Chapter 2: Ordinary Least Squares

Revised on September 25, 2007

1 Estimating Single-Independent-Variable Models with OLS

We now discuss how to estimate the unknown parameters β_0 and β_1. Estimating these two unknown parameters by choosing β_0 and β_1 in

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i \tag{1}$$

is equivalent to fitting a straight line to the data. Visually inspecting the data and fitting a line is not a good idea as it is too arbitrary and imprecise. We choose β_0 and β_1 in such a way that the residual

$$e_i = Y_i - \hat{Y}_i \tag{2}$$
$$Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i) \tag{3}$$

is as "small" as possible under a certain criterion. Note that the residual, e_i, is the vertical deviation of the corresponding data point from the fitted line. There are at least three different criteria that one could use:

(1) **Minimizing the sum of the residuals**: The problem of minimizing the sum of residuals is not well defined, we can make the sum of residuals as small as we please if we are allowed to make the sum negative.

(2) **Minimizing the sum of absolute values of the residuals**: This criterion is technically difficult to use because the absolute value function is not

differentiable at zero. Some people, however, still use it. We will not use it in this course.

(3) **Minimizing the sum of the squared residuals**: This method of choosing $\hat{\beta}_0$ and $\hat{\beta}_1$ is called the *least squares principle*. This criterion is algebraically easy to handle and results in a unique line of fit. This is the criterion we shall use for finding out the best line of fit. The estimators of β_0 and β_1 based on this method are called the *Ordinary Least Squares* (OLS) estimators.

Algebraically, the least squares principle tells us to choose $\hat{\beta}_0$ and $\hat{\beta}_1$, so that we minimize *residual sum of squares (RSS)*

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)^2 \qquad (4)$$

Thus we see that the RSS, which we want to minimize, depends upon the $\hat{\beta}_0$ and $\hat{\beta}_1$ we choose. We minimize the RSS by choosing $be\hat{t}a_0$ and $\hat{\beta}_1$ appropriately.

1.1 How Does OLS Work?

Using calculus, we obtain the following formula:

$$\hat{\beta}_1 = \frac{[(\sum_{i=1}^{N}(X_i - \bar{X})) \cdot (Y_i - \bar{Y})]}{\sum_{i=1}^{N}(X_i - \bar{X})^2}. \tag{5}$$

$$\hat{\beta}_0 = \bar{Y} - \beta_2 \bar{X} \qquad (5.16). \tag{6}$$

where $\bar{X}$ is the sample average of X, or $\sum X_i/N$, and $\bar{Y}$ is the sample average of Y, or $\sum Y_i/N$. Note that for each different data set, we will get different estimates of β_0 and β_1.

1.2 Illustrations of OLS Estimation

Let us see how we apply the OLS formulae in Equations (5) and (6) to Table 1.B. in order to obtain OLS estimates. One way is to use the following type of a table:

Table 2.A The OLS estimates from Steve's Data

Raw Data			Calculations			
i	X_i	Y_i	$Y_i - \bar{Y}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	25	44.5	-4.67	-14	196	65.38
2	40	49	-0.17	1	1	-0.17
3	52	54	4.83	13	169	62.79
Sum	117	147.5	-0.01	0	366	128
Mean	39	49.17				

Note that the sum of the column (4) is not exactly zero because of the rounding error, but that it is close to zero. Now applying Equation (5) to these results: divide the sum of column (7) by the sum of column (6) to compute

$$\hat{\beta}_1 = 128/366 = 0.349727 \approx 0.35. \tag{7}$$

Then use Equation (6) to get

$$\hat{\beta}_0 = 49.17 - 0.35 \cdot 39 = 35.52. \tag{8}$$

These results are often reported in the following format:

$$\hat{Y}_i = 35.52 + 0.35 X_i \tag{9}$$

Exercise (OLS for the data in Table 1.B): Interpret the estimated regression coefficients. Plot the sample regression. Plot the population regression.

Exercise (OLS for the data in Table 1.C): Compute the OLS regression coefficient estimates, using the data in Table 1.C. Report the results in the format given in Equation(9). Interpret the estimated regression coefficients. Plot the sample regression. Plot the population regression.

1.3 Examples

There are two types of data that are often used in economics. One type is *time series data*, which are collected over a period of time, such as the data on GDP, consumption, etc. The other type is *cross-section* (or *cross-sectional*) data, which are data on one or more variables collected at one point in time. Steve's ice cream stand data in Table 1.B is time series. The example in Section 2.1.3 of Studenmund's book is cross-section.

1.4 Notable Features of OLS Estimators

Note the following features of the OLS estimators.

(1) The estimated regression obtained by the OLS passes through the sample mean values of X and Y. From Equation (6), we have

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \bar{X}. \tag{10}$$

(2) The sample mean value of the residuals is always equal to zero (except for the rounding error):

$$\bar{e} = \frac{1}{N} \sum_{i=1}^{N} e_i = 0. \tag{11}$$

This is because

$$\bar{e} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = \bar{Y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{X} = 0 \tag{12}$$

from Equation (6).