Introductory Econometrics

Lecture 7: Estimating the variance of errors

Instructor: Ma, Jun

Renmin University of China

September 27, 2021

The importance of σ^2

- The variance of $\hat{\beta}$ depends on unknown $\sigma^2 = \mathrm{E}\left[U_i^2\right]$:

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

- If U's were observable, we could estimate σ^2 by $\frac{1}{n}\sum_{i=1}^n U_i^2$.
- Note that $\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n U_i^2\right] = \frac{1}{n}\sum_{i=1}^n \mathrm{E}\left[U_i^2\right] = \frac{1}{n}\sum_{i=1}^n \sigma^2 = \sigma^2$.
- However, such an estimator is infeasible.
- Instead, we have $\hat{U}$'s:

$$\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i.$$

- A feasible estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n \hat{U}_i^2.$$

- It turns out that $\hat{\sigma}^2$ is biased .

An unbiased estimator of σ^2

An unbiased estimator of σ^2 is:

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} \hat{U}_i^2.$$

For the unbiasedness of s^2, we will use the following assumptions:

1. $Y_i = \alpha + \beta X_i + U_i$,
2. $\mathrm{E}\left[U_i | X_1, \ldots, X_n\right] = 0$ for all i's,
3. $\mathrm{E}\left[U_i^2 | X_1, \ldots, X_n\right] = \sigma^2$ for all i's,
4. $\mathrm{E}\left[U_i U_j | X_1, \ldots, X_n\right] = 0$ for all $i \neq j$.

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} \hat{U}_i^2.$$

- In order to construct $\hat{U}_i$, first we need to estimate two parameters α and β :

$$\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i,$$

  and s^2 adjusts for the estimation of α and β by dividing by $n-2$ instead of n.

- To show that $\mathrm{E}\left[s^2\right] = \sigma^2$ (unbiasedness), we need to show that

$$\mathrm{E}\left[\sum_{i=1}^{n} \hat{U}_i^2\right] = (n-2)\,\sigma^2.$$

- We need to express $\hat{U}$'s using U's, since $\mathrm{E}\left[U_i^2\right] = \sigma^2$.

Expansion of $\hat{U}_i$

From

$$\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i, \text{ and}$$

we have

$$\begin{aligned} \hat{U}_i &= Y_i - \left(\bar{Y} - \hat{\beta}\bar{X}\right) - \hat{\beta} X_i \\ &= \left(Y_i - \bar{Y}\right) - \hat{\beta}\left(X_i - \bar{X}\right). \end{aligned} \quad (1)$$

Next,

$$\begin{aligned} Y_i &= \alpha + \beta X_i + U_i, \\ \bar{Y} &= \alpha + \beta \bar{X} + \bar{U}, \text{ and} \\ Y_i - \bar{Y} &= \beta\left(X_i - \bar{X}\right) + U_i - \bar{U}. \end{aligned} \quad (2)$$

By combining (1) and (2),

$$\hat{U}_i = \left(U_i - \bar{U}\right) - \left(\hat{\beta} - \beta\right)\left(X_i - \bar{X}\right).$$

Expansion of $\sum_{i=1}^{n} \hat{U}_i^2$

From

$$\hat{U}_i = \left(U_i - \bar{U}\right) - \left(\hat{\beta} - \beta\right)\left(X_i - \bar{X}\right),$$

We have

$$\begin{aligned}
\hat{U}_i^2 &= \left[\left(U_i - \bar{U}\right) - \left(\hat{\beta} - \beta\right)\left(X_i - \bar{X}\right)\right]^2 \\
&= \left(U_i - \bar{U}\right)^2 + \left(\hat{\beta} - \beta\right)^2 \left(X_i - \bar{X}\right)^2 - \\
&\quad -2\left(\hat{\beta} - \beta\right)\left(X_i - \bar{X}\right)\left(U_i - \bar{U}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{i=1}^{n} \hat{U}_i^2 &= \sum_{i=1}^{n} \left(U_i - \bar{U}\right)^2 + \left(\hat{\beta} - \beta\right)^2 \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \\
&\quad -2\left(\hat{\beta} - \beta\right) \sum_{i=1}^{n} \left(X_i - \bar{X}\right)\left(U_i - \bar{U}\right).
\end{aligned}$$

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

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

We will show that

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

and therefore,

$$\mathrm{E}\left[\sum_{i=1}^{n} \hat{U}_i^2\right] = (n-1)\,\sigma^2 + \sigma^2 - 2\sigma^2 = (n-2)\,\sigma^2.$$

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

First,

$$\begin{aligned}\sum_{i=1}^{n}\left(U_i-\bar{U}\right)^2 &= \sum_{i=1}^{n}\left(U_i-\bar{U}\right)U_i \\ &= \sum_{i=1}^{n}U_i^2-\bar{U}\sum_{i=1}^{n}U_i.\end{aligned}$$

$$
\begin{aligned}
\sum_{i=1}^{n}\left(U_i-\bar{U}\right)^2 &= \sum_{i=1}^{n} U_i^2-\frac{1}{n}\left(\sum_{i=1}^{n} U_i\right)^2 \\
&= \sum_{i=1}^{n} U_i^2-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} U_i U_j \\
&= \sum_{i=1}^{n} U_i^2-\frac{1}{n}\left(\sum_{i=1}^{n} U_i^2+\sum_{i=1}^{n}\sum_{j\neq i} U_i U_j\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left[\sum_{i=1}^{n}\left(U_i-\bar{U}\right)^2\right] &= \sum_{i=1}^{n}\mathrm{E}\left[U_i^2\right]-\frac{1}{n}\left(\sum_{i=1}^{n}\mathrm{E}\left[U_i^2\right]+\sum_{i=1}^{n}\sum_{j\neq i}\mathrm{E}\left[U_i U_j\right]\right) \\
&= \sum_{i=1}^{n}\sigma^2-\frac{1}{n}\left(\sum_{i=1}^{n}\sigma^2+0\right)=n\sigma^2-\frac{1}{n}n\sigma^2 \\
&= (n-1)\,\sigma^2.
\end{aligned}
$$

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

- First, note that since conditionally on X's,

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

  we have that conditionally on X's,

$$\begin{aligned}\mathrm{E}\left[\left(\hat{\beta}-\beta\right)^2\right] &= \mathrm{E}\left[\left(\hat{\beta}-\mathrm{E}\left[\hat{\beta}\right]\right)^2\right] \\ &= \mathrm{Var}\left[\hat{\beta}\right].\end{aligned}$$

- The conditional variance of $\hat{\beta}$ given $X_1,\ldots,X_n$ is

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

- Thus,

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

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

- First,
$$\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(U_i-\bar{U}\right)=\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i.$$
- Next,
$$\hat{\beta}-\beta=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}.$$
- Thus,
$$\begin{aligned}&\left(\hat{\beta}-\beta\right)\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(U_i-\bar{U}\right)\\ =\;&\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i\\ =\;&\frac{1}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}\right)U_i\right)^2.\end{aligned}$$

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

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

Conditional on X's,

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

Estimation of the variance of $\hat{\beta}$

- The variance of $\hat{\beta}$ (conditional on X's):

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

  It is unknown because σ^2 is unknown.

- The estimator of σ^2 :

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{U}_i^2 = \frac{1}{n-2}\sum_{i=1}^{n}\left(Y_i - \hat{\alpha} - \hat{\beta}X_i\right)^2.$$

- The estimator for the variance of $\hat{\beta}$:

$$\widehat{\operatorname{Var}}\left[\hat{\beta}\right] = \frac{s^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}.$$

- The standard errorof β:

$$\operatorname{SE}\left[\hat{\beta}\right] = \sqrt{\widehat{\operatorname{Var}}\left[\hat{\beta}\right]} = \sqrt{\frac{s^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}}.$$