Economics 326
Methods of Empirical Research in Economics

Lecture 5: Gauss-Markov Theorem

Vadim Marmer
University of British Columbia

January 27, 2011

There are many alternatives estimators

- The OLS estimator is not the only estimator we can construct. There are alternative estimators with some desirable properties.
- Example: Using only the first two observations, suppose that $X_2 \neq X_1$.

$$\tilde{\beta} = \frac{Y_2 - Y_1}{X_2 - X_1}.$$

- $\tilde{\beta}$ is linear:

$$\tilde{\beta} = c_1 Y_1 + c_2 Y_2,$$

where

$$c_1 = -\frac{1}{X_2 - X_1} \text{ and } c_2 = \frac{1}{X_2 - X_1}.$$

Unbiasedness of $\tilde{\beta}$

- If $Y_i = \alpha + \beta X_i + U_i$ and $E\left(U_i | X_1, \ldots, X_n\right) = 0$, then $\tilde{\beta}$ is unbiased:

$$
\begin{aligned}
\tilde{\beta} &= \frac{Y_2 - Y_1}{X_2 - X_1} \\
&= \frac{(\alpha + \beta X_2 + U_2) - (\alpha + \beta X_1 + U_1)}{X_2 - X_1} \\
&= \frac{\beta\left(X_2 - X_1\right)}{X_2 - X_1} + \frac{U_2 - U_1}{X_2 - X_1} \\
&= \beta + \frac{U_2 - U_1}{X_2 - X_1}, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
E\left(\tilde{\beta} | X_1, X_2\right) &= \beta + E\left(\frac{U_2 - U_1}{X_2 - X_1} | X_1, X_2\right) \\
&= \beta + \frac{E\left(U_2 | X_1, X_2\right) - E\left(U_1 | X_1, X_2\right)}{X_2 - X_1} \\
&= \beta.
\end{aligned}
$$

An optimality criterion

- Among all linear and unbiased estimators, an estimator with the smallest variance is called the Best Linear Unbiased Estimator(BLUE).
- Note that the statement is conditional on X's:
  - The estimators are unbiased conditionally on X's.
  - The variance is conditional on X's.

Gauss-Markov Theorem

Suppose that

- $Y_i = \alpha + \beta X_i + U_i$.
- $E\left(U_i | X_1, \ldots, X_n\right) = 0$.
- $E\left(U_i^2 | X_1, \ldots, X_n\right) = \sigma^2$ for all $i = 1, \ldots, n$ (homoskedasticity).
- For all $i \neq j$, $E\left(U_i U_j | X_1, \ldots, X_n\right) = 0$.

Then, conditionally on X's, the OLS estimators are BLUE.

Gauss-Markov Theorem

- We already know that the OLS estimator $\hat{\beta}$ is linear and unbiased (conditionally on X's).
- Let $\tilde{\beta}$ be any other estimator of β such that
  - $\tilde{\beta}$ is linear:
    $$\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i,$$
    where c's depend only on X's.
  - $\tilde{\beta}$ is unbiased:
    $$E\tilde{\beta} = \beta,$$
    where expectation is conditional on X's.
- We need to show that for any such $\tilde{\beta} \neq \hat{\beta}$,
  $$Var\left(\tilde{\beta}\right) > Var\left(\hat{\beta}\right),$$
  where the variance is conditional on X's.

An outline of the proof

1. First, we are going to show that the c's in $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$ satisfy $\sum_{i=1}^{n} c_i = 0$ and $\sum_{i=1}^{n} c_i X_i = 1$.
2. Using the results of Step 1, we will show that conditionally on X's, $Cov\left(\tilde{\beta}, \hat{\beta}\right) = Var\left(\hat{\beta}\right)$.
3. Using the results of Step 2, we will show that conditionally on X's, $Var\left(\tilde{\beta}\right) \geq Var\left(\hat{\beta}\right)$.
4. Lastly, we will show that $Var\left(\tilde{\beta}\right) = Var\left(\hat{\beta}\right)$ if and only if $\tilde{\beta} = \hat{\beta}$.

Proof: Step 1

- Since $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$,

$$\begin{aligned}\tilde{\beta} &= \sum_{i=1}^{n} c_i \left(\alpha + \beta X_i + U_i\right) \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i.\end{aligned}$$

- Conditionally on X's,

$$\begin{aligned}E\tilde{\beta} &= E\left(\alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i\right) \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i E U_i \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.\end{aligned}$$

Proof: Step 1

- From the linearity we have that, conditionally on X's,

$$E\tilde{\beta} = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- From the unbiasedness we have that conditionally on X's,

$$\beta = E\tilde{\beta} = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- Since this has to be true for any α and β, it follows now that

$$\begin{aligned} \sum_{i=1}^{n} c_i &= 0, \\ \sum_{i=1}^{n} c_i X_i &= 1. \end{aligned}$$

Proof: Step 2

- We have

$$
\begin{aligned}
\tilde{\beta} &= \beta + \sum_{i=1}^{n} c_i U_i, \text{ with } \sum_{i=1}^{n} c_i = 0, \sum_{i=1}^{n} c_i X_i = 1. \\
\hat{\beta} &= \beta + \sum_{i=1}^{n} w_i U_i, \text{ with } w_i = \frac{X_i - \bar{X}}{\sum_{j=1}^{n} \left(X_j - \bar{X}\right)^2}.
\end{aligned}
$$

- Conditionally on X's,

$$
\begin{aligned}
Cov\left(\tilde{\beta}, \hat{\beta}\right) &= E\left[\left(\tilde{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)\right] \\
&= E\left[\left(\sum_{i=1}^{n} c_i U_i\right)\left(\sum_{i=1}^{n} w_i U_i\right)\right] \\
&= \sum_{i=1}^{n} c_i w_i E\left(U_i^2\right) + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E\left(U_i U_j\right).
\end{aligned}
$$

Proof: Step 2

$$Cov\left(\tilde{\beta}, \hat{\beta}\right) = \sum_{i=1}^{n} c_i w_i E\left(U_i^2\right) + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E\left(U_i U_j\right).$$

- Since $E\left(U_i^2\right) = \sigma^2$ for all i's:

$$\sum_{i=1}^{n} c_i w_i E\left(U_i^2\right) = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

- Since $E\left(U_i U_j\right) = 0$ for all $i \neq j$,

$$\sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E\left(U_i U_j\right) = 0.$$

- Thus,

$$Cov\left(\tilde{\beta}, \hat{\beta}\right) = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

Proof: Step 2

Conditionally on X's:

$$Cov\left(\tilde{\beta}, \hat{\beta}\right) = \sigma^2 \sum_{i=1}^{n} c_i w_i \text{ and } w_i = \frac{X_i - \bar{X}}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2}.$$

$$\begin{aligned}
Cov\left(\tilde{\beta}, \hat{\beta}\right) &= \sigma^2 \sum_{i=1}^{n} c_i \frac{X_i - \bar{X}}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \sum_{i=1}^{n} c_i \left(X_i - \bar{X}\right) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \left(\sum_{i=1}^{n} c_i X_i - \bar{X} \sum_{i=1}^{n} c_i\right) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \left(1 + \bar{X} \cdot 0\right) \\
&= Var\left(\hat{\beta}\right).
\end{aligned}$$

Proof: Step 3

- We know now that for any linear and unbiased $\tilde{\beta}$,
$$Cov\left(\tilde{\beta}, \hat{\beta}\right) = Var\left(\hat{\beta}\right).$$

- Let's consider $Var\left(\tilde{\beta} - \hat{\beta}\right)$:
$$\begin{aligned} Var\left(\tilde{\beta} - \hat{\beta}\right) &= Var\left(\tilde{\beta}\right) + Var\left(\hat{\beta}\right) - 2Cov\left(\tilde{\beta}, \hat{\beta}\right) \\ &= Var\left(\tilde{\beta}\right) + Var\left(\hat{\beta}\right) - 2Var\left(\hat{\beta}\right) \\ &= Var\left(\tilde{\beta}\right) - Var\left(\hat{\beta}\right). \end{aligned}$$

- But since $Var\left(\tilde{\beta} - \hat{\beta}\right) \geq 0$,
$$Var\left(\tilde{\beta}\right) - Var\left(\hat{\beta}\right) \geq 0$$
or
$$Var\left(\tilde{\beta}\right) \geq Var\left(\hat{\beta}\right).$$

Proof: Step 4 (Uniqueness)

Suppose that $Var\left(\tilde{\beta}\right) = Var\left(\hat{\beta}\right)$.

- Then,
$$Var\left(\tilde{\beta} - \hat{\beta}\right) = Var\left(\tilde{\beta}\right) - Var\left(\hat{\beta}\right) = 0.$$
- Thus, $\tilde{\beta} - \hat{\beta}$ is not random or
$$\tilde{\beta} - \hat{\beta} = \text{constant}.$$
- This constant also has to be zero because
$$\begin{aligned} E\tilde{\beta} &= E\hat{\beta} + \text{constant} \\ &= \beta + \text{constant}, \end{aligned}$$
and in order for $\tilde{\beta}$ to be unbiased
$$\text{constant}=0 \text{ or } \tilde{\beta} = \hat{\beta}.$$