Classical Normal Linear Regression Model (CNLRM)

Classical theory of statistical inference

- Estimation
- Hypothesis testing

The Probability Distribution of Disturbances

$$\hat{\beta}_2 = \sum k_i Y_i$$

where $$k_i = \frac{x_i}{\sum x_i^2} = \frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$$

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\hat{\beta}_2 = \sum k_i (\beta_1 + \beta_2 X_i + u_i)$$

Linearity and Unbiasedness Properties of Least-Squares Estimators

$$\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$= \frac{\sum x_i y_i}{\sum x_i^2}$$

$x_i = X_i - \bar{X}$

$y_i = Y_i - \bar{Y}$

$$\sum x_i y_i = \sum x_i (Y_i - \bar{Y})$$

$$= \sum x_i Y_i - \bar{Y} \sum x_i$$

$$= \sum x_i Y_i - \bar{Y} \sum (X_i - \bar{X}) \quad \rightarrow 0$$

$$= \sum x_i Y_i$$

$$\hat{\beta}_2 = \frac{\sum x_i Y_i}{\sum x_i^2}$$

$$\hat{\beta}_2 = \frac{\sum x_i Y_i}{\sum x_i^2} = \sum k_i Y_i$$

where $$k_i = \frac{x_i}{\sum x_i^2}$$

$\hat{\beta}_2$ is a linear estimator because it is a linear function of Y; actually it is a weighted average of Y_i with k_i serving as the weights.

Properties of the weights k_i:

① Since the X_i are assumed to be nonstochastic, the k_i are nonstochastic too.

② $\sum k_i = 0$

E.g. $\sum k_i = \sum \left(\frac{x_i}{\sum x_i^2} \right) = \frac{1}{\sum x_i^2} \sum x_i$

$= 0$

3. $\sum k_i^2 = \dfrac{1}{\sum x_i^2}$

$$k_i = \frac{x_i}{\sum x_i^2}$$

$$\sum \left(\frac{x_i}{\sum x_i^2} \right)^2$$

$$= \frac{x_1^2}{(\sum x_1^2)^2} + \frac{x_2^2}{(\sum x_2^2)^2} + \dots \frac{x_n^2}{(\sum x_n^2)^2}$$

$$= \frac{\sum x_i^2}{(\sum x_i^2)^2}$$

$$= \frac{\cancel{\sum x_i^2}}{\cancel{(\sum x_i^2)}(\sum x_i^2)} = \frac{1}{\sum x_i^2}$$

4. $\sum k_i x_i = \sum k_i X_i = 1$.

These properties can be directly verified from the definition of k_i.

$$k_i = \frac{x_i}{\sum x_i^2}$$

$$\sum k_i x_i = \frac{\sum x_i^2}{\sum x_i^2} = 1$$

Now substitute the PRF $Y_i = \beta_1 + \beta_2 X_i + u_i$

$$\hat{\beta}_2 = \sum k_i Y_i$$

$$\hat{\beta}_2 = \sum k_i(\beta_1 + \beta_2 X_i + u_i)$$

$$= \beta_1 \sum k_i^{\nearrow 0} + \beta_2 \sum k_i X_i^{\nearrow 1} + \sum k_i u_i$$

$$= \beta_2 + \sum k_i u_i$$

where use is made of the properties of k_i

Now taking expectation of equation (4) on both sides and noting that k_i, being nonstochastic, can be treated as constants, we obtain

$$E(\hat{\beta}_2) = \beta_2 + \sum k_i \underbrace{E(u_i)}_{=0 \text{ by assumption}}$$

$$= \beta_2$$

The Normality Assumption for

The classical normal linear regression model assumes that each u_i is distributed normally with

$$E(u_i) = 0$$

$$E[u_i - E(u_i)]^2 = E(u_i^2) = \sigma^2$$

$$E\{[(u_i - E(u_i)][(u_j - E(u_j)]\} = E(u_i u_j) = 0 \quad i \neq j$$

$$u_i \sim N(0, \sigma^2)$$

Where **N** stands for normal distribution

$$u_i \sim NID(0, \sigma^2)$$

Where **NID** stands for normally and independent distributed

Properties of OLS Estimators under the Normality Assumption

- Unbiased
- Minimum variance unbiased or efficient estimators
- Consistency

  Sample size increases $\longrightarrow$ the estimators converge to their true population values

Properties of OLS Estimators under the Normality Assumption

- is normally distributed with $\hat{\beta}_1$

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

$$\text{var}(\hat{\beta}_1): \quad \sigma^2_{\hat{\beta}_1} = \frac{\sum X_i^2}{n \sum x_i^2} \sigma^2$$

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

Properties of OLS Estimators under the Normality Assumption

By the properties of the normal distribution, the variable Z,

$$Z = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}}$$

follows the standard normal distribution, that is, a normal distribution with zero mean and unit variance

$$Z \sim N(0,1)$$

Properties of OLS Estimators under the Normality Assumption

- is normally distributed with $\hat{\beta}_2$

$$Mean: \; E(\hat{\beta}_2) = \beta_2$$

$$\text{var}(\hat{\beta}_2): \; \sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\sum x_i^2}$$

$$\hat{\beta}_2 \sim N(\beta_2, \sigma^2_{\hat{\beta}_2})$$

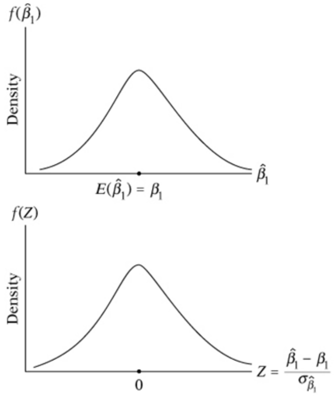
Properties of OLS Estimators under the Normality Assumption

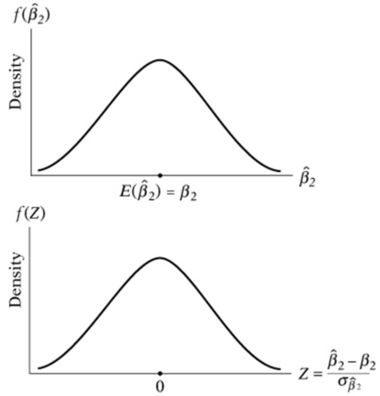
By the properties of the normal distribution, the variable Z,

$$Z = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}}$$

follows the standard normal distribution

The probability distributions of $\hat{\beta}_1$ and $\hat{\beta}_2$





Properties of OLS Estimators under the Normality Assumption

- $(n-2)(\hat{\sigma}^2/\sigma^2)$ is distributed as the χ^2 (chi square) distribution with (n-2) degree of freedom

- $(\hat{\beta}_1, \hat{\beta}_2)$ are distributed independently of $\hat{\sigma}^2$

- $\hat{\beta}_1$ *and* $\hat{\beta}_2$ have minimum variance in the entire class of unbiased estimators, whether linear or not

Source

Gujarati, D.N. (2009) Basic Econometrics. 5th ed. Singapore, McGraw-Hill.