Lecture 4: Properties of OLS

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

April 5, 2026

Properties of Estimators

1. Random
2. Mean
3. Variance
4. Distribution

OLS Estimators as Random Variables

- The model

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

  Conditioning on $X_1, \dots, X_n$ allows us to treat all the X_i's as fixed, but Y_i is still random.
- To save on writing, we will use the notation

$$\mathrm{E}\left[\cdot \mid \mathbf{X}\right] = \mathrm{E}\left[\cdot \mid X_1, \dots, X_n\right].$$

  That is, $\mathbf{X} = (X_1, \dots, X_n)$.
- 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 the random Y_i's even after conditioning on the X_i's.

Linearity of Estimators

- 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

Definition and Claim

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

$$\mathrm{E}\left[\hat{\beta}\right] = \beta.$$

Proof Step 1: Decomposition into signal and noise

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

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

Proof Step 2: Conditioning on Regressors

- Once we condition on $\mathbf{X}$, all the X_i'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,

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

Proof Step 3

- Thus, with $\mathrm{E}\left[U_i \mid \mathbf{X}\right] = 0$, we have

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

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

Strong Exogeneity of Regressors

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

$$\begin{aligned}
&\mathrm{E}\left[U_1 \mid \mathbf{X}\right] = \mathrm{E}\left[U_1 \mid X_1\right], \\
&\mathrm{E}\left[U_2 \mid \mathbf{X}\right] = \mathrm{E}\left[U_2 \mid X_2\right] \text{ etc.}
\end{aligned}$$

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

Variance of the Slope Estimator

Variance Formula and Homoskedasticity

- If $Y_i = \alpha + \beta X_i + U_i$, $\mathrm{E}\left[U_i \mid \mathbf{X}\right] = 0$, and

$$\mathrm{E}\left[U_i^2 \mid \mathbf{X}\right] = \sigma^2 = \text{constant},$$

  and for $i \neq j$

$$\mathrm{E}\left[U_i U_j \mid \mathbf{X}\right] = 0,$$

  then

$$\mathrm{Var}\left(\hat{\beta} \mid \mathbf{X}\right) = \frac{\sigma^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}.$$

- The assumption $\mathrm{E}\left[U_i^2 \mid \mathbf{X}\right] = \sigma^2 = \text{constant}$ is called (conditional) homoskedasticity.
- The assumption $\mathrm{E}\left[U_i U_j \mid \mathbf{X}\right] = 0$ for $i \neq j$ can be replaced by the assumption that the observations are independent.

Determinants of Variance

$$\mathrm{Var}\left(\hat{\beta} \mid \mathbf{X}\right) = \frac{\sigma^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}.$$

- 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 the X_i's are more dispersed.

Derivation of Variance

- We have $\hat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})U_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$ and $\mathrm{E}\left[\hat{\beta} \mid \mathbf{X}\right] = \beta$.

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

- Expanding the square,

$$\begin{aligned}
\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)U_i\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(X_i - \bar{X}\right)\left(X_j - \bar{X}\right)U_iU_j \\
&= \sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 U_i^2 \\
&\quad + \sum_{i=1}^{n}\sum_{j \neq i}\left(X_i - \bar{X}\right)\left(X_j - \bar{X}\right)U_iU_j.
\end{aligned}$$

- Since $\mathrm{E}\left[U_iU_j \mid \mathbf{X}\right] = 0$ for $i \neq j$,

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

- We have

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

  and therefore,

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

Distribution of the Slope Estimator

Normality of the OLS Estimator

- 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 the X_i's, $\hat{\beta}$ is also normally distributed conditional on the X_i's.
- Conditional on $\mathbf{X}$

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