LECTURE 10

SIMPLE REGRESSION MODEL - II

Outline of today's lecture:

I. Mean and Variance of the Dependent Variable Y

The dependent variable Y has mean

$$E(Y_t) = \beta_0 + \beta_1 X_t$$

and variance

$$Var(Y_t) = E\left[Y_t - E(Y_t)\right]^2 = E(u_t^2) = \sigma^2$$

1. *Let us show that the mean of Y_t is $\beta_0 + \beta_1 X_t$.* →

By definition the mean of Y_t is its expected value.

Given that $Y_t = \beta_0 + \beta_1 X_t + u_t$. Taking the expected values we get

$$E(Y_t) = E\left[\beta_0 + \beta_1 X_t + u_t\right]$$
$$E(Y_t) = E\left[\beta_0 + \beta_1 X_t\right] + E\left[u_t\right]$$

Given that β_0 and β_1 are parameters and by Assumption 2 the values of X_t's are a set of fixed numbers (in the process of hypothetical sampling)

$$E\left[\beta_0+\beta_1 X_t\right]=\beta_0+\beta_1 X_t$$

Furthermore, by Assumption 3, $E(u_t)=0$.

Therefore,

$$E(Y_t)=\beta_0+\beta_1 X_t$$

2. *Let us now show that the variance of Y_t is σ^2*

$$Var(Y_t)=E\left[Y_t-E(Y_t)\right]^2$$
$$Var(Y_t)=E\left[\beta_0+\beta_1 X_t+u_t-\beta_0+\beta_1 X_t\right]^2$$
$$Var(Y_t)=E\left[u_t\right]^2$$

By Assumption 4, the u_t's are homoscedastic, that is, they have the constant variance σ^2

$$Var(Y_t)=E\left[u_t\right]^2=\sigma^2$$

II. Ordinary Least Squares (OLS) Estimation

- The two-variable *population regression function* is given by

$$Y_t=\beta_0+\beta_1 X_t+u_t,$$

  but we do not observe it.

- Hence we estimate it from the *sample regression function*

$$Y_t = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 X_t}_{\hat{Y}_t} + \hat{u}_t .$$

or $Y_t = \hat{Y}_t + \hat{u}_t$.

- We can rewrite the sample regression function as

$$\hat{u}_t = Y_t - \hat{Y}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t .$$

- In other words, the residuals are the differences between the actual and the estimated Y_t values.

- With T observations, we want to choose $\hat{\beta}_0$ and $\hat{\beta}_1$ such that the sum of the residuals is minimized: $\sum_{t=1}^{T} \hat{u}_t = \sum_{t=1}^{T} (Y_t - \hat{Y}_t)$.

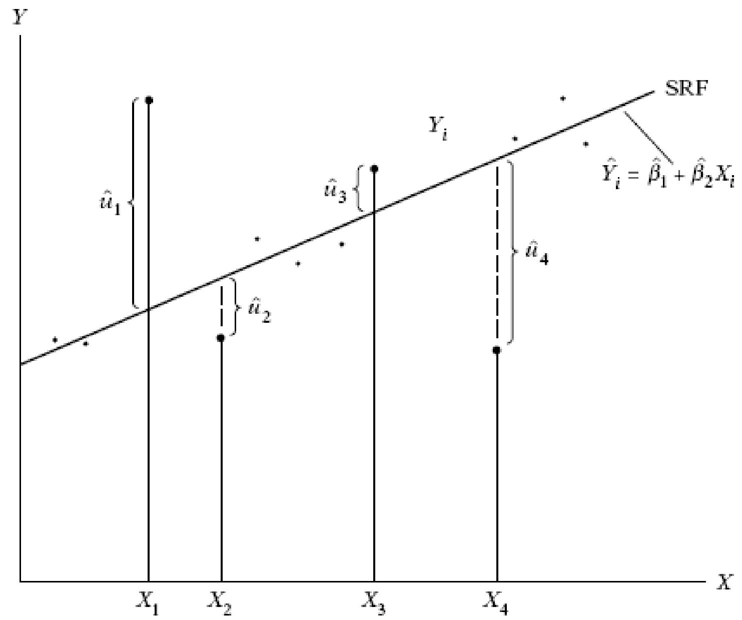

FIGURE 3.1 Least-squares criterion.

- this turns out not to be a very good rule because some residuals are negative and some are positive (and they would cancel each other), and

- all residuals have the same weight (importance) even though some are small and some are large.

  ✓ To overcome these problems, we use the squares of the residuals instead of their own values.

Ordinary Least Squares (OLS) criterion:

$$\text{Minimize } \sum \hat{u}_t^{\,2} = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t)^2$$
$$\text{wrt } \quad \hat{\beta}_1 \text{ and } \hat{\beta}_2$$

The necessary condition for a minimum is that the first derivatives of the function be equal to zero.

Partial differentiation yields

$$\frac{\partial(\sum \hat{u}_t^{\,2})}{\partial \hat{\beta}_0} = -2\sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) = 0 \qquad (1)$$
$$\frac{\partial(\sum \hat{u}_t^{\,2})}{\partial \hat{\beta}_1} = -2\sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) X_t = 0 \qquad (2)$$

From (1)

$$\sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) = 0$$
$$\sum_{t=1}^{T} Y_t - T.\hat{\beta}_0 - \hat{\beta}_1 \sum_{t=1}^{T} X_t = 0$$
$$T.\hat{\beta}_0 = \sum_{t=1}^{T} Y_t - \hat{\beta}_1 \sum_{t=1}^{T} X_t$$
$$\hat{\beta}_0 = \frac{\sum_{t=1}^{T} Y_t}{T} - \hat{\beta}_1 \frac{\sum_{t=1}^{T} X_t}{T}$$
$$\Rightarrow \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

From (2)

$$\sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) X_t = 0$$

$$\sum_{t=1}^{T} Y_t X_t - \hat{\beta}_0 \sum_{t=1}^{T} X_t - \hat{\beta}_1 \sum_{t=1}^{T} X_t^2 = 0$$

$$\sum_{t=1}^{T} Y_t X_t = \hat{\beta}_0 \sum_{t=1}^{T} X_t + \hat{\beta}_1 \sum_{t=1}^{T} X_t^2$$

$$\sum_{t=1}^{T} Y_t X_t = \left[\bar{Y} - \hat{\beta}_1 \bar{X} \right] \underbrace{\sum_{t=1}^{T} X_t}_{T.\bar{X}} + \hat{\beta}_1 \sum_{t=1}^{T} X_t^2$$

$$\sum_{t=1}^{T} Y_t X_t = T.\bar{Y}.\bar{X} - \hat{\beta}_1 .T.\bar{X}.\bar{X} + \hat{\beta}_1 \sum_{t=1}^{T} X_t^2$$

$$\sum_{t=1}^{T} Y_t X_t - T.\bar{Y}.\bar{X} = \hat{\beta}_1 \sum_{t=1}^{T} X_t^2 - \hat{\beta}_1 .T.\bar{X}.\bar{X}$$

$$\sum_{t=1}^{T} Y_t X_t - T.\bar{Y}.\bar{X} = \hat{\beta}_1 \left[\sum_{t=1}^{T} X_t^2 - T.\bar{X}^2 \right]$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{t=1}^{T} Y_t X_t - T.\bar{Y}.\bar{X}}{\sum_{t=1}^{T} X_t^2 - T.\bar{X}^2}$$

Note that

$$\sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T} (X_t - \bar{X})^2 = \sum_{t=1}^{T} X_t^2 - 2\sum_{t=1}^{T} X_t \bar{X} + \sum_{t=1}^{T} \bar{X}^2$$

$$= \sum_{t=1}^{T} X_t^2 - 2\bar{X} \sum_{t=1}^{T} X_t + T.\bar{X}^2 = \sum_{t=1}^{T} X_t^2 - 2T\bar{X}^2 + T.\bar{X}^2$$

$$= \sum_{t=1}^{T} X_t^2 - T\bar{X}^2$$

and

$$\sum_{t=1}^{T} x_t y_t = \sum_{t=1}^{T} (X_t - \bar{X})(Y_t - \bar{Y}) = \sum_{t=1}^{T} (X_t Y_t - X_t \bar{Y} - \bar{X} Y_t + \bar{Y}\bar{X})$$
$$= \sum_{t=1}^{T} X_t Y_t - \bar{X} \sum_{t=1}^{T} Y_t - \bar{Y} \sum_{t=1}^{T} X_t + \sum_{t=1}^{T} \bar{Y}\bar{X}$$
$$= \sum_{t=1}^{T} X_t Y_t - \bar{X} T \bar{Y} - \bar{Y} T \bar{X} + T \bar{Y}\bar{X} = \sum_{t=1}^{T} X_t Y_t - T \bar{X}\bar{Y}$$

Hence the OLS estimator can also be written as in mean-deviation form as follows:

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2}$$

A. Mean of $\hat{\beta}_1$

We assume that we draw repeated samples of size *T* from the population of *Y* and *X*, and for each sample we estimate the parameters $\hat{\beta}_0$ and $\hat{\beta}_1$. This is known as hypothetical repeated sampling procedure. If all the possible samples are taken, then the mean value of $\hat{\beta}_1$ will be its expected value, *(mean* $\hat{\beta}_1$*)=E(*$\hat{\beta}_1$*)*. To find the value of the mean in terms of the observations of our sample of *Y* and *X* we work as follows.

We found that $\hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2}$.

Substituting $y_t = Y_t - \bar{Y}$ we obtain

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2} = \frac{\sum_{t=1}^{T} x_t (Y_t - \bar{Y})}{\sum_{t=1}^{T} x_t^2} = \frac{\sum_{t=1}^{T} x_t Y_t}{\sum_{t=1}^{T} x_t^2} - \frac{\bar{Y}\sum_{t=1}^{T} x_t}{\sum_{t=1}^{T} x_t^2}$$

But by definition, the sum of the deviations of a variable from its mean is identically equal to zero, $\sum_{t=1}^{T} x_t = 0$. Therefore

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_t Y_t}{\sum_{t=1}^{T} x_t^2} = \sum_{t=1}^{T} \left(\frac{x_t}{\sum_{t=1}^{T} x_t^2} . Y_t \right)$$

By **assumption 2** of the method of least squares, the values of *X* are a set of fixed values, which do not change from sample to sample. Consequently the ratio $\frac{x_t}{\sum_{t=1}^{T} x_t^2}$ will be constant from sample to sample, and if we denote the ratio by a_t we may write the estimator $\hat{\beta}_1$ in the form $\hat{\beta}_1 = \sum_{t=1}^{T} a_t Y_t$.

By substituting the value of $Y_t = \beta_0 + \beta_1 X_t + u_t$ and rearranging the factors we find

$$\hat{\beta}_1 = \sum_{t=1}^{T} a_t \left(\beta_0 + \beta_1 X_t + u_t \right)$$

$$\hat{\beta}_1 = \beta_0 \sum_{t=1}^{T} a_t + \beta_1 \sum_{t=1}^{T} a_t X_t + \sum_{t=1}^{T} a_t u_t$$

Note (and show) that $\sum_{t=1}^{T} a_t = 0$ and $\sum_{t=1}^{T} a_t X_t = 1$.

Therefore, the equation above reduces to

$$\hat{\beta}_1 = \beta_1 + \sum_{t=1}^{T} a_t u_t$$

which implies that $\hat{\beta}_1$ is a linear estimator because it is a linear function of *Y*; actually it is a weighted average of Y_t with a_t serving as the weights.

Taking expected values yields

$$E(\hat{\beta}_1) = E(\beta_1) + \sum_{t=1}^{T} a_t E\left[u_t\right]$$

The significance of the assumption of *constant X values* is seen in the above manipulations, in that the operation of taking expected values is applied to *u* and *Y* values but not to *X*.

Since β_1 (the true population parameter) is constant, we can write $E(\beta_1) = \beta_1$. Finally using **assumption 3**, we have $E\left[u_t\right] = 0$.

Hence, the equation reduces to

$$\underbrace{E(\hat{\beta}_1)}_{mean\, of\ \hat{\beta}_1} = \beta_1$$

which implies that the mean of OLS estimate $\hat{\beta}_1$ is equal to the true value of the population parameter β_1**.**

This implies that the $\hat{\beta}_1$ is an *unbiased* estimator.

B. Variance of $\hat{\beta}_1$

It can be proved that $Var(\hat{\beta}_1) = E\left[\hat{\beta}_1 - E(\hat{\beta}_1)\right]^2 = E\left[\hat{\beta}_1 - \beta_1\right]^2 = \dfrac{\sigma^2}{\sum_{t=1}^{T} x_t^2}$.

To show this, recall that we established $\hat{\beta}_1 = \sum_{t=1}^{T} a_t Y_t$ where $a_t = \dfrac{x_t}{\sum_{t=1}^{T} x_t^2}$ = constant weights in the process of hypothetical repeated sampling.

Therefore,

$$Var(\hat{\beta}_1) = Var(\sum_{t=1}^{T} a_t Y_t) = \sum_{t=1}^{T} a_t^2\, Var(Y_t)$$

given that $a_t = \dfrac{x_t}{\sum_{t=1}^{T} x_t^2}$ are constant weights, independent of the values of Y_t by Assumption 2.

However, recall that $Var(Y_t) = \sigma^2$.

Therefore,

$$Var(\hat{\beta}_1) = \sigma^2 \sum_{t=1}^{T} a_t^2 = \sigma^2 \left[\sum_{t=1}^{T} \frac{x_t^2}{\left(\sum_{t=1}^{T} x_t^2 \right)^2} \right] = \sigma^2 \left[\frac{\sum_{t=1}^{T} x_t^2}{\left(\sum_{t=1}^{T} x_t^2 \right)^2} \right] = \frac{\sigma^2}{\sum_{t=1}^{T} x_t^2}$$

C. Mean of $\hat{\beta}_0$

In the last lecture, we have established that

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Substituting $\hat{\beta}_1 = \sum_{t=1}^{T} a_t Y_t$ we obtain:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$
$$\hat{\beta}_0 = \bar{Y} - \bar{X} \sum_{t=1}^{T} a_t Y_t$$
$$\hat{\beta}_0 = \frac{\sum_{t=1}^{T} Y_t}{T} - \bar{X} \sum_{t=1}^{T} a_t Y_t$$

Taking Y_t as the common factor, we may write:

$$\hat{\beta}_0 = \sum_{t=1}^{T} \left[\frac{1}{T} - \bar{X} a_t \right] Y_t$$

Here, denoting $\frac{1}{T} - \bar{X} a_t = b_t$, we can write the equation as

$$\hat{\beta}_0 = \sum_{t=1}^{T} b_t Y_t$$

which implies that $\hat{\beta}_0$ is a linear estimator.

Taking expected values

$$E(\hat{\beta}_0) = \sum_{t=1}^{T} \left[\frac{1}{T} - \bar{X} a_t \right] E(Y_t)$$

given that T, $\bar{X}$ and a_t are constant from sample to sample.

In the last lecture, we have derived that $E(Y_t) = \beta_0 + \beta_1 X_t$.

Therefore

$$E(\hat{\beta}_0) = \sum_{t=1}^{T} \left[\frac{1}{T} - \bar{X} a_t \right] (\beta_0 + \beta_1 X_t)$$

$$E(\hat{\beta}_0) = \sum_{t=1}^{T} \left[\frac{\beta_0}{T} - \bar{X} a_t \beta_0 + \frac{\beta_1 X_t}{T} - \bar{X} a_t \beta_1 X_t \right]$$

$$E(\hat{\beta}_0) = \sum_{t=1}^{T} \frac{\beta_0}{T} - \bar{X} \beta_0 \underbrace{\sum_{t=1}^{T} a_t}_{0} + \beta_1 \sum_{t=1}^{T} \frac{X_t}{T} - \bar{X} \beta_1 \underbrace{\sum_{t=1}^{T} a_t X_t}_{1}$$

$$E(\hat{\beta}_0) = \beta_0 + \beta_1 \bar{X} - \beta_1 \bar{X}$$

$$E(\hat{\beta}_0) = \beta_0$$

Hence $\hat{\beta}_0$ *is an <u>unbiased estimator</u> of* β_0.

A similar proof can be done using the notation $\hat{\beta}_0 = \sum_{t=1}^{T} b_t Y_t$, with the help of $\sum_{t=1}^{T} b_t = 1$ and $\sum_{t=1}^{T} b_t X_t = 0$ (Show that $\sum_{t=1}^{T} b_t = 1$ and $\sum_{t=1}^{T} b_t X_t = 0$).

D. Variance of $\hat{\beta}_0$

We established that

$$\hat{\beta}_0 = \sum_{t=1}^{T} b_t Y_t$$

$$\hat{\beta}_0 = \sum_{t=1}^{T} \left[\frac{1}{T} - \bar{X} a_t \right] Y_t$$

Therefore

$$Var(\hat{\beta}_0) = Var\left[\sum_{t=1}^{T} b_t Y_t\right]$$

$$Var(\hat{\beta}_0) = Var\left[Y_t\right]\sum_{t=1}^{T} b_t^2$$

$$Var(\hat{\beta}_0) = \sigma^2 \sum_{t=1}^{T} b_t^2$$

$$Var(\hat{\beta}_0) = \sigma^2 \sum_{t=1}^{T}\left(\frac{1}{T} - \bar{X}a_t\right)^2$$

$$Var(\hat{\beta}_0) = \sigma^2 \sum_{t=1}^{T}\left(\frac{1}{T^2} - \frac{2\bar{X}a_t}{T} + \bar{X}^2 a_t^2\right)$$

Thus,

$$Var(\hat{\beta}_0) = \sigma^2\left[\frac{1}{T} - \frac{2\bar{X}}{T}\sum_{t=1}^{T} a_t + \bar{X}^2 \sum_{t=1}^{T} a_t^2\right]$$

$$Var(\hat{\beta}_0) = \sigma^2\left[\frac{1}{T} - \frac{2\bar{X}}{T}\sum_{t=1}^{T} a_t + \bar{X}^2 \sum_{t=1}^{T} a_t^2\right]$$

Since $\sum_{t=1}^{T} a_t = 0$ and $\sum_{t=1}^{T} a_t^2 = \frac{1}{\sum_{t=1}^{T} x_t^2}$, we obtain

$$Var(\hat{\beta}_0) = \sigma^2\left[\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^{T} x_t^2}\right] \Rightarrow Var(\hat{\beta}_0) = \sigma^2\left[\frac{\sum_{t=1}^{T} x_t^2 + T\bar{X}^2}{T\sum_{t=1}^{T} x_t^2}\right]$$

Recall that $\sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T} X_t^2 - T\bar{X}^2$. This implies that $\sum_{t=1}^{T} x_t^2 + T\bar{X}^2 = \sum_{t=1}^{T} X_t^2$. Hence, the variance of $\hat{\beta}_0$ is obtained as follows:

$$Var(\hat{\beta}_0) = \sigma^2 \frac{\sum_{t=1}^{T} X_t^2}{T\sum_{t=1}^{T} x_t^2}$$

E. Covariance of $\hat{\beta}_0$ and $\hat{\beta}_1$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = E\left[\hat{\beta}_0 - E(\hat{\beta}_0)\right]\left[\hat{\beta}_1 - E(\hat{\beta}_1)\right]$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = E\left[\underbrace{\hat{\beta}_0 - E(\hat{\beta}_0)}_{\substack{-\bar{X}(\hat{\beta}_1 - \beta_1) \\ Why?\,See\,below!}}\right]\left[\hat{\beta}_1 - \underbrace{E(\hat{\beta}_1)}_{\beta_1}\right]$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = E\left[-\bar{X}(\hat{\beta}_1 - \beta_1)\right]\left[\hat{\beta}_1 - \beta_1\right]$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\bar{X}.\underbrace{E\left[\hat{\beta}_1 - \beta_1\right]^2}_{Var(\hat{\beta}_1)}$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\bar{X}.\underbrace{Var(\hat{\beta}_1)}_{\sigma^2 / \sum_{t=1}^{T} x_t^2}$$

$$\Rightarrow Cov(\hat{\beta}_0, \hat{\beta}_1) = -\bar{X}\frac{\sigma^2}{\sum_{t=1}^{T} x_t^2}$$

Obtaining the covariance expression, we have used the following equality: $\hat{\beta}_0 - E(\hat{\beta}_0) = -\bar{X}(\hat{\beta}_1 - \beta_1)$. Let us show how we can obtain this relationship, below.

Recall that:

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$, which yields $E(\hat{\beta}_0) = \bar{Y} - E(\hat{\beta}_1)\bar{X}$.Hence, we get:

$E(\hat{\beta}_0) = \bar{Y} - \beta_1\bar{X}$

Therefore,

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

$$E(\hat{\beta}_0) = \overline{Y} - \beta_1 \overline{X}$$

$$\hat{\beta}_0 - E(\hat{\beta}_0) = -\hat{\beta}_1 \overline{X} + \beta_1 \overline{X}$$

$$\Rightarrow \hat{\beta}_0 - E(\hat{\beta}_0) = -\overline{X}(\hat{\beta}_1 - \beta_1)$$