n\bar{X} = 0$. That is, sum of the deviations from mean is zero.

2. $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \dfrac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}$. Now $\bar{X} = \dfrac{1700}{10} = 170$; $\bar{Y} = \dfrac{1110}{10} = 111$

$$\hat{\beta}_1 = \frac{205500 - (10)(170)(111)}{322000 - (10)(170)^2} = 0.5091; \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X} = 111 - (0.5091)(170) = 24.4530$$

The estimated regression line therefore is: $\hat{Y}_i = 24.4530 + 0.5091X_i$

$\hat{\beta}_1$ is the marginal propensity to consume. In this example, it indicates that \$1 increase in weekly income increases the weekly consumption expenditure amounts to about 51 cents.

$$R^2 = \hat{\beta}_1^2 \frac{\sum(X_i - \bar{X})^2}{\sum(Y_i - \bar{Y})^2} = \hat{\beta}_1^2 \frac{\sum X_i^2 - n\bar{X}^2}{\left(\sum Y_i^2 - n\bar{Y}^2\right)} = (0.5091)^2\left[\frac{322000 - (10)(170)^2}{132100 - (10)(111)^2}\right] = 0.9621$$

The value of R^2 of 0.9621means that about 96 percent of the variation in the weekly consumption expenditure is explained by income.

To Calculate $Se(\hat{\beta}_0)$ and $Se(\hat{\beta}_1)$

$$Se(\hat{\beta}_0) = \sqrt{\frac{\hat{\sigma}^2 n^{-1} \sum x_i^2}{\sum (x_i - \bar{x})^2}} \quad \text{— (1)}$$

$\sum x_i^2 = 322{,}000$

$\sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 322{,}000 - (10)(170)^2$

$= 322{,}000 - 289{,}000 = 33{,}000$

$\hat{\sigma}^2 = ? = \frac{1}{n-2} \sum \hat{u}_i^2 = \frac{SSR}{n-2}$

Recall $R^2 = 1 - \frac{SSR}{SST} \Rightarrow \frac{SSR}{SST} = 1 - R^2 = 1 - 0.9621$

$\Rightarrow \frac{SSR}{SST} = 0.0379 \Rightarrow SSR = (0.0379)(SST)$ — ②

Now, $SST = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2$

$= 132{,}100 - (10)(111)^2 = 132{,}100 - 123{,}210$

$= 8890$

Substituting into (2), we have

$SSR = (0.0379)(8890) = 336.931$

Then, $\hat{\sigma}^2 = \frac{SSR}{n-2} = \frac{336.931}{10-2} = \frac{336.931}{8} = 42.1164$

From (1)

$$Se(\hat{\beta}_0) = \sqrt{\frac{(42.1164)(322{,}000)}{(10)(33{,}000)}} = \sqrt{\frac{(42.1164)(322)}{(10)(33)}}$$

$$\Rightarrow Se(\hat{\beta}_0) = \sqrt{\frac{13561.481}{330}} = \sqrt{41.0954} = 6.4106$$

$$Se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\Sigma(x_i - \bar{x})^2}} = \sqrt{\frac{42.1164}{33,000}}$$

$$= \sqrt{0.0012763} = 0.0357$$

3(i). The regression equation of $c_1 y_i$ on $c_2 x_i$ may be written as: $c_1 y_i = \tilde{\beta}_0 + \tilde{\beta}_1 (c_2 x_i) + v_i$

The regression equation of y_i on x_i is written as $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + u_i$

Using OLS method, $\tilde{\beta}_1 = \dfrac{\sum (c_1 y_i - c_1 \bar{y})(c_2 x_i - c_2 \bar{x})}{\sum (c_2 x_i - c_2 \bar{x})^2} = \dfrac{\sum c_1 (y_i - \bar{y}) c_2 (x_i - \bar{x})}{\sum c_2^{\,2} (x_i - \bar{x})^2}$

$$= \frac{c_1 c_2 \sum (y_i - \bar{y})(x_i - \bar{x})}{c_2^{\,2} \sum (x_i - \bar{x})^2} = \frac{c_1 \sum (y_i - \bar{y})(x_i - \bar{x})}{c_2 \sum (x_i - \bar{x})^2} = \left(\frac{c_1}{c_2}\right)\hat{\beta}_1$$

$$\tilde{\beta}_0 = c_1 \bar{y} - \tilde{\beta}_1 c_2 \bar{x} = c_1 \bar{y} - \left(\frac{c_1}{c_2}\right)\hat{\beta}_1 c_2 \bar{x} = c_1 \left(\bar{y} - \hat{\beta}_1 \bar{x}\right) = c_1 \hat{\beta}_0$$

3(ii). Let the regression equation of $(c_1 + y_i)$ on $(c_2 + x_i)$ be $(c_1 + y_i) = \tilde{\beta}_0 + \tilde{\beta}_1 (c_2 + x_i) + w_i$. (Note that, because c_1 and c_2 are constants, $\bar{c}_1 = c_1$ and $\bar{c}_2 = c_2$. Applying OLS method,

$$\tilde{\beta}_1 = \frac{\sum [(c_1 + y_i) - (c_1 + \bar{y})][(c_2 + x_i) - (c_2 + \bar{x})]}{\sum [(c_2 + x_i) - (c_2 - \bar{x})]^2} = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = \hat{\beta}_1$$

$$\tilde{\beta}_0 = \left[(c_1 + \bar{y}) - \tilde{\beta}_1 (c_2 + \bar{x})\right] = c_1 + \left(\bar{y} - \tilde{\beta}_1 \bar{x}\right) - c_2 \tilde{\beta}_1 = c_1 + \left(\bar{y} - \hat{\beta}_1 \bar{x}\right) - c_2 \hat{\beta}_1$$

$$= \hat{\beta}_0 + c_1 - c_2 \hat{\beta}_1$$

3(iii). If $y_i = \beta_0 + \beta_1 x_i + u_i$, $i = 1,2,......,n$, then OLS estimates of the parameters are:

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$, ... (3.1), and

$\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$, ...(3.2)

Define $X_i = \dfrac{x_i - a}{b}$, ...(3.3)

The new regression equation is: $y_i = \beta_0 + \beta_1 X_i + u_i$, ... (3.4)

The OLS estimates of the parameters of the new regression (3.4) are:

$\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1 \bar{X}$, ... (3.5)

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}, \text{ .... (3.6)}$$

Substituting (3.3) into (3.6), we have

$$\tilde{\beta}_1 = \frac{\sum (y_i - \bar{y})(\frac{x_i - a}{b} - \frac{\bar{x} - a}{b})}{\sum (\frac{x_i - a}{b} - \frac{\bar{x} - a}{b})^2} =$$

$$\frac{\frac{1}{b}\sum (y_i - \bar{y})(x_i - a - \bar{x} + a)}{\frac{1}{b^2}\sum \ (x_i - a - \bar{x} + a)^2} = \frac{\frac{1}{b}\sum (y_i - \bar{y})(x_i - \bar{x})}{\frac{1}{b^2}\sum (x_i - \bar{x})^2} = (b)\frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = b\hat{\beta}_1$$

Thus, $\tilde{\beta}_1 = b\hat{\beta}_1$

From (3.5), $\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1\bar{X} =$

$$\bar{y} - b\hat{\beta}_1\bar{X} = \bar{y} - b\hat{\beta}_1(\frac{\bar{x} - a}{b}) = \bar{y} - \hat{\beta}_1(\bar{x} - a) = \bar{y} - \hat{\beta}_1\bar{x} + a\hat{\beta}_1 = \hat{\beta}_0 + a\hat{\beta}_1$$

Thus, $\tilde{\beta}_0 = \hat{\beta}_0 + a\hat{\beta}_1$

4. (i) A larger rank of a law school means that the school has less prestige; this lowers starting salaries. For example, a rank of 100 means there are 99 schools thought to be better.

(ii) $\beta_1 > 0$, $\beta_2 > 0$. Both *LAST* and *GPA* are measures of the quality of the entering class. No matter where better students attend law school, we expect them to earn more, on average. $\beta_3 > 0$, $\beta_4 > 0$. The number of volumes in the law library and the tuition cost are both measures of the school quality. (Cost is less obvious than library volumes, but should reflect quality of the faculty, physical plant, and so on.)

(iii) This is just the coefficient on *GPA*, multiplied by 100: 0.248x100 = 24.8%

(iv) This is an elasticity: a one percent increase in library volumes implies a 0.095% increase in predicted median starting salary, other things equal.

(v) It is definitely better to attend a law school with a lower rank. Δ [log (*salary*)] = -0.0033 Δ(*rank*). If law school A has a ranking 20 less than law school B, the predicted difference in starting salary is 100(0.0033)(20) = 6.6% higher for law school A.