Lecture 6: Estimating the variance of errors

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

The importance of σ^2

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

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

- If U's were observable, we could estimate σ^2 by $\frac{1}{n}\sum_{i=1}^n U_i^2$, which is unbiased, but infeasible.
- Using residuals $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$ gives a feasible estimator

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

  but $\hat{\sigma}^2$ is biased.

An unbiased estimator of σ^2

- An unbiased estimator of σ^2 is

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

- Assumptions:
  1. $Y_i = \alpha + \beta X_i + U_i$.
  2. $E\left(U_i \mid X_1, \dots, X_n\right) = 0$ for all i.
  3. $E\left(U_i^2 \mid X_1, \dots, X_n\right) = \sigma^2$ for all i.
  4. $E\left(U_i U_j \mid X_1, \dots, X_n\right) = 0$ for all $i \neq j$.
- Since $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$, dividing by $n-2$ adjusts for estimating α and β.

Expressing $\hat{U}_i$ in terms of U_i

From

$$\begin{aligned} \hat{U}_i &= \left(Y_i - \bar{Y}\right) - \hat{\beta}\left(X_i - \bar{X}\right), \\ Y_i - \bar{Y} &= \beta\left(X_i - \bar{X}\right) + U_i - \bar{U}, \end{aligned}$$

we get

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

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

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

Thus,

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

To show $E\sum_{i=1}^n \hat{U}_i^2 = (n-2)\,\sigma^2$, we verify the three expectations below.

Expectation of $\sum_{i=1}^{n}\left(U_i-\bar{U}\right)^2$

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

Taking expectations and using the assumptions,

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

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

Because $E\hat{\beta}=\beta$ (conditionally on X),

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

Hence,

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

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

Note that

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

and

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

Therefore,

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

Conditionally on X's,

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

Putting it together

Using the three expectations above,

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

so s^2 is unbiased for σ^2.

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

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

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

- Estimator of σ^2:

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

- Estimator of the variance of $\hat{\beta}$:

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

- Standard error of $\hat{\beta}$:

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