War is too important to be left to the generals

Georges Clemenceau

While Theorem 5.1 shows that

\[
\lim_{n \to \infty} \sqrt{n} (\hat{\beta}_n - \beta_0) \overset{d}{\to} N(0, I)
\]

the results cannot be used in practice since \( \hat{\beta}_n \) is unknown.

To remedy this difficulty, we need a series of lemmas.

Lemma 5.2 (Hölder's inequality). Let \( p, q > 1 \) and \( \frac{1}{p} + \frac{1}{q} = 1 \). Then,

\[
\mathbb{E}|XY| \leq (\mathbb{E}|X|^p)^{\frac{1}{p}} (\mathbb{E}|Y|^q)^{\frac{1}{q}}.
\]

Remark 3. Let \( \exists \ i > 1 \), \( \exists i, \ldots, m \) for some positive integer \( m \). Then,

\[
\mathbb{E}|X_1 \cdots X_m| \leq (\mathbb{E}|X_1|^{\lambda_1})^{\frac{1}{\lambda_1}} \cdots (\mathbb{E}|X_m|^{\lambda_m})^{\frac{1}{\lambda_m}}.
\]
Lemma 5.3. Assume the same assumptions as in Theorem 5.1 except that (iv) and (vii) are strengthened to
\[
(iv') \sup_{x \geq 1} E |X_t|^\delta < C_1 < \infty , \quad \text{and}
\]
\[
(vii') \sup_{x \geq 1} \max_{s \leq k} E |X_{s-k}^t|^\delta < C_2 < \infty , \quad \text{respectively}.
\]

Then,
\[
\hat{V}_n - V_n \xrightarrow{p} 0 ,
\]

where
\[
\hat{V}_n = \frac{1}{n} \sum_{k=1}^{n} X_k X_k^\prime \hat{\varepsilon}_k^2
\]
with
\[
\hat{\varepsilon}_k = Y_k - X_k^\prime \hat{\gamma}_n .
\]

Proof. \[
(\hat{V}_n - V_n)_{ij} = \frac{1}{n} \sum_{k=1}^{n} (X_{ki} X_{kj} \hat{\varepsilon}_k^2 - E (X_{ki} X_{kj} \varepsilon_k^2))
\]
\[
\frac{1}{n} \sum_{k=1}^{n} \left[ \frac{1}{n} \sum_{k=1}^{n} \left( X_{k,i} X_{k,j} \right)^2 \right] \]

\[
= \frac{1}{n} \sum_{k=1}^{n} \left( X_{k,i} X_{k,j} \right)^2 - E \left( X_{k,i} X_{k,j} \right)^2 \\
= -2 \left( \hat{\beta}_n - \beta_0 \right) \frac{1}{n} \sum_{k=1}^{n} X_{k,i} X_{k,j} X_k \eta_k \\
+ \left( \hat{\beta}_n - \beta_0 \right) \left( \frac{1}{n} \sum_{k=1}^{n} X_{k,i} X_{k,j} X_k \eta_k \right) \left( \hat{\beta}_n - \beta_0 \right) = (I) + (II) + (III). \quad (5.12)
\]

Since for any \( \varepsilon > 0 \),

\[
P( |(I) | > \varepsilon ) \leq \frac{1}{n^2} \sum_{k=1}^{n} E \left( X_{k,i} X_{k,j} \right)^2 - E \left( X_{k,i} X_{k,j} \right)^2 \right)^2
\]

(why?)

\[
\left( \frac{x+y}{2} \right)^2 \leq 2x^2 + 2y^2 \leq \frac{2}{n^2} \sum_{k=1}^{n} \left[ E \left( X_{k,i} X_{k,j} \right)^2 \right] + \left( E \left( X_{k,i} X_{k,j} \right)^2 \right) \]

(Hölder's inequality)

\[
\leq \frac{2}{n^2} \sum_{k=1}^{n} \left[ \left( \frac{1}{n} \sum_{k=1}^{n} X_{k,i} \right)^{\frac{1}{2}} \left( \frac{1}{n} \sum_{k=1}^{n} X_{k,j} \right)^{\frac{1}{2}} \right] + \left( \frac{1}{n} \sum_{k=1}^{n} X_{k,i} \right)^{\frac{1}{2}} \left( \frac{1}{n} \sum_{k=1}^{n} X_{k,j} \right)^{\frac{1}{2}} \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \right)^{\frac{1}{2}}
\]

(iv) and

\[
O \left( \frac{1}{n} \right) = o(1).
\]
we have

\[(Z) = o_p(1) \quad (5.13)\]

By Hölder's inequality and (vi') and (vii'),

\[
\frac{1}{n} \sum_{k=1}^{n} x_{ki} x_{kj} x_{ki} = O_p(1) .
\]

This fact and the fact that

\[\hat{\beta}_n - \beta_0 = o_p(1) \quad \text{(See Theorem 3.1)}\]

imply

\[(\Pi) = o_p(1) \quad (5.15)\]

Similarly, it can be shown that

\[(\Pi) = o_p(1) \quad (5.15)\]

In view of (5.12) - (5.15), the claimed result follows.
Lemma 5.4. Assume the same assumptions as in Theorem 5.3.

Then

\[ \hat{P}_n - P_n = o_p(1), \quad (5.16) \]
\[ \hat{P}_n^{-1} - P_n^{-1} = o_p(1), \quad (5.17) \]

where

\[ \hat{P}_n = \hat{M}_n \hat{V}_n \hat{M}_n^{-1}. \]

proof. By an argument similar to that used in the proof of Lemma 3.3,

\[ \| \hat{M}_n^{-1} - M_n^{-1} \| = o_p(1). \quad (5.18) \]

Also observe that

\[ \hat{P}_n - P_n = \hat{M}_n \hat{V}_n \hat{M}_n^{-1} - M_n^{-1} V_n M_n \]
\[ = (\hat{M}_n^{-1} - M_n^{-1}) \hat{V}_n \hat{M}_n^{-1} + M_n^{-1} (\hat{V}_n - V_n) \hat{M}_n^{-1} + M_n^{-1} V_n (\hat{M}_n^{-1} - M_n^{-1}) \]
\[ = (I) + (II) + (III) \]
By (5.18) and Lemma 5.3,

\[(I) = (\hat{M}_n-M_n^{-1}) (\hat{V}_n-V_n) (\hat{M}_n-M_n^{-1})
+ (\hat{M}_n-M_n^{-1}) (\hat{V}_n-V_n) M_n^{-1}
+ (\hat{M}_n-M_n^{-1}) V_n (\hat{M}_n-M_n^{-1})
+ (\hat{M}_n-M_n^{-1}) V_n M_n^{-1} = o_p(1) o_p(1) o_p(1)
+ o_p(1) o_p(1) o(1) + o_p(1) o(1) o_p(1) + o_p(1) o(1) o(1)
= o_p(1).\]

Similarly, we have

\[(II) = o_p(1) \quad \text{and} \quad (III) = o_p(1),\]

and hence (5.16).

Finally, (5.17) can be shown by an argument similar to that used in Lemma 3.3.
Remark 5.

\[(5.17) \text{ implies } \]
\[
\frac{-\frac{1}{2}}{\hat{\beta}_n - \beta_0} = o_p(1) \quad (5.19)
\]

Theorem 5.5. Under the assumptions of Lemma 5.3,
\[
\sqrt{n} (\hat{\beta}_n - \beta_0) \xrightarrow{d} N(0, I_k) \quad (5.20)
\]
and
\[
\frac{-1}{\hat{\beta}_n - \beta_0} (\hat{\beta}_n - \beta_0) \xrightarrow{d} \chi^2(k) \quad (5.21)
\]

Remark 5. (Continuous mapping theorem).

If \(X_n \xrightarrow{d} X\) and \(g\) is a continuous function,
then \(g(X_n) \xrightarrow{d} g(X)\).
Proof of Theorem 5.5.

To show (5.20), observe that

$$
\| \hat{P}_n \sqrt{n} (\hat{\beta}_n - \beta) - \hat{\beta}_n \sqrt{n} (\beta - \beta_0) \|
$$

$$
\leq \| \hat{P}_n - P \| \| \sqrt{n} (\hat{\beta}_n - \beta_0) \|
$$

$$
\leq \| \hat{P}_n - P \|\| P \| \| \sqrt{n} (\hat{\beta}_n - \beta_0) \|
$$

(If $X_n \rightarrow X$, then $X_n = O_p(\cdot)$)

$$
= o_p(1) \circ O(1) \circ O_p(1) = o_p(1).
$$

(5.22)

By (5.22) and Slutsky's theorem,

$$
\hat{P}_n \sqrt{n} (\hat{\beta}_n - \beta_0) \quad \text{and} \quad \hat{\beta}_n \sqrt{n} (\beta - \beta_0)
$$

have the same limiting distribution, which is $N(0, I)$.

Finally, (5.21) is ensured by (5.20) and the continuous mapping theorem.
Theorem 5.6 (Wald test)

Let the assumptions of Lemma 5.3 hold. Consider the null hypothesis $H_0: R \beta_0 = \gamma$, where $R$ is a known matrix with non-singular rank $q \leq K$ and $\gamma$ is a known vector (scalar).

Then, under $H_0$,

$$W_n = n (R \hat{\beta}_n - \gamma) (R \hat{\beta}_n - \gamma)' (R \hat{\beta}_n - \gamma)^{-1} \rightarrow \chi^2(q),$$

where $\hat{\beta}_n = R \hat{\beta}_n R'$.

**Proof.** Observe that

$$\sqrt{n} (R \hat{\beta}_n - \gamma) = \sqrt{n} R (\hat{\beta}_n - \beta_0)$$

$$= R \hat{\beta}_n^{1/2} \sqrt{n} \hat{\beta}_n^{-1/2} (\hat{\beta}_n - \beta_0).$$
This fact, continuous mapping theorem, and Slutsky's theorem yield,

\[(R \hat{\beta}_n R')^{1/2} \sqrt{n} (R \hat{\beta}_n - R) \xrightarrow{d} N(0, I_y),\]

and hence

\[n (R \hat{\beta}_n - R) \hat{\beta}_n^{-1} (R \hat{\beta}_n - R) \xrightarrow{d} X^2(q).\]