Classical Normal Linear Regression Model (CNLRM)

Classical theory of statistical inference

- Estimation
  - Point e.g. $\hat{\beta}_1$, $\hat{\beta}_2$, $\bar{X}$, $\hat{\sigma}^2$
  - Interval
- Hypothesis testing
  - t-test
  - F-test

Recall in regression analysis our objective is not only to estimate the sample regression function (SRF)

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

The Probability Distribution of Disturbances

$$\hat{\beta}_2 = \sum k_i Y_i \quad (*)$$

where

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

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

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

$\hat{\beta}_2$ is a linear estimator because it is linear function of Y

Linearity and Unbiasedness Properties of Least-Squares Estimators

$$\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \quad (1)$$

$x_i = X_i - \bar{X}$

$y_i = Y_i - \bar{Y}$

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

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

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

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

$$= \sum x_i Y_i$$

$$\boxed{\hat{\beta}_2 = \frac{\sum x_i Y_i}{\sum x_i^2}} \quad (3)$$

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

$x_i = X_i - \bar{X}$

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

where $k_i = \dfrac{x_i}{\sum x_i^2}$ $\circledast$

$\hat{\beta}_2$ 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: $\dfrac{x_i}{\sum x_i^2}$

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

② $\sum k_i = 0$

E.g. $\sum k_i = \sum \left(\dfrac{x_i}{\sum x_i^2} \right) = \dfrac{1}{\sum x_i^2} \sum x_i = 0$

$x_i = X_i - \bar{X}$

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

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

$\sum \left(\frac{x_i}{\sum x_i^2} \right)^2$ 1

1st 2nd . . . . n^{th}

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

$= \frac{\sum x_i^2}{(\sum x_i^2)^2}$ 3

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

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

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

$x_i = X_i - \bar{X}$

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

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

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

$k_i = ?$

$$\hat{\beta}_2 = \sum k_i Y_i \quad (1)$$

$$\hat{\beta}_2 = \sum k_i (\beta_1 + \beta_2 X_i + u_i) \quad (2)$$

$$= \beta_1 \cancel{\sum k_i}^{\,0} + \beta_2 \cancel{\sum k_i X_i}^{\,1} + \sum k_i u_i \quad (3)$$

$$= \beta_2 + \sum k_i u_i \quad (4)$$

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

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

$$= \beta_2$$

Minimum variance property of least squares estimators

(See appendix 3A.6)

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

P.109 Why the normality assumption?

1. u_i represent the combined influence (on the dependent variable) of a large number of independent variables that are not explicitly introduced in the regression model. As noted, we hope that the influence of these omitted variables is small and at best random.

Central limit theorem (CLT) → if there are a large number of independent and identically distributed random variables, then, with a few exceptions, the distribution of their sum tends to a normal distribution as the number of such variables increases indefinitely. It is the CLT that provides a theoretical justification for the assumption of normality of u

$wage_i = \beta_1 + \beta_2 edu + u_i$

ability

2. A variant of the CLT states that even if the number of variables is not very large or if these variables are not strictly independent, their sum may still be normally distributed.

3. With normality assumption, the probability distributions of OLS estimators can be easily derived. One property of the normal distribution is that any linear function of normally distributed variables is itself normally distributed.

OLS estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ are linear functions of u_i

Therefore, if u_i are normally distributed, so are $\hat{\beta}_1$ and $\hat{\beta}_2$, which make our task of hypothesis testing very straightforward

4. The normal distribution is a comparatively simple distribution involving only two parameters (mean and variance).

5. If we are dealing with a small (or finite, sample size, say data less than 100 observations, the normality assumption assumes a critical role.

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

$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$

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

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

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

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

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 $\hat{\beta}_2$

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

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

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

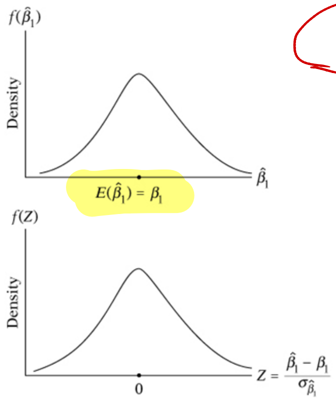
Properties of OLS Estimators under the Normality Assumption

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

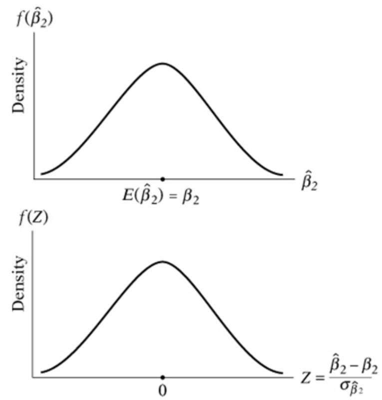
$$Z = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}}$$

follows the standard normal distribution

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



$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$



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.