Introductory Econometrics

Lecture 6: Gauss-Markov Theorem

Instructor: Ma, Jun

Renmin University of China

October 5, 2022

There are many alternatives estimators

- The OLS estimator is not the only estimator we can construct. There are alternative estimators with some desirable properties.
- Example: Using only the first two observations, suppose that $X_2 \neq X_1$.
$$\tilde{\beta} = \frac{Y_2 - Y_1}{X_2 - X_1}.$$
- $\tilde{\beta}$ is linear:
$$\tilde{\beta} = c_1 Y_1 + c_2 Y_2,$$
where
$$c_1 = -\frac{1}{X_2 - X_1} \text{ and } c_2 = \frac{1}{X_2 - X_1}.$$

Unbiasedness of $\tilde{\beta}$

- If $Y_i = \alpha + \beta X_i + U_i$ and $\mathrm{E}[U_i \mid X_1, \ldots, X_n] = 0$, then $\tilde{\beta}$ is unbiased:

$$\begin{aligned}
\tilde{\beta} &= \frac{Y_2 - Y_1}{X_2 - X_1} \\
&= \frac{(\alpha + \beta X_2 + U_2) - (\alpha + \beta X_1 + U_1)}{X_2 - X_1} \\
&= \frac{\beta (X_2 - X_1)}{X_2 - X_1} + \frac{U_2 - U_1}{X_2 - X_1} \\
&= \beta + \frac{U_2 - U_1}{X_2 - X_1}, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left[\tilde{\beta} \mid X_1, X_2\right] &= \beta + \mathrm{E}\left[\frac{U_2 - U_1}{X_2 - X_1} \mid X_1, X_2\right] \\
&= \beta + \frac{\mathrm{E}[U_2 \mid X_1, X_2] - \mathrm{E}[U_1 \mid X_1, X_2]}{X_2 - X_1} \\
&= \beta.
\end{aligned}$$

An optimality criterion

- Among all linear and unbiased estimators, an estimator with the smallest variance is called the Best Linear Unbiased Estimator (BLUE).
- Note that the statement is conditional on X's:
  - The estimators are unbiased conditionally on X's.
  - The variance is conditional on X's.

Gauss-Markov Theorem

Suppose that

- $Y_i = \alpha + \beta X_i + U_i$.
- $\mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] = 0$.
- $\mathrm{E}\left[U_i^2 \mid X_1, \ldots, X_n\right] = \sigma^2$ for all $i = 1, \ldots, n$ (homoskedasticity).
- For all $i \neq j$, $\mathrm{E}\left[U_i U_j \mid X_1, \ldots, X_n\right] = 0$.

Then, conditionally on X's, the OLS estimators are BLUE.

Gauss-Markov Theorem

- We already know that the OLS estimator $\hat{\beta}$ is linear and unbiased (conditionally on X's).
- Let $\tilde{\beta}$ be any other estimator of β such that
  - $\tilde{\beta}$ is linear:
    $$\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i,$$
    where c's depend only on X's.
  - $\tilde{\beta}$ is unbiased:
    $$\mathrm{E}\left[\tilde{\beta}\right] = \beta,$$
    where expectation is conditional on X's.
- We need to show that for any such $\tilde{\beta} \neq \hat{\beta}$,
  $$\mathrm{Var}\left[\tilde{\beta}\right] > \mathrm{Var}\left[\hat{\beta}\right],$$
  where the variance is conditional on X's.

An outline of the proof*

- First, we are going to show that the c's in $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$ satisfy $\sum_{i=1}^{n} c_i = 0$ and $\sum_{i=1}^{n} c_i X_i = 1$.
- Using the results of Step 1, we will show that conditionally on X's, $\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] = \text{Var}\left[\hat{\beta}\right]$.
- Using the results of Step 2, we will show that conditionally on X's, $\text{Var}\left[\tilde{\beta}\right] \geq \text{Var}\left[\hat{\beta}\right]$.
- Lastly, we will show that $\text{Var}\left[\tilde{\beta}\right] = \text{Var}\left[\hat{\beta}\right]$ if and only if $\tilde{\beta} = \hat{\beta}$.

Proof: Step 1*

- Since $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$,

$$\begin{aligned}\tilde{\beta} &= \sum_{i=1}^{n} c_i \left(\alpha + \beta X_i + U_i\right) \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i.\end{aligned}$$

- Conditionally on X's,

$$\begin{aligned}\mathrm{E}\left[\tilde{\beta}\right] &= \mathrm{E}\left[\alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i\right] \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i \mathrm{E}\left[U_i\right] \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.\end{aligned}$$

Proof: Step 1*

- From the linearity we have that, conditionally on X's,

$$\mathrm{E}\left[\tilde{\beta}\right] = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- From the unbiasedness we have that conditionally on X's,

$$\beta = \mathrm{E}\left[\tilde{\beta}\right] = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- Since this has to be true for any α and β, it follows now that

$$\begin{aligned} \sum_{i=1}^{n} c_i &= 0, \\ \sum_{i=1}^{n} c_i X_i &= 1. \end{aligned}$$

Proof: Step 2*

- We have

$$\tilde{\beta} = \beta + \sum_{i=1}^{n} c_i U_i, \text{ with } \sum_{i=1}^{n} c_i = 0, \sum_{i=1}^{n} c_i X_i = 1.$$

- Conditionally on X's,

$$\begin{aligned} \text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] &= \text{E}\left[\left(\tilde{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)\right] \\ &= \text{E}\left[\left(\sum_{i=1}^{n} c_i U_i\right)\left(\sum_{i=1}^{n} w_i U_i\right)\right] \\ &= \sum_{i=1}^{n} c_i w_i \text{E}\left[U_i^2\right] + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j \text{E}\left[U_i U_j\right]. \end{aligned}$$

Proof: Step 2*

$$\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] = \sum_{i=1}^{n} c_i w_i \text{E}\left[U_i^2\right] + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j \text{E}\left[U_i U_j\right].$$

- Since $\text{E}\left[U_i^2\right] = \sigma^2$ for all i's:

$$\sum_{i=1}^{n} c_i w_i \text{E}\left[U_i^2\right] = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

- Since $\text{E}\left[U_i U_j\right] = 0$ for all $i \neq j$,

$$\sum_{i=1}^{n} \sum_{j \neq i} c_i w_j \text{E}\left[U_i U_j\right] = 0.$$

- Thus,

$$\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

Proof: Step 2*

Conditionally on X's:

$$\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] = \sigma^2 \sum_{i=1}^{n} c_i w_i \text{ and } w_i = \frac{X_i - \bar{X}}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2}.$$

$$\begin{aligned}
\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] &= \sigma^2 \sum_{i=1}^{n} c_i \frac{X_i - \bar{X}}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \sum_{i=1}^{n} c_i \left(X_i - \bar{X}\right) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \left(\sum_{i=1}^{n} c_i X_i - \bar{X} \sum_{i=1}^{n} c_i\right) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \left(1 + \bar{X} \cdot 0\right) \\
&= \text{Var}\left[\hat{\beta}\right].
\end{aligned}$$

Proof: Step 3*

- We know now that for any linear and unbiased $\tilde{\beta}$,

$$\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] = \text{Var}\left[\hat{\beta}\right].$$

- Let's consider $\text{Var}\left[\tilde{\beta} - \hat{\beta}\right]$:

$$\begin{aligned}
\text{Var}\left[\tilde{\beta} - \hat{\beta}\right] &= \text{Var}\left[\tilde{\beta}\right] + \text{Var}\left[\hat{\beta}\right] - 2\text{Cov}\left[\tilde{\beta}, \hat{\beta}\right] \\
&= \text{Var}\left[\tilde{\beta}\right] + \text{Var}\left[\hat{\beta}\right] - 2\text{Var}\left[\hat{\beta}\right] \\
&= \text{Var}\left[\tilde{\beta}\right] - \text{Var}\left[\hat{\beta}\right].
\end{aligned}$$

- But since $\text{Var}\left[\tilde{\beta} - \hat{\beta}\right] \geq 0$,

$$\text{Var}\left[\tilde{\beta}\right] - \text{Var}\left[\hat{\beta}\right] \geq 0$$

or

$$\text{Var}\left[\tilde{\beta}\right] \geq \text{Var}\left[\hat{\beta}\right].$$

Proof: Step 4 (Uniqueness)*

Suppose that $\text{Var}\left[\tilde{\beta}\right] = \text{Var}\left[\hat{\beta}\right]$.

- Then,

$$\text{Var}\left[\tilde{\beta} - \hat{\beta}\right] = \text{Var}\left[\tilde{\beta}\right] - \text{Var}\left[\hat{\beta}\right] = 0.$$

- Thus, $\tilde{\beta} - \hat{\beta}$ is not random or

$$\tilde{\beta} - \hat{\beta} = \text{constant}.$$

- This constant also has to be zero because

$$\begin{aligned} \text{E}\left[\tilde{\beta}\right] &= \text{E}\left[\hat{\beta}\right] + \text{constant} \\ &= \beta + \text{constant}, \end{aligned}$$

and in order for $\tilde{\beta}$ to be unbiased

$$\text{constant=0 or } \tilde{\beta} = \hat{\beta}.$$