DATA 5600 4/11/2022

Wooldridge chp. 2 Continued

Recall that we defined the OLS (ordinary least squares) estimators for β_0 and β_1 as

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

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

for any $\hat{\beta}_0$ and $\hat{\beta}_1$ define a fitted value for y

when $x = x_i$ as

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

This is the value we predict for y when $x = x_i$ for a given intercept and slope.

There is a fitted value for every observation in the sample.

The residual for observation i is the difference between the actual y_i and its fitted value:

$$\hat{u}_i = y_i - \hat{y} = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

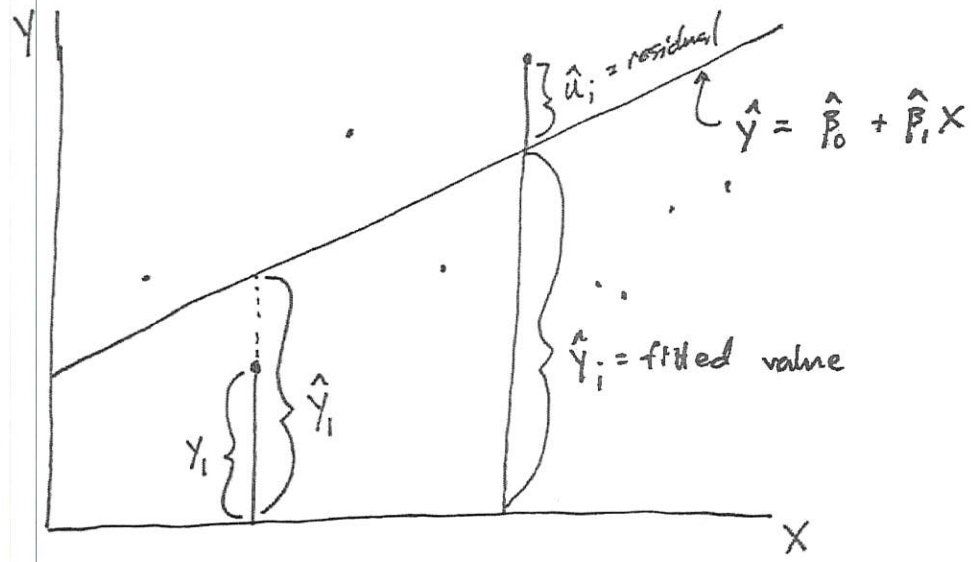
Now suppose we choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to make the sum of squared residuals

$$\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

as small as possible

Fitted Values & Residuals

graph



Once we have obtained $\hat{\beta}_0$ and $\hat{\beta}_1$ by OLS we form the <u>OLS regression line</u>:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

In most cases the slope estimate, which we can write as

$$\hat{\beta}_1 = \Delta \hat{Y} / \Delta X$$

is of primary interest. It tells us the amount by which $\hat{Y}$ changes when X increases by one unit. Equivalently

$$\Delta \hat{Y} = \hat{\beta}_1 \Delta X$$

Ex: 2.3 p. 33

For the population of CEO's, let y be annual salary in thousands of dollars.

So, for example $y = 856.3$ indicates a salary of $856,300 and $y = 1,452.6$ indicates a salary of $1,452,600.

Let x be the average return on equity (roe) for the CEO's firm for the previous three years. ROE is defined in terms of net income as a percentage of common equity. For example, if $x = 10$ then average return on equity is 10%.

To study the relationship between roe and ceo compensation,

we propose the simple model

$$salary = \beta_0 + \beta_1 roe + u$$

The slope parameter β_1 measures the change in annual salary, in thousands of dollars, when return on equity increases by one percentage point.

Because a high roe is good ~~for~~ a firm we expect $\beta_1 > 0$

Using data on CEO salaries for a sample of $n = 209$ for 1990 collected from Business Week we obtain

⑦

$$\widehat{salary} = 963.191 + 18.501\, roe$$

Q: How do we interpret this?

If roe = 0 then predicted salary is the intercept 963.191, which equals $963,191 since salary is measured in thousands.

Next we can write the predicted change in salary as a function of the change in roe:

$$\Delta \widehat{salary} = 18.501\, \Delta roe$$

If roe changes by one percentage point $\Delta roe = 1$ then salary is predicted to change by about 18.5 or $18,500

We can use this to predict salaries at different values of roe. Suppose, for example, that roe = 30 then

$$\widehat{salary} = 963.191 + 18.501(30)$$

$$= 1,518,221$$

which is just over \$1.5 million

Writing y_i as

$$y_i = \underset{\substack{\uparrow \\ \text{fitted} \\ \text{value}}}{\hat{y}_i} + \underset{\substack{\uparrow \\ \text{residual}}}{\hat{u}_i}$$

$$\left\{ \begin{array}{l} \text{NB: can write} \\ \hat{u}_i = y_i - \hat{y}_i \end{array} \right.$$

We can define total sum of squares (SST)

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

the explained sum of squares (SSE)

$$SSE = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

and the residual sum of squares (SSR)

$$SSR = \sum_{i=1}^{n} \hat{u}_i^2$$

SST measures the total sample variation in y_i - that is how spread out the y_i are in the sample.

NB: dividing SST by (n-1) gives the sample variance of y

SSE measures the sample variation in $\hat{Y}_i$

SSR measures the sample variation in $\hat{u}_i$

We can write

$$SST = SSE + SSR$$

{ NB: see p. 39 for proof

Goodness of Fit

We define R-squared of the regression as

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$$

R^2 is the ratio of the explained variation compared to total variation, it is interpreted as the fraction of the sample variation in y that is explained by X

$0 < R^2 < 1$

We usually multiply by $100 R^2$ to get the percentage of the sample variation in y that is explained by X

Statistical Properties of the OLS Estimators

Now that we have mathementical rules (estimators) for $\hat{\beta}_0$ and $\hat{\beta}_1$ defined we can assess their statistical properties.

This will be a discussion of the sampling distributions (PDF) of $\hat{\beta}_0$ and $\hat{\beta}_1$ over different random samples from the population.

We will start by establishing that OLS gives unbiased estimates. But we will need to state a few assumptions that we will rely on.

Assumption SLR.1 - Linear in Parameters

In the population model, the explained variable y, is related to the explanatory variable, x, and the error, u as

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

where β_0 and β_1 are the population intecept and slope, respectively.

NB: y is related to x linearly in β's (this is really as stance regarding the "true" population DGP)

Assumption SLR.2 – Random Sampling

We have a random sample of size n, $\{(x_i, y_i) : i = 1, 2, \ldots, n\}$, following the population model

$$Y = \beta_0 + \beta_1 X + u$$

Assumption SLR.3 – Sample Variation in the Explanator Variable

The sample outcomes on X, namely $\{X_i, i = 1, \ldots, n\}$ are not all the same value.

Assumption SLR.4 - Zero Conditional Mean

The error u has an expected value of zero given any value of the explanatory variable.

In other words,

$$E(u/x) = 0$$

NB: This is a statement about the joint distribution of x and u. Essentially, it amounts to an assumption that x and u are independent, and that $E(u) = 0$

Unbiasedness of OLS

Using SLR.1 through SLR.4

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

and

$$E(\hat{\beta}_1) = \beta_1$$

in other words, $\hat{\beta}_0$ is unbiased for β_0 and, $\hat{\beta}_1$ is unbiased for $\hat{\beta}_1$

To start with we can write

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

slope estimator as

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

Which we can write as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{SST_x} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i)}{SST_x}$$

in which $SST_x = \sum_{i=1}^{n}[x_i - \bar{x}]^2$ (as we have defined it)

Using the properties of the summation operator $\sum$, we can write the numerator as

$$\sum_{i=1}^{n}(x_i - \bar{x})\beta_0 + \sum_{i=1}^{n}(x_i - \bar{x})\beta_1 x_i + \sum_{i=1}^{n}(x_i - \bar{x})u_i$$

$$= \beta_0 \sum(x_i - \bar{x}) + \beta_1 \sum(x_i - \bar{x})x_i + \sum(x_i - \bar{x})u_i$$

Recall that $\sum(x_i - \bar{x}) = 0$ and that $\sum(x_i - \bar{x})x_i = \sum(x_i - \bar{x})^2 = SST_x$

Now we can write the numerator as $\beta_1 SST_x + \sum_{i=1}^{n}(x_i - \bar{x})u_i$

Putting this over the denominator gives

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_1^1 (x_i - \hat{x}) u_i}{SST_x}$$

$$= \beta_1 + \frac{1}{SST_x} \sum_1^1 d_i u$$

Where $d_i = x_i - \bar{x}$

And now we can look at $E(\hat{\beta}_1)$

$$E(\hat{\beta}_1) = \beta_1 + E\left[\frac{1}{SST_x} \sum d_i u_i \right]$$

$$= \beta_1 + \frac{1}{SST_x} \sum_{i=1}^{n} E(d_i u_i)$$

$$= \beta_1 + \frac{1}{SST_x} \sum d_i E(u_i)$$

$$= \beta_1 + \frac{1}{SST_x} \sum d_i (0)$$

$$= \beta_1$$

Thus, $\hat{\beta}_1$ is unbiased for β_1

~~[illegible] proof for [illegible]~~

~~[illegible]~~

decide to show it!

The proof for $\hat{\beta}_0$ is now straight forward. We can average

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad , \quad i=1,\dots,n$$

across i to ge

$$\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{u}$$

and plug this into $\hat{\beta}_0$ to get

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \beta_0 + \beta_1 \bar{X} + \bar{u} - \hat{\beta}_1 \bar{X} = \beta_0 + (\beta_1 - \hat{\beta}_1)\bar{X} + \bar{u}$$

Then conditional on the values of the x_i

$$E(\hat{\beta}_0) = \beta_0 + E[(\beta_1 - \hat{\beta}_1)\bar{x}] + E(\bar{u})$$

$$= \beta_0 + E[(\beta_1 - \hat{\beta}_1)\bar{x}]$$

Since $E(\bar{u}) = 0$. Now since we know that $E(\hat{\beta}_1) = \beta_1$ we can say that $E(\beta_1 - \hat{\beta}_1) = \beta_1 - \beta_1 = 0$ so that

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

Thus $\hat{\beta}_0$ is unbiased for β_0

Variances of the OLS Estimators

We will state an additional assumption

Assumption SLR. 5 - Homoskedasticity

The error u has the same variance given any value of the explanatory variable. In other words,

$$Var(u|x) = \sigma^2$$

Because $Var(u|x) = E(u^2|x) - [E(u|x)]^2$ and $E(u|x) = 0$

we can say that $E(u^2|x) = Var(u) = \sigma^2$. I.E. b/c

$E(u|x) = 0$ σ^2 is also the unconditional variance of u.

It is often called the error variance or disturbance variance.

It is useful to write SLR.4 and SLR.5 in terms of

the conditional mean and conditional variance of y!

$$E(y|x) = \beta_0 + \beta_1 x$$

$$Var(y|x) = \sigma^2$$

Sampling Variances of the OLS Estimators

NB: it is sampling not sample variance that we are talking about here!

Under assumptions SLR.1 through SLR.2

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{1}{SST_x}\sigma^2$$

and

$$\mathrm{Var}(\hat{\beta}_0) = \frac{\sigma^2 \frac{1}{n} \sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

NB: We will not show the proof. The mathematically minded may see pp. 54- 55 for the proof.

An unbiased Estimator of σ^2

Under assumptions SLR.1 through SLR.5

$$E(\hat{\sigma}^2) = \sigma^2 \qquad \text{(i.e. an unbiased estimator)}$$

in which

$$\sigma^2 = \frac{1}{(n-2)} \sum_{i=1}^{n} \hat{u}_i^2 = \frac{SSR}{n-2}$$

and

$$se(\hat{\beta}_1) = \hat{\sigma} \Big/ \sqrt{SST_X} = \hat{\sigma} \Big/ \left(\sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{1/2}$$

↳ standard error of $\hat{\beta}_1$

NB: Recall that

$\hat{u}_i = Y_i - \hat{Y}_i$ and

$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$