Lecture 6: Estimating the variance of errors

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

April 5, 2026

The importance of σ^2

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

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

- If U's were observable, we could estimate σ^2 by $\frac{1}{n}\sum_{i=1}^n U_i^2$, which is unbiased. This is not possible as U's are unobservable.
- Using sample residuals instead, $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$, gives a feasible estimator:

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

- $\hat{\sigma}^2$ 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. $\mathrm{E}\left[U_i \mid \mathbf{X}\right] = 0$ for all i.
  3. $\mathrm{E}\left[U_i^2 \mid \mathbf{X}\right] = \sigma^2$ for all i.
  4. $\mathrm{E}\left[U_i U_j \mid \mathbf{X}\right] = 0$ for all $i \neq j$.
- Since $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$, dividing by $n-2$ adjusts for estimating two parameters: α and β.

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

- $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$
- $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$, so

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

- Also:

$$Y_i - \bar{Y} = \beta\left(X_i - \bar{X}\right) + U_i - \bar{U},$$

- We have the following relationship between $\hat{U}_i$ and U_i:

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

- $\hat{U}_i$ is related to U_i, but it is contaminated by the estimation errors.

Expanding $\sum \hat{U}_i^2$

- We have:

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

- The squared residual is:

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

- Thus,

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

- To show $\mathrm{E}\left[\sum_{i=1}^{n} \hat{U}_i^2 \mid \mathbf{X}\right] = (n-2)\,\sigma^2$, we verify the three terms on the RHS:
- **Claim 1:**

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

- **Claim 2:**

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

- **Claim 3:**

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

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

- Variance of $\hat{\beta}$ conditional on $\mathbf{X}$:

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

- Estimator of σ^2:

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

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

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

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

$$\begin{aligned}
\text{se}\left(\hat{\beta}\right) &= \sqrt{\widehat{\text{Var}}\left(\hat{\beta}\right)} \\
&= \sqrt{\frac{s^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}}.
\end{aligned}$$

Example in R

- Regress hourly wage on years of education using the `wage1` dataset from Wooldridge:

```r
library(wooldridge)
data("wage1")
fit <- lm(wage ~ educ, data = wage1)
summary(fit)
```

```

Call:
lm(formula = wage ~ educ, data = wage1)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.3396 -2.1501 -0.9674  1.1921 16.6085 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.90485    0.68497  -1.321    0.187    
educ         0.54136    0.05325  10.167   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.378 on 524 degrees of freedom
Multiple R-squared:  0.1648,	Adjusted R-squared:  0.1632 
F-statistic: 103.4 on 1 and 524 DF,  p-value: < 2.2e-16
```

- `Std. Error` column: standard errors $\text{se}\left(\hat{\alpha}\right)$ and $\text{se}\left(\hat{\beta}\right)$
- `Residual standard error`: $s = \sqrt{s^2}$; so $s^2 = 3.378^2 \approx 11.41$
- The estimate s^2 can also be extracted directly:

```r
summary(fit)$sigma^2
```

```
[1] 11.41352
```

Proof of Claim 1

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

Taking conditional expectations and using the assumptions,

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

Proof of Claim 2

Because $\mathrm{E}\left[\hat{\beta}\mid\mathbf{X}\right]=\beta$,

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

Hence,

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

Proof of Claim 3

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,

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

Conditionally on $\mathbf{X}$,

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

Putting it together

Using the three expectations above,

$$\begin{aligned}
\mathrm{E}\left[\sum_{i=1}^{n}\hat{U}_i^2 \mid \mathbf{X}\right] &= (n-1)\,\sigma^2+\sigma^2-2\sigma^2 \\
&= (n-2)\,\sigma^2,
\end{aligned}$$

so s^2 is unbiased for σ^2.