Econ 361: Advanced Econometrics

Properties of Estimators

Bivariate Linear Regression Model

Size N (Bivariate) Sample: $\{ \ (Y_i = y_i, X_i = x_i) \ \}_{i=1}^{N}$

OLS Predictor of Y given X: $\hat{Y}_{ols} = a^{ols} + b^{ols} X$

Sometimes also expressed as $\hat{Y}_{ols} = \underbrace{b_0^{ols}}_{\text{“intercept”}} + \underbrace{b_1^{ols}}_{\text{“slope”}} X$

OLS Estimat**ors**

$$b_0^{ols} = \underbrace{\underbrace{\frac{1}{N}\sum_{i=1}^{N} Y_i}_{\bar{Y}_N} - b_1^{ols} \underbrace{\frac{1}{N}\sum_{i=1}^{N} X_i}_{\bar{X}_N}}_{\text{analogous to } E[Y] - \beta^* E[X]}$$

$$b_1^{ols} = \underbrace{\frac{\frac{1}{N}\sum_{i=1}^{N} X_i Y_i - \bar{X}_N \bar{Y}_N}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2}}_{\text{analogous to } \beta^* = \mathrm{Cov}(X,Y)/\mathrm{Var}(X)}$$

Estimators, as functions of random variables, are themselves random variables.

OLS Estimat**es**

$$\text{“}\, b_0^{ols} \,\text{”} \quad = \quad \underbrace{\frac{1}{N}\sum_{i=1}^{N} y_i}_{\bar{y}_N} \; - \; b_1^{ols} \underbrace{\frac{1}{N}\sum_{i=1}^{N} x_i}_{\bar{x}_N}$$

$$\text{“}\, b_1^{ols} \,\text{”} \quad = \quad \frac{\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \bar{x}_N \bar{y}_N}{\frac{1}{N}\sum_{i=1}^{N} x_i^2 - (\bar{x}_N)^2}$$

Estimates, as functions of known constants, are themselves constants.

Expected Value of an Estimator: Bivariate OLS Example

Assume the **Linearity Condition** holds: $E[Y_i|X] = \beta_0 + \beta_1 X_i$

$$
\begin{aligned}
E[b_1^{ols}|X] &= E[\frac{\frac{1}{N}\sum_{i=1}^{N} X_i Y_i - \bar{X}_N \bar{Y}_N}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2}|X] \\
&= \frac{1}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2} \; E[\frac{1}{N}\sum_{i=1}^{N} X_i Y_i - \bar{X}_N \bar{Y}_N|X] \\
&= \left(\frac{1}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2}\right) \left(\frac{1}{N}\sum_{i=1}^{N} X_i \; E[Y_i|X] - \bar{X}_N E[\bar{Y}_N|X]\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\frac{1}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2}\right) \\
&\quad \left(\underbrace{\frac{1}{N}\sum_{i=1}^{N} X_i(\beta_0 + \beta_1 X_i)}_{=\bar{X}_N\beta_0+\beta_1\frac{1}{N}\sum_{i=1}^{N} X_i^2} - \bar{X}_N(\beta_0 + \beta_1 \bar{X}_N)\right) \\
&= \left(\frac{1}{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2}\right) \beta_1 \left(\frac{1}{N}\sum_{i=1}^{N} X_i^2 - (\bar{X}_N)^2\right) \\
&= \beta_1
\end{aligned}
$$

$$
\begin{aligned}
E[b_0^{ols}|X] &= E[\bar{Y}_N - b_1^{ols}\bar{X}_N|X] \\
&= E[\bar{Y}_N|X] - \beta_1\bar{X}_N \\
&= (\beta_0 + \beta_1\bar{X}_N) - \beta_1\bar{X}_N \\
&= \beta_0
\end{aligned}
$$

Note: $E[\bar{Y}_N|X] = E[\frac{1}{N}\sum_{i=1}^{N} Y_i|X] = \frac{1}{N}\sum_{i=1}^{N} E[Y_i|X]$

$$
= \frac{1}{N}\sum_{i=1}^{N}(\beta_0 + \beta_1 X_i) = \beta_0 + \beta_1\bar{X}_N
$$

Expected Value of an Estimator: Multivariate OLS Example

Assume **Full Rank** and **Linearity Conditions** hold:

$b^{ols} = (X'X)^{-1}X'Y$ and $E[Y|X] = X\beta$

$$
\begin{aligned}
E[b^{ols}|X] &= E[(X'X)^{-1}X'Y|X] \\
&= (X'X)^{-1}X' \underbrace{E[Y|X]}_{=X\beta} = \underbrace{(X'X)^{-1}X'X}_{I_N}\beta = \beta
\end{aligned}
$$

Variance of an Estimator: Multivariate OLS Example

Assume **all three Gauss-Markov Assumptions** hold:

$b^{ols} = (X'X)^{-1}X'Y$ and $E[Y|X] = X\beta$ and $\text{Var}(Y|X) = \sigma^2 I_N$

$$
\begin{aligned}
\text{Var}[b^{ols}|X] &= \text{Var}[\ \overbrace{(X'X)^{-1}X'}^{A}\, Y|X\ ] \\
&= \underbrace{(X'X)^{-1}X'}_{A}\ \underbrace{\text{Var}(Y|X)}_{\sigma^2 I_N}\, \underbrace{X(X'X)^{-1}}_{A'} \\
&= \sigma^2(X'X)^{-1}X' I_N X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}
\end{aligned}
$$

Note: $(X'X)$ is a symmetric matrix, as is $(X'X)^{-1}$

Variance of an Estimator: Bivariate OLS Example

Assume **all three Gauss-Markov Assumptions** hold:

$b^{ols} = (X'X)^{-1}X'Y$ and $E[Y|X] = X\beta$ and Var$(Y|X) = \sigma^2 I_N$

$$
\begin{aligned}
\text{Var}(b_1^{ols}|X) &= \frac{\sigma^2}{\sum_{i=1}^N (X_i - \bar{X}_N)^2} \\
\text{Var}(b_0^{ols}|X) &= \frac{\sigma^2}{N} \frac{\sum_{i=1}^N X_i^2}{\sum_{i=1}^N (X_i - \bar{X}_N)^2}
\end{aligned}
$$

Leave the full derivation as an exercise to reader