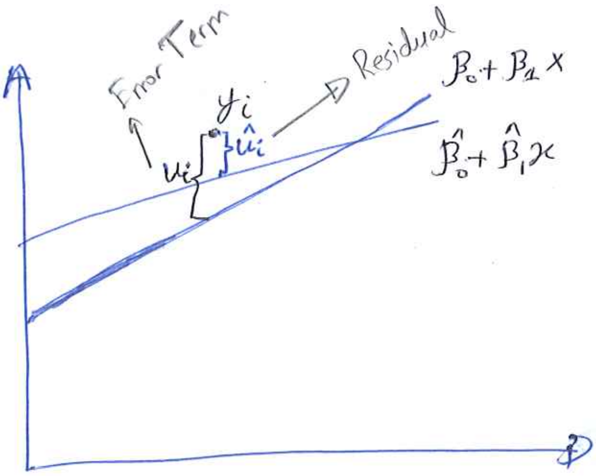
# Econometrics (1/March/2016)

$$y = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 x}_{\hat{y}} + \hat{u}$$




$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i)}{\sum (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 x$$

A1. Linearity in Parameter is important.

For example, $y = x_1^{\beta_1} x_2^{\beta_2}$

$$\log y = \cancel{\beta_1 x_1 + \beta_2}\ \beta_1(\log x_1) + \beta_2(\log x_2)$$

$$W = \beta_1 Z_1 + \beta_2 Z_2$$

A2. Randomness $\Rightarrow$ (i.i.d observations)

A3. Conditional Independence $\rightarrow$ In case of violation we take an instrumental ~~and~~ variable or Difference in Difference. (or Twin Studies)

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

Properties of the OLS:

$$\sum \hat{u}_i = 0 \;/\; \sum \hat{u}_i x_i = 0$$

$$\sum \hat{u}_i \hat{y}_i = 0 \Rightarrow \bar{\hat{y}} = \hat{y}$$

$$SST = SSR + SSE$$

$$y = \frac{\beta_1}{\beta_0} x$$

$$\ln(y) = \ln(\beta_1) + \ln(x) - \ln(\beta_0$$

A4. $SST_x > 0$

A5. Homoskedacity (Opposite: Heteroskedacity)

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

Constant Variance

A6. Absence of large outliers

## Expected Value of the OLS

$$y = \beta_0 + \hat{\beta}_1 x + u$$

- Here $x$ is not stochastic because it is exogenous and therefore, given.

$$\hat{\beta}_1 = \frac{\Sigma (x_i - \bar{x})}{SST_x} y_i$$

$$= \frac{\Sigma (x_i - \bar{x})}{SST_x} [\beta_0 + \beta_1 x + u_x]$$

Hence

$$\hat{\beta}_1 = \underbrace{\beta_0 \frac{\Sigma (x_i - \bar{x})}{SST_x}}_{=0} + \underbrace{\beta_1 \frac{\Sigma (x_i - \bar{x}) x_i}{SST_x}}_{= \beta_1 \frac{SST_x}{SST_x} = \beta_1} + \underbrace{\frac{\Sigma (x_i - \bar{x}) u_i}{SST_x}}_{= \frac{E[(\Sigma (x_i - \bar{x})) u_i]}{SST_x} = 0}$$

$$k_i = \frac{x_i - \bar{x}}{SST_x} \implies \Sigma k_i = 0 \; / \; \Sigma k_i x_i = 1 \; / \; \Sigma k_i^2 = \frac{1}{SST_x}$$

$$\hat{\beta}_1 = \sum \frac{(x_i - \bar{x})}{SST_X} y_i \Longrightarrow \hat{\beta}_1 = \sum k_i y_i$$

$$E(\hat{\beta}_1) = E\left(\sum k_i y_i\right) = E\left(\sum k_i [\beta_0 + \beta_1 x + u_i]\right)$$

$$= \underbrace{\beta_0 \sum k_i}_{0} + \underbrace{\beta_1 \sum k_i x_i}_{\beta_1} + \underbrace{E\left(\sum k_i u_i\right)}_{0}$$

$$= \beta_1$$

$$y_i = \beta_0 + \beta_1 x + u_i$$

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

$$Var(y_i) = Var(\beta_0 + \beta_1 x + u_i)$$

$$= 0 + 0 + Var(u_i | x)$$

$$= \sigma^2$$

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})}{SST_x} y_i = \sum k_i y_i$$

$$Var(\hat{\beta}_1) = Var\left(\sum k_i y_i\right) = \sum k_i^2 Var(y_i)$$

$$= \sum k_i^2 \sigma^2 = \frac{\sigma^2}{SST_x}$$

$$Var(\hat{\beta}_0) = Var(\bar{y} - \hat{\beta}_1 \bar{x})$$

$$= Var(\bar{y}) + \frac{\bar{x}^2 \sigma^2}{SST_X}$$ ?

$$= \frac{\sigma^2}{n} + \frac{x^2 \sigma^2}{SST_X}$$

$$= \sigma^2 \left[ \frac{1}{n} + \frac{x^2}{SST_X} \right]$$

Gauss-Markov Theorem

Under A1-A6

OLS estimators $(\hat{\beta}_0, \hat{\beta}_1)$ are Best Linear Unbiased Estimators

Proof:

$$\tilde{\beta}_1 = \sum g_i y_i$$

Since linear $\sum g_i = 0$ and $\sum g_i x_i = 1$

$$Var(\tilde{\beta}_1) = Var(\sum g_i y_i) = \sum g_i^2 \sigma^2$$

$$= \sum (g_i + k_i - k_i)^2 \sigma^2$$

$$= \sum (g_i - k_i)^2 + \sum k_i \sigma^2 \geq \sum k_i \sigma^2$$

$= OLS$