Statistical Assumptions and Properties of the OLS estimator

- SLR.1 $y_i = \beta_0 + \beta_1 x_i + u_i \qquad i = 1 \ldots n$
- SLR.2 $\{(x_i, y_i), i = 1..n\}$ is a random sample
- SLR.3 $\sum (x_i - \overline{x})^2 > 0$
- SLR.4 $E(u_i | x_1, x_2, \ldots x_n) = 0 \quad i = 1 \ldots n$

- SLR.3$'$ $\sum (x_i - \overline{x})^2 > 0$ and $\{x_i\}$ are fixed in repeated sampling

Rks:

- SLR.2 says that the observations are all independent.
- SLR.4 says that each disturbance is mean independent of *all* the explanatory variables x_i. With SLR.2, this reduces to $E(u_i|x_i) = 0$.
- SLR.3 is usually strengthened to SLR.3$'$ in economics. This allows us to treat the explanatory variables as fixed numbers and strictly speaking guarantees SLR.4. In practice, economists view SLR.4 as the crucial assumption, and invoke SLR.3$'$ just to avoid saying "conditional on X". This is a worthwhile simplification but it should not confuse you!

:Properties of $\widehat{\beta}_1$ under SLR.1-SLR.4

$$\widehat{\beta}_1 = \frac{\sum(x_i - \overline{x})y_i}{\sum(x_i - \overline{x})^2} \qquad (*) \qquad \text{by SLR.3}$$

$$= \frac{\sum(x_i - \overline{x})(\beta_0 + \beta_1 x_i + u_i)}{\sum(x_i - \overline{x})^2} \quad \text{by SLR.1}$$

$$= \frac{\beta_0 \cdot 0 + \beta_1 \sum(x_i - \overline{x})x_i + \sum(x_i - \overline{x})u_i)}{\sum(x_i - \overline{x})^2}$$

$$= \beta_1 + \sum_i \left(\frac{x_i - \overline{x}}{\sum(x_i - \overline{x})^2} \right) u_i$$

$$\equiv \beta_1 + \sum_i w_i u_i \qquad (**)$$

Rks:

- The weights satisfy $\sum w_i = 0$ and $\sum w_i^2 = 1/\sum(x_i - \overline{x})^2$
- $(*)$ is the formula used to compute the estimator
- $(**)$ is the formula used to derive the sampling properties

:Expected value of $\hat{\beta}_1$

Define $\underline{x} = (x_1, x_2, \ldots x_n)$.

$$E(\hat{\beta}_1|\underline{x}) = \beta_1 + E(\sum_i w_i u_i|\underline{x})$$

$$= \beta_1 + \sum_i w_i E(u_i|\underline{x}) \quad \text{given } w_i = f(\underline{x})$$

$$= \beta_1 + 0 \qquad \text{by SLR.4}$$

$\therefore$ SLR.1-SLR.4 $\Rightarrow \hat{\beta}_1$ is conditionally unbiased (given $\underline{x}$).

Rks:

- $E(\hat{\beta}_1) = E\left[E(\hat{\beta}_1|\underline{x})\right] = E(\beta_1) = \beta_1$. So conditionally unbiased implies unbiased.
- Prove $E(\hat{\beta}_0) = \beta_0$.

: Variance of $\widehat{\beta}_1$

- SLR.5 $Var(u_i|\underline{x}) = \sigma^2$ (conditional homoskedasticity)

Rks:

- In presence of SLR.2, SLR.5 is equivalent to $Var(u_i|x_i) = \sigma^2$
- If $Var(u_i|x_i) = \sigma^2(x_i)$ we have conditional heteroskedasticity
- Because $Var(u_i) = E[Var(u_i|\underline{x})] + Var[E(u_i|\underline{x})]$, with random sampling we always have unconditional homoskedasticity ($Var(u_i)$ is a constant)!

From

$$\hat{\beta}_1 = \beta_1 + \sum_i w_i u_i$$

We get

$$\begin{aligned} Var(\hat{\beta}_1 | \underline{x}) &= \sum_i w_i^2 Var(u_i | \underline{x}) + 2 \sum_{i<j} w_i w_j Cov(u_i, u_j | \underline{x}) \\ &= \sum_i w_i^2 Var(u_i | x_i) \qquad \text{by SLR.2} \\ &= \sum_i w_i^2 \sigma^2 \qquad \text{by SLR.5} \\ &= \frac{\sigma^2}{\sum (x_i - \overline{x})^2} \end{aligned}$$

Rks:

- In general, $Var(\widehat{\beta}_1|\underline{x})$ is much more interesting than $Var(\widehat{\beta}_1)$. But under SLR.3$'$, (i.e. the regressors are "fixed"), then these two objects are the same. In economics, we often write $Var(\widehat{\beta}_1) = \sigma^2/\sum(x_i - \overline{x})^2$ even though we know the regresssors are not fixed.

Definition: For SLR, OLS estimator of σ^2 is

$$\widehat{\sigma}^2 = \frac{1}{n-2}\sum \widehat{u}_i^2$$

Under SLR.1-SLR.5, $E(\widehat{\sigma}^2) = \sigma^2$. (proof with ch3)

Definition: The "standard error of the regression" is

$$\widehat{\sigma} = \sqrt{\widehat{\sigma}^2}$$

Definition: The "standard error of $\widehat{\beta}_1$ is its estimated standard deviation, i.e.

$$s.e.\left(\widehat{\beta}_1\right) = \sqrt{\frac{\widehat{\sigma}^2}{\sum(x_i - \overline{x})^2}}$$

Rk: In ch 3, we'll go beyond first two moments of the OLS estimators.

Statistical Properties using Matrix Notation

:Preliminaries

a. Expectation of a random matrix

Let Y be an (mxn) matrix of r.v.'s, i.e. $Y = (y_{ij})$ where y_{ij} is a r.v.

Definition: $E(Y) \in \mathbb{R}^{mxn}$ with $[E(Y)]_{ij} = E(y_{ij})$

Ex: $Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$ $E(Y) = \begin{bmatrix} E(y_{11}) & E(y_{12}) \\ E(y_{21}) & E(y_{22}) \end{bmatrix}$

Properties: Suppose Y^1 and Y^2 are (mxn) matrix of r.v.'s, and $A^i \in \mathbb{R}^{pxm}$, $B^i \in \mathbb{R}^{nxr}$ for $i = 1,2$. Then

- $E(A^1Y^1B^1 + A^2Y^2B^2) = A^1E(Y^1)B^1 + A^2E(Y^2)B^2$

b. Covariance matrix

Let y be an $(mx1)$ vector of r.v.'s.

Definition: The covariance matrix of y, $V(y) \in \mathbb{R}^{mxm}$, is given by

$$V(y) = E((y - Ey)(y - Ey)')$$

- $y - Ey$ is an $(mx1)$ vector of centered r.v.'s, and its outer product is the (mxm) matrix of r.v.'s given by

$$
(y - Ey)(y - Ey)'
$$

$$
= \begin{bmatrix} y_1 - Ey_1 \\ \vdots \\ y_m - Ey_m \end{bmatrix} \begin{bmatrix} y_1 - Ey_1 \cdots & y_m - Ey_m \end{bmatrix}
$$

$$
= \begin{bmatrix} (y_1 - Ey_1)^2 & \cdots & (y_1 - Ey_1)(y_m - Ey_m) \\ \vdots & \ddots & \vdots \\ (y_m - Ey_m)(y_1 - Ey_1) & \cdots & (y_m - Ey_m)^2 \end{bmatrix}
$$

Therefore,

$$V(y) = \begin{bmatrix} var(y_1) & \cdots & cov(y_1, y_m) \\ \vdots & \ddots & \vdots \\ cov(y_1, y_m) & \cdots & var y_m) \end{bmatrix}$$

So $[V(y)] = cov(y_i, y_j)$

- a.k.a. the "variance-covariance" or the "dispersion" matrix
- $V(y)$ is symmetric.

Properties: Suppose y is an $(mx1)$ vector of r.v.'s, $a \in \mathbb{R}^m$, and $B \in \mathbb{R}^{rxm}$ (or $B \in \mathbb{R}$).

- 1. $V(a + y) = V(y)$
- 2. $V(By) = BV(y)B'$

:Sampling distribution of $\widehat{\beta}$

I will rewrite the statistical assumptions slightly as

- S1 $y_i = \beta_0 + \beta_1 x_i + u_i \quad i = 1\ldots n$
- S2 $\sum(x_i - \overline{x})^2 > 0$
- S3 $E(u_i|x_1, x_2, \ldots x_n) = E(u_i|\underline{x}) = 0 \quad i = 1\ldots n$
- S4 $E(u_i u_j|\underline{x}) = \delta_{ij}\sigma^2$ where δ_{ij} is Kronecker's delta fn.
  - a) $\forall i \quad E(u_i^2|\underline{x}) = \sigma^2$
  - b) $\forall i \neq j \quad E(u_i u_j|\underline{x}) = 0$

Rks:

- SLR.1⇔S1
- SLR.3⇔S2
- SLR.2 and SLR.4⇒S3
- SLR.2 and SLR.5⇒S4

:In matrix notation

- S1 $y = X\beta + u$
- S2 $X'X$ is invertible
- S3 $E(u|X) = 0$
- S4 $V(u|X) = \sigma^2 I_n$

Using S1 and S2

$$\begin{aligned}\widehat{\beta} &= (X'X)^{-1}X'y \\ &= (X'X)^{-1}X'(X\beta + u) \\ &= \beta + (X'X)^{-1}X'u\end{aligned}$$

Define $L = (X'X)^{-1}X'$, $P = X(X'X)^{-1}X'$, $M = I - P$

:First Moment

$$
\begin{aligned}
E(\widehat{\beta}|X) &= E(\beta + Lu|X) \\
&= \beta + LE(u|X) \\
&= \beta \qquad \text{by S3}
\end{aligned}
$$

:Second Moment

$$
\begin{aligned}
V(\widehat{\beta}|X) &= V(\beta + Lu|X) \\
&= V(Lu|X) \\
&= LV(u|x)L' \quad \text{by S4} \\
&= (X'X)^{-1}X'(\sigma^2 I_n)X(X'X)^{-1} \\
&= \sigma^2 (X'X)^{-1}X'X(X'X)^{-1} \\
&= \sigma^2 (X'X)^{-1}
\end{aligned}
$$