Introductory Econometrics

Lecture 5: Properties of OLS

Instructor: Ma, Jun

Renmin University of China

September 16, 2021

The OLS estimators are random variables

- The model

$$\begin{aligned} Y_i &= \alpha + \beta X_i + U_i, \\ \mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] &= 0. \end{aligned}$$

  Conditioning on X in $\mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] = 0$ allows us to treat all X's as fixed, but Y is still random.

- The estimators

$$\hat{\beta} = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) Y_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \text{and } \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

  are random because they are functions of random data.

The estimators are linear

- Since $\hat{\beta} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)Y_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$, we can write $\hat{\beta} = \sum_{i=1}^{n} w_i Y_i$, where

$$w_i = \frac{X_i - \bar{X}}{\sum_{l=1}^{n}\left(X_l - \bar{X}\right)^2}.$$

  After conditioning on X's, w_i's are not random.

- For $\hat{\alpha}$,

$$\begin{aligned}
\hat{\alpha} &= \bar{Y} - \hat{\beta}\bar{X} \\
&= \frac{1}{n}\sum_{i=1}^{n} Y_i - \left(\sum_{i=1}^{n} w_i Y_i\right)\bar{X} \\
&= \sum_{i=1}^{n}\left(\frac{1}{n} - \bar{X}w_i\right)Y_i \\
&= \sum_{i=1}^{n}\left(\frac{1}{n} - \bar{X}\frac{X_i - \bar{X}}{\sum_{l=1}^{n}\left(X_l - \bar{X}\right)^2}\right)Y_i.
\end{aligned}$$

Unbiasedness

- $\hat{\beta}$ is called an unbiased estimator if $\mathrm{E}\left[\hat{\beta}\right] = \beta$.
- Suppose that $Y_i = \alpha + \beta X_i + U_i$, $\mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] = 0$. Then $\mathrm{E}\left[\hat{\beta}\right] = \beta$.

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) Y_i}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} \\
&= \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)\left(\alpha + \beta X_i + U_i\right)}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} \\
&= \alpha \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} + \beta \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) X_i}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} + \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} \\
&= \alpha \frac{0}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} + \beta \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} + \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2},
\end{aligned}$$

- or

$$\hat{\beta} = \beta + \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}.$$

Unbiasedness

- Once we condition on $X_1, \ldots, X_n$, all X's in

$$\hat{\beta} = \beta + \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}$$

can be treated as fixed.

- Thus,

$$\begin{aligned}
\mathrm{E}\left[\hat{\beta} \mid X_1, \ldots, X_n\right] &= \mathrm{E}\left[\beta + \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \mid X_1, \ldots, X_n\right] \\
&= \beta + \mathrm{E}\left[\frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) U_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \mid X_1, \ldots, X_n\right] \\
&= \beta + \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) \mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right]}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}.
\end{aligned}$$

Unbiasedness

- Thus, with $\mathrm{E}[U_i \mid X_1, \ldots, X_n] = 0$, we have

$$\begin{aligned} \mathrm{E}\left[\hat{\beta} \mid X_1, \ldots, X_n\right] &= \beta + \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) \mathrm{E}[U_i \mid X_1, \ldots, X_n]}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} \\ &= \beta + \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right) \cdot 0}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} = 0. \end{aligned}$$

- By the LIE, $\mathrm{E}\left[\hat{\beta}\right] = \mathrm{E}\left[\mathrm{E}\left[\hat{\beta} \mid X_1, \ldots, X_n\right]\right] = \mathrm{E}[\beta] = \beta$.

Strong exogeneity of regressors

- The regressor X is strongly exogenous if $\mathrm{E}[U_i \mid X_1, \ldots, X_n] = 0$.
- Alternatively, we can assume that $\mathrm{E}[U_i \mid X_i] = 0$ and all observations are independent:

$$\begin{aligned} \mathrm{E}[U_1 \mid X_1, \ldots, X_n] &= \mathrm{E}[U_1 \mid X_1], \\ \mathrm{E}[U_2 \mid X_1, \ldots, X_n] &= \mathrm{E}[U_2 \mid X_2] \text{ and etc.} \end{aligned}$$

- The OLS estimator is in general biased if the strong exogeneity assumption is violated.

Variance of $\hat{\beta}$

- If $Y_i = \alpha + \beta X_i + U_i, \mathrm{E}[U_i \mid X_1, \ldots, X_n] = 0$, and
$$\mathrm{E}\left[U_i^2 \mid X_1, \ldots, X_n\right] = \sigma^2 = \text{constant},$$
and for $i \neq j$
$$\mathrm{E}\left[U_i U_j \mid X_1, \ldots, X_n\right] = 0,$$
Then
$$\mathrm{Var}\left[\hat{\beta} \mid X_1, \ldots, X_n\right] = \frac{\sigma^2}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}.$$
- The assumption $\mathrm{E}\left[U_i^2 \mid X_1, \ldots, X_n\right] = \sigma^2$ =constant is called (conditional) homoskedasticity.
- The assumption $\mathrm{E}\left[U_i U_j \mid X_1, \ldots, X_n\right] = 0$ for $i \neq j$ can be replaced by the assumption that the observations are independent.

Variance of $\hat{\beta}$

$$\mathrm{Var}\left[\hat{\beta} \mid X_1, \ldots, X_n\right] = \frac{\sigma^2}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}.$$

- The variance of $\hat{\beta}$ is positively related to the variance of the errors $\sigma^2 = \mathrm{Var}\left[U_i\right]$.
- The variance of $\hat{\beta}$ is smaller when X's are more dispersed.

Variance of $\hat{\beta}$

- We are going to condition on X's and will treat them as constants. All expectations below are implicitly conditional on X's.
- We have $\hat{\beta} = \beta + \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$ and $\mathrm{E}\left[\hat{\beta}\right] = \beta,$ conditional on X's,

$$
\begin{aligned}
\mathrm{Var}\left[\hat{\beta}\right] &= \mathrm{E}\left[\left(\hat{\beta} - \mathrm{E}\left[\hat{\beta}\right]\right)^2\right] \\
&= \mathrm{E}\left[\left(\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right)^2\right] \\
&= \left(\frac{1}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right)^2 \mathrm{E}\left[\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i\right)^2\right].
\end{aligned}
$$

Variance of $\hat{\beta}$

- $$\begin{aligned} & \left(\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i\right)^2 \\ = & \sum_{i=1}^{n}\sum_{j=1}^{n}\left(X_i-\bar{X}\right)\left(X_j-\bar{X}\right)U_iU_j \\ = & \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2U_i^2+\sum_{i=1}^{n}\sum_{j\neq i}\left(X_i-\bar{X}\right)\left(X_j-\bar{X}\right)U_iU_j. \end{aligned}$$
- Since $\mathrm{E}\left[U_iU_j\right]=0$ for $i\neq j$,

$$\begin{aligned} \mathrm{E}\left[\left(\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i\right)^2\right] &= \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\mathrm{E}\left[U_i^2\right]+0 \\ &= \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\sigma^2 \end{aligned}$$

Variance of $\hat{\beta}$

We have

$$\mathrm{Var}\left[\hat{\beta}\right] = \left(\frac{1}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right)^2 \mathrm{E}\left[\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i\right)^2\right],$$

$$\mathrm{E}\left[\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i\right)^2\right] = \sigma^2 \sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2,$$

and therefore,

$$\begin{aligned}\mathrm{Var}\left[\hat{\beta}\right] &= \left(\frac{1}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right)^2 \sigma^2 \sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \\ &= \left(\frac{1}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right)\sigma^2\end{aligned}$$

Normality of $\hat{\beta}$

- Assume that U_i's are jointly normally distributed conditional on X's.
- Then $Y_i = \alpha + \beta X_i + U_i$ are also jointly normally distributed.
- Since $\hat{\beta} = \sum_{i=1}^{n} w_i Y_i$, where $w_i = \frac{X_i - \bar{X}}{\sum_{l=1}^{n}\left(X_l - \bar{X}\right)^2}$ depend only on X's, $\hat{\beta}$ is also normally distributed conditional on X's.
- Conditional on $X_1, \ldots, X_n$

$$
\begin{aligned}
\hat{\beta} &\sim \mathrm{N}\left(\mathrm{E}\left[\hat{\beta}\right], \mathrm{Var}\left[\hat{\beta}\right]\right) \\
&\sim \mathrm{N}\left(\beta, \frac{\sigma^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right).
\end{aligned}
$$