Applied Econometrics, OLS

Simple regression model: $y = \beta_0 + \beta_1 x_1 + u$

To calculate OLS estimate of the slope:

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{sample\ covariance\ between\ x\ and\ y}{sample\ variance\ of\ x}$$

To calculate OLS estimate of the intercept:

$$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x}$$

Recalling from probability and statistics:

Given n numbers $\{x_i: i=1, 2, \ldots, n\}$

(1) The average or mean of x is

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

(A.5 in Wooldridge Appendix A)

(2) The sum of squared deviations from the mean

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i{}^2 - n(\bar{x})^2$$

(A.7 in Wooldridge Appendix A)

(3) The sample variance

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(C.5 in Wooldridge Appendix C)

(4) The sample standard deviation is the square root of the sample variance

Given a data set on two variables, $\{(x_i, y_i): i=1, 2, \ldots, n\}$

(5) The sum of the products of deviation in x and deviation in y

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})y_i = \sum_{i=1}^{n} x_i y_i - n(\bar{x})(\bar{y})$$

(A.8 in Wooldridge Appendix A)

(6) The sample covariance

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

(C.14 in Wooldridge Appendix C)

PRACTICE EXERCISES

Estimate the relationship between y and x using OLS, i.e., calculate the OLS estimate of the intercept and slope in the equation

$y = \beta_0 + \beta_1 x_1 + u$

If your answers are not whole numbers, then round them to the nearest hundredth.

Exercise 1

where the summary statistics for a sample are as follows:

$\bar{x} = 30; \bar{y} = 20; \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 50; \sum_{i=1}^{n}(x_i - \bar{x})^2 = 25$

Exercise 2

where the summary statistics for a sample are as follows:

$\bar{x} = 10; \bar{y} = 7$; sample covariance between x and y = 3; sample variance of x = 6

Exercise 3

where the data are as follows:

x	y
1	2
2	4
3	4
4	7

Exercise 4

where the data are as follows:

x	y
2	6
3	2
7	7
8	4
5	4