Unbiasedness of s^2

Linda Maokomatanda, Tim Mensinger, Markus Schick

University of Bonn

4. November 2019

Content

1. Introduction (Linda)
   - Assumptions
   - Statement of Theorem
2. Proof (Markus)
3. Wrap-up (Tim)

Assumption 1 - Linearity

$$\boldsymbol{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$$

Assumption 2 - Strict Exogeneity

$$\mathbb{E}\left[\boldsymbol{\varepsilon} \mid \mathbf{X}\right] = \mathbf{0}$$

Assumption 3 - Rank Condition

$$\text{rank}(\mathbf{X}) = K \quad \text{a.s.}$$

Assumption 4 - Spherical Error

$$\mathbb{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \sigma^2 \, \mathbf{id}_n$$

for some *fixed* $\sigma > 0$

Statement of Theorem

Theorem

Under assumptions 1-4, we have that

$$\mathbb{E}\left[s^2 \mid \mathbf{X}\right] = \sigma^2$$

where $s^2 = \hat{\varepsilon}'\hat{\varepsilon}/(n-K)$

Proof

WTS:

$$\mathbb{E}\left[s^2 \mid \mathbf{X}\right] = \mathbb{E}\left[\hat{\varepsilon}'\hat{\varepsilon}/(n-K) \mid \mathbf{X}\right] = \sigma^2$$

Strategy:

1. Show $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon'\mathbf{M}\varepsilon$
2. Show $\mathbb{E}\left[\varepsilon'\mathbf{M}\varepsilon \mid \mathbf{X}\right] = \sigma^2 \operatorname{trace}(\mathbf{M})$
3. Show $\operatorname{trace}(\mathbf{M}) = n - K$

Remember that $\hat{\varepsilon} = \mathbf{M}\boldsymbol{y}$. Using $\boldsymbol{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$ and Lemma 3.1.1 it follows trivially that

$$\begin{aligned}
\hat{\varepsilon}'\hat{\varepsilon} &= (\mathbf{My})'\mathbf{My} \\
&= (\mathbf{M}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}))'\mathbf{M}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\
&= (\mathbf{M}\boldsymbol{\varepsilon})'\mathbf{M}\boldsymbol{\varepsilon} \\
&= \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}
\end{aligned}$$

2.

Note that $\mathbf{M} = \mathbf{M}(\mathbf{X})$ and $\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon} = \sum_{i,j=1}^{n} m_{ij}\varepsilon_i\varepsilon_j$. Therefore

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon} \mid \mathbf{X}\right] &= \sum_{i,j=1}^{n} m_{ij}\mathbb{E}\left[\varepsilon_i\varepsilon_j \mid \mathbf{X}\right] \\
&= \sum_{i=1}^{n} m_{ii}\sigma^2 \\
&= \sigma^2 \operatorname{trace}(\mathbf{M})
\end{aligned}$$

3.

$$
\begin{aligned}
\text{trace}(\mathbf{M}) &= \text{trace}(\mathbf{id}_n - \boldsymbol{P}) \\
&= \text{trace}(\mathbf{id}_n) - trace(\boldsymbol{P}) \\
&= n - \text{trace}(\boldsymbol{P}) \\
&= n - \text{trace}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \\
&= n - \text{trace}(\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}) \\
&= n - \text{trace}(\mathbf{id}_K) \\
&= n - K
\end{aligned}
$$

Proof

Combining 1 - 3 we get

$$\begin{aligned}\mathbb{E}\left[s^2 \mid \mathbf{X}\right] &= \mathbb{E}\left[\hat{\varepsilon}'\hat{\varepsilon}/(n-K) \mid \mathbf{X}\right] \\ &= \mathbb{E}\left[\varepsilon'\mathbf{M}\varepsilon \mid \mathbf{X}\right]/(n-K) \\ &= \sigma^2 \operatorname{trace}(\mathbf{M})/(n-K) \\ &= \sigma^2\end{aligned}$$

which was what we wanted

Wrap-Up

strong assumptions $\implies$ *weak* results