Estimation of μ, γ_k, ρ_k and P_k

Ching-Kang Ing

Institute of Statistics, National Tsing Hua University, Hsinchu, Taiwan

Outline

1. Estimation of μ
2. Estimation of γ_k
3. Estimation of ρ_k
4. Estimation of P_k
5. CLT for $\bar{Z}$

Estimation of μ

Question 1

Having observed $Z_1, \ldots, Z_n$, how does one estimate $\mu = E(Z_t)$?

Ans: Using $\bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t$.

Question 2

Does $\bar{Z}$ possess desirable properties?

(i) Unbiasedness: $E(\bar{Z}) = \mu$ (easy to verify)

(ii) Consistency (i.e. for any $\epsilon > 0$, $P(|\bar{Z} - \mu| > \epsilon) \to 0$, as $n \to \infty$)

- By Chebyshev's inequality, we only need to show that

$$\begin{aligned}
\mathrm{Var}(\bar{Z}) &= E((\bar{Z}-\mu)^2) && \text{(by the unbiasedness of } \bar{Z}) \\
&= E(\bar{Z}^2) && \text{(WLOG, we may assume } \mu = 0) \\
&\overset{\text{why?}}{=} \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{i-j} \\
&\overset{\rho_j=\frac{\gamma_j}{\gamma_0}}{=} \frac{\gamma_0}{n}\sum_{t=-n+1}^{n-1}\left(1-\frac{|t|}{n}\right)\rho_t \to 0, \text{ as } n\to\infty, && (1)
\end{aligned}$$

which is guaranteed by

$$\rho_n \to 0 \text{ as } n \to \infty. \text{ (why?)} \qquad (2)$$

- In the following, we also call "consistency" as "convergence in probability", which is denoted by $\bar{Z} \xrightarrow{pr.} \mu$.

Question 3

Can you give a more precise expression for $\mathrm{Var}(\bar{Z})$?

- By (1), we have

$$n\mathrm{Var}(\bar{Z}) = \gamma_0 \sum_{t=-n+1}^{n-1} \rho_t - \frac{\gamma_0}{n} \sum_{t=-n+1}^{n-1} |t|\rho_t = \gamma_0 \sum_{t=-n+1}^{n-1} \rho_t - \frac{2\gamma_0}{n} \sum_{t=1}^{n-1} t\rho_t.$$

- If we assume

$$\sum_{t=-\infty}^{\infty} |\rho_t| < \infty, \tag{3}$$

then

$$\sum_{t=-n+1}^{n-1} \rho_t \xrightarrow{n\to\infty} \sum_{t=-\infty}^{\infty} \rho_t \text{ (why?)} \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n-1} t\rho_t \xrightarrow{n\to\infty} 0 \text{ (why?)}.$$

As a result,

$$\lim_{n\to\infty} n\mathrm{Var}(\bar{Z}) = \gamma_0 \sum_{t=-\infty}^{\infty} \rho_t. \tag{4}$$

Estimation of γ_k

Question 4

How do we estimate γ_k?

Ans: Using $\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z})$.

Question 5

Does $\hat{\gamma}_k \xrightarrow{pr.} \gamma_k$ hold? (A big question!!)

We will analyze this question through several steps.

Step 1

- Define $\tilde{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \mu)(Z_{t+k} - \mu)$. We will first show that

$$\hat{\gamma}_k - \tilde{\gamma}_k \xrightarrow{pr.} 0.$$

- To see this, we express $\hat{\gamma}_k$ as

$$\begin{aligned} &\frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \mu + \mu - \bar{Z})(Z_{t+k} - \mu + \mu - \bar{Z}) \\ &= \frac{\mu - \bar{Z}}{n}\sum_{t=1}^{n-k}(Z_{t+k} - \mu) + \frac{\mu - \bar{Z}}{n}\sum_{t=1}^{n-k}(Z_t - \mu) + \frac{(n-k)(\mu - \bar{Z})^2}{n} + \tilde{\gamma}_k. \end{aligned} \tag{5}$$

- By an argument similar to that used to prove $\bar{Z} \xrightarrow{pr.} \mu$, we have

$$\frac{1}{n}\sum_{t=1}^{n-k}(Z_{t+k} - \mu) \xrightarrow{pr.} 0 \quad \text{and} \quad \frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \mu) \xrightarrow{pr.} 0.$$

Consequently, $\hat{\gamma}_k - \tilde{\gamma}_k \xrightarrow{pr.} 0$ follows.

Step 2

- We will show that $\tilde{\gamma}_k \xrightarrow{pr.} \gamma_k$. The proof of this result is more involved.
- We can again assume $\mu = 0$. Then,

$$\tilde{\gamma}_k - \gamma_k = \frac{1}{n} \sum_{t=1}^{n-k} (Z_t Z_{t+k} - \gamma_k) - \frac{k\gamma_k}{n}.$$

- Since $\frac{k\gamma_k}{n} \to 0$, we only need to show that

$$\frac{1}{n-k} \sum_{t=1}^{n-k} (Z_t Z_{t+k} - \gamma_k) \xrightarrow{pr.} 0. \tag{6}$$

- Define

$$y_{t,k} = Z_t Z_{t+k} - \gamma_k \quad \text{and} \quad V_{i,k} = E(y_{t,k} y_{t+i,k}),$$

noting that $y_{t,k}$ is second-order weakly stationary (covariance stationary) if we assume $\{Z_t\}$ is "fourth-order weakly stationary".

Recall

$\{Z_t\}$ is called qth-order, $q \geq 2$, weakly stationary if $E(Z_t) = \mu$ for all t, and for any $t_1, \ldots, t_q$ and k,

$$E[(Z_{t_1} - \mu) \cdots (Z_{t_q} - \mu)] = E[(Z_{t_1+k} - \mu) \cdots (Z_{t_q+k} - \mu)].$$

- In fact, it can shown that

$$V_{i,k} = V_{-i,k} \text{ for } i = 0, 1, 2, \ldots, \quad \text{(why?)} \tag{7}$$

if $\{Z_t\}$ is fourth-order stationary, which will be assumed in the rest of the proof of (6).

- Define $N = n - k$. Then,

$$E\left(\frac{1}{n-k}\sum_{t=1}^{n-k}(Z_t Z_{t+k} - \gamma_k)\right)^2$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} V_{i-j,k} \qquad \text{(Since } y_{t,k} \text{ is covariance stationary)}$$

$$= \frac{1}{N}\sum_{t=-N+1}^{N-1}\left(1-\frac{|t|}{N}\right)V_{t,k} \to 0, \text{ as } n \to \infty, \qquad (8)$$

provided

$$V_{n,k} \to 0, \text{ as } n \to \infty. \text{ (why?)} \qquad (9)$$

- By (9) and Chebyshev's inequality, we obtain (6).

Step 3

Since $\hat{\gamma}_k - \tilde{\gamma}_k \xrightarrow{pr.} 0$ and $\tilde{\gamma}_k - \gamma_k \xrightarrow{pr.} 0$, it follows that

$$\hat{\gamma}_k - \gamma_k \xrightarrow{pr.} 0. \quad \text{(or, equivalently, } \hat{\gamma}_k \xrightarrow{pr.} \gamma_k\text{)}$$

Remark

When $\{Z_t\}$ is a Gaussian process, we obtain by Isserlis's Theorem that

$$\begin{aligned} V_{n,k} &= E(Z_t Z_{t+k} Z_{t+n} Z_{t+n+k}) - \gamma_k^2 \\ &= E(Z_t Z_{t+k}) E(Z_{t+n} Z_{t+n+k}) + E(Z_t Z_{t+n}) E(Z_{t+k} Z_{t+n+k}) \\ &\quad + E(Z_t Z_{t+n+k}) E(Z_{t+k} Z_{t+n}) - \gamma_k^2 \\ &= \gamma_n^2 + \gamma_{n+k}\gamma_{n-k}, \end{aligned}$$

and hence a sufficient condition for (9) to hold (under Gaussianity) is

$$\gamma_n \to 0 \text{ as } n \to \infty. \quad \text{(why?)} \quad \text{(This condition is equivalent to (2).)} \qquad (10)$$

Question 6

Can you give a more precise expression for the mean squared error $E(\hat{\gamma}_k - \gamma_k)^2$?

- Note that since $E(\hat{\gamma}_k) \neq \gamma_k$, $E(\hat{\gamma}_k - \gamma_k)^2 \neq \mathrm{Var}(\hat{\gamma}_k)$. However, we should also note that the difference between $E(\hat{\gamma}_k - \gamma_k)^2$ and $\mathrm{Var}(\hat{\gamma}_k)$ is "small".
- In the following, I will show that

$$\text{Bartlett's formula}: \quad E\{n(\hat{\gamma}_k - \gamma_k)^2\} \xrightarrow{n\to\infty} \sum_{t=-\infty}^{\infty} (\gamma_t^2 + \gamma_{t+k}\gamma_{t-k}), \qquad (11)$$

provided $\{Z_t\}$ is Gaussian.

- Note first that

$$\begin{aligned} nE(\hat{\gamma}_k - \gamma_k)^2 &= nE(\hat{\gamma}_k - \tilde{\gamma}_k + \tilde{\gamma}_k - \gamma_k)^2 \\ &= n\{E(\hat{\gamma}_k - \tilde{\gamma}_k)^2 + E(\tilde{\gamma}_k - \gamma_k)^2 + 2E[(\hat{\gamma}_k - \tilde{\gamma}_k)(\tilde{\gamma}_k - \gamma_k)]\}. \end{aligned}$$

- By (3) and (5), it is not difficult to show that

$$nE(\hat{\gamma}_k - \tilde{\gamma}_k)^2 \to 0 \text{ as } n \to \infty. \quad (12)$$

- If we can prove (Here, we have implicitly assumed that $\sum_{i=\infty}^{\infty} \gamma_i^2 < \infty$.)

$$\lim_{n\to\infty} nE(\tilde{\gamma}_k - \gamma_k)^2 = \sum_{t=-\infty}^{\infty} (\gamma_t^2 + \gamma_{t+k}\gamma_{t-k}), \quad (13)$$

  then (11) follows from (12), (13) and the Cauchy-Schwarz inequality. (why?)
- To show (13), note first that

$$\begin{aligned} E[n(\tilde{\gamma}_k - \gamma_k)^2] &= \frac{1}{n} E\left(\sum_{t=1}^{N} y_{t,k}\right)^2 - 2k\gamma_k E\left(\frac{1}{n}\sum_{t=1}^{n-k} y_{t,k}\right) + \frac{k^2\gamma_k^2}{n} \\ &= \frac{N}{n} E\left(\frac{1}{N}\left(\sum_{t=1}^{N} y_{t,k}\right)^2\right) + \frac{k^2\gamma_k^2}{n}. \quad (14) \end{aligned}$$

- By (7) and (8), it follows that

$$\frac{1}{N}E\left(\sum_{t=1}^{N} y_{t,k}\right)^2 = \sum_{t=-N+1}^{N-1} V_{t,k} - \frac{2}{N}\sum_{t=1}^{N-1} tV_{t,k} \xrightarrow{n\to\infty} \sum_{t=-\infty}^{\infty} V_{t,k}, \qquad (15)$$

provided

$$\sum_{t=-\infty}^{\infty} |V_{t,k}| < \infty. \qquad (16)$$

- By Isserlis's Theorem,

$$\sum_{t=-\infty}^{\infty} V_{t,k} \overset{\text{why?}}{=} \sum_{t=-\infty}^{\infty} (\gamma_t^2 + \gamma_{t+k}\gamma_{t-k}). \qquad (17)$$

- Moreover, (16) is ensured by

$$\sum_{t=-\infty}^{\infty} \gamma_t^2 < \infty. \quad \text{(why?)} \tag{18}$$

- Consequently, (13) follows from (14), (15), (17), provided (18) holds true.

Remark

In fact, it can be shown that

$$\lim_{n\to\infty} n\mathrm{Var}(\hat{\gamma}_k) = \sum_{t=-\infty}^{\infty} (\gamma_t^2 + \gamma_{t+k}\gamma_{t-k}),$$

under the same assumptions.

Estimation of ρ_k

Question 7

How to estimate ρ_k?

Ans: Using $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$.

Question 8

Does $\hat{\rho}_k \xrightarrow{pr.} \rho_k$?

Ans: Yes. (since $\hat{\gamma}_k \xrightarrow{pr.} \gamma_k$ and $\hat{\gamma}_0 \xrightarrow{pr.} \gamma_0$)

Question 9

How to derive Bartlett's formula for $\lim_{n\to\infty} n\text{Var}(\hat{\rho}_k)$?

- Note first that

$$\lim_{n\to\infty} n\text{Cov}(\hat{\gamma}_k, \hat{\gamma}_{k+j}) = \sum_{t=-\infty}^{\infty} (\gamma_t\gamma_{t+j} + \gamma_{t+j+k}\gamma_{k-t}),$$

whose proof is sketched as follows.

- Straightforward calculations yield

$$\begin{aligned} & n\text{Cov}(\hat{\gamma}_k, \hat{\gamma}_{k+j}) \\ &= nE[(\hat{\gamma}_k - \gamma_k)(\hat{\gamma}_{k+j} - \gamma_{k+j})] - nE(\hat{\gamma}_k - \gamma_k)E(\hat{\gamma}_{k+j} - \gamma_{k+j}) \\ &= nE[(\hat{\gamma}_k - \gamma_k)(\hat{\gamma}_{k+j} - \gamma_{k+j})] + o(1). \end{aligned}$$

Moreover,

$$\begin{aligned} nE[(\hat{\gamma}_k - \gamma_k)(\hat{\gamma}_{k+j} - \gamma_{k+j})] = &\, nE[(\hat{\gamma}_k - \tilde{\gamma}_k)(\hat{\gamma}_{k+j} - \tilde{\gamma}_{k+j})]^{①} \\ &+ nE[(\hat{\gamma}_k - \tilde{\gamma}_k)(\tilde{\gamma}_{k+j} - \gamma_{k+j})]^{②} \\ &+ nE[(\tilde{\gamma}_k - \gamma_k)(\hat{\gamma}_{k+j} - \tilde{\gamma}_{k+j})]^{③} \\ &+ nE[(\tilde{\gamma}_k - \gamma_k)(\tilde{\gamma}_{k+j} - \gamma_{k+j})]^{④} \end{aligned}$$

①: $nE[(\hat{\gamma}_k - \tilde{\gamma}_k)(\hat{\gamma}_{k+j} - \tilde{\gamma}_{k+j})] \xrightarrow{\text{why?}} 0$

②: by the Cauchy-Schwarz Inequality,

$$nE[(\hat{\gamma}_k - \tilde{\gamma}_k)(\tilde{\gamma}_{k+j} - \gamma_{k+j})] \le n\sqrt{E(\hat{\gamma}_k - \tilde{\gamma}_k)^2 E(\tilde{\gamma}_{k+j} - \gamma_{k+j})^2} \to 0$$

③: $\to 0$ (similar to ②)

④:

$$
\begin{aligned}
nE[(\tilde{\gamma}_k - \gamma_k)(\tilde{\gamma}_{k+j} - \gamma_{k+j})] &= \frac{1}{n} E\left[\sum_{t=1}^{n-k} \sum_{i=1}^{n-k-j} (Z_t Z_{t+k} - \gamma_k)(Z_i Z_{i+k+j} - \gamma_{k+j})\right] \\
&\simeq \frac{1}{n} E\left[\sum_{t=1}^{n} \sum_{i=1}^{n} (Z_t Z_{t+k} - \gamma_k)(Z_i Z_{i+k+j} - \gamma_{k+j})\right] \\
&= \frac{1}{n} \sum_{t=-n+1}^{n-1} (n - |t|)\eta_t,
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_t &= E[(Z_i Z_{i+k} - \gamma_k)(Z_{i+t} Z_{i+t+k+j} - \gamma_{k+j})] \\
&= E(Z_i Z_{i+k} Z_{i+t} Z_{i+t+k+j}) - \gamma_k \gamma_{k+j} \\
&= \gamma_t \gamma_{t+j} + \gamma_{t+j+k} \gamma_{k-t}, \qquad \text{(assuming Gaussianity)}
\end{aligned}
$$

implying that

$$
\begin{aligned}
\lim_{n\to\infty} nE[(\tilde{\gamma}_k - \gamma_k)(\tilde{\gamma}_{k+j} - \gamma_{k+j})] &= \lim_{n\to\infty} \sum_{t=-n+1}^{n-1} \eta_t - \frac{1}{n} \sum_{t=-n+1}^{n-1} |t|\eta_t \\
&= \sum_{t=-\infty}^{\infty} \gamma_t \gamma_{t+j} + \gamma_{t+j+k} \gamma_{k-t}.
\end{aligned}
$$

- Thus, we have

$$\lim_{n\to\infty} n\mathrm{Cov}(\hat{\gamma}_k, \hat{\gamma}_{k+j}) = \sum_{t=-\infty}^{\infty} (\gamma_t\gamma_{t+j} + \gamma_{t+j+k}\gamma_{k-t}) \stackrel{\text{def}}{\equiv} u_{k,k+j},$$

which is Bartlett's formula for the covariance of $\hat{\gamma}_k$ and $\hat{\gamma}_{k+j}$.

- Now, write $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = f(\hat{\gamma}_0, \hat{\gamma}_k)$. By Taylor's expansion, one has

$$\frac{\hat{\gamma}_k}{\hat{\gamma}_0} \sim \frac{\gamma_k}{\gamma_0} + \left(\frac{\partial f(\gamma_0, \gamma_k)}{\partial \gamma_0}, \frac{\partial f(\gamma_0, \gamma_k)}{\partial \gamma_k}\right) \begin{pmatrix} \hat{\gamma}_0 - \gamma_0 \\ \hat{\gamma}_k - \gamma_k \end{pmatrix}.$$

- Therefore,

$$\sqrt{n}(\hat{\rho}_k - \rho_k) \sim \frac{1}{\gamma_0}(-\rho_k, 1)\sqrt{n} \begin{pmatrix} \hat{\gamma}_0 - \gamma_0 \\ \hat{\gamma}_k - \gamma_k \end{pmatrix}$$

yielding

$$n\mathrm{Var}(\hat{\rho}_k) \sim \frac{1}{\gamma_0^2}(-\rho_k, 1) \begin{pmatrix} u_{00} & u_{0k} \\ u_{k0} & u_{kk} \end{pmatrix} \begin{pmatrix} -\rho_k \\ 1 \end{pmatrix}.$$

- Straightforward algebraic manipulations yield

$$\lim_{n\to\infty} n\mathrm{Var}(\hat{\rho}_k) = \sum_{t=-\infty}^{\infty} (\rho_t^2 + \rho_{t+k}\rho_{t-k} - 4\rho_k\rho_t\rho_{t+k} + 2\rho_k^2\rho_t^2).$$

- In fact, it can be shown that

$$\sqrt{n}\begin{pmatrix} \hat{\gamma}_0 - \gamma_0 \\ \vdots \\ \hat{\gamma}_k - \gamma_k \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} u_{00} & \cdots & u_{0k} \\ \vdots & \ddots & \vdots \\ u_{k0} & \cdots & u_{kk} \end{pmatrix}\right).$$

This and Taylor's expansion give

$$\sqrt{n}(\hat{\rho}_k - \rho_k) \xrightarrow{d} N\left(0, \sum_{t=-\infty}^{\infty} (\rho_t^2 + \rho_{t+k}\rho_{t-k} - 4\rho_k\rho_t\rho_{t+k} + 2\rho_k^2\rho_t^2)\right).$$

Question 10

How to construct a test of the null hypothesis:

$$\mathrm{H}_0 : \rho_1 = \rho_2 = \cdots = 0 \quad \text{vs} \quad \mathrm{H}_1 : \sim \mathrm{H}_0$$

at the asymptotic 0.05 significance level?

By Bartlett's formula for $\hat{\rho}_k$, $k \geq 1$,

$$\sqrt{n}\hat{\rho}_k \xrightarrow{d} N(0,1),$$

provided H_0 holds true. Therefore, the testing rule:

$$\text{Reject } \mathrm{H}_0 \text{ if } |\hat{\rho}_k| > \frac{1.96}{\sqrt{n}}$$

has an asymptotic 0.05 significance level, i.e.:

$$P_{\mathrm{H}_0}\left(|\hat{\rho}_k| > \frac{1.96}{\sqrt{n}}\right) \to 0.05 \quad \text{as} \quad n \to \infty.$$

Question 11

How to construct a test of the null hypothesis:

$$\mathrm{H}_0 : \rho_k = \rho_{k+1} = \cdots = 0 \quad \text{vs} \quad \mathrm{H}_1 : \sim \mathrm{H}_0$$

at the asymptotic 0.05 significance level?

- By Bartlett's formula for $\hat{\rho}_k$, $k \geq 1$,

$$\sqrt{n}\hat{\rho}_k \xrightarrow{d} N\left(0, \sum_{i=-k+1}^{k-1} \rho_i^2\right),$$

  provided H_0 holds true.
- Moreover, since $\hat{\rho}_i \xrightarrow{pr.} \rho_i$ for all i, the testing rule:

$$\text{Reject } \mathrm{H}_0 \text{ if } |\hat{\rho}_k| > 1.96\sqrt{\frac{\sum_{i=-k+1}^{k-1} \hat{\rho}_i^2}{n}}$$

  has an asymptotic 0.05 significant level.

Estimation of P_k

Question 12

How to estimate P_k?

- Consider the following alignment of data:

$$\boldsymbol{y} = \begin{pmatrix} Z_{k+1} - \bar{Z} \\ \vdots \\ Z_n - \bar{Z} \end{pmatrix}, \boldsymbol{X} = \begin{pmatrix} Z_k - \bar{Z} & \cdots & Z_2 - \bar{Z} \\ \vdots & & \vdots \\ Z_{n-1} - \bar{Z} & \cdots & Z_{n-k+1} - \bar{Z} \end{pmatrix}, \text{ and } \boldsymbol{w} = \begin{pmatrix} Z_1 - \bar{Z} \\ \vdots \\ Z_{n-k} - \bar{Z} \end{pmatrix}.$$

- Define $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{w}$, $\hat{\boldsymbol{\eta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, $\hat{\boldsymbol{w}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\eta}}$.
- We have two estimators,

$$\hat{P}_k^{(1)} = \frac{(\boldsymbol{w} - \hat{\boldsymbol{w}})'(\boldsymbol{y} - \hat{\boldsymbol{y}})}{\|\boldsymbol{w} - \hat{\boldsymbol{w}}\| \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|} \quad \text{and} \quad \hat{P}_k^{(2)} = \frac{(\boldsymbol{w} - \hat{\boldsymbol{w}})'(\boldsymbol{y} - \hat{\boldsymbol{y}})}{\|\boldsymbol{w} - \hat{\boldsymbol{w}}\|^2},$$

where $\|\cdot\|$ denotes the Euclidean norm.

Question 13

Does $\hat{P}_k^{(i)} \xrightarrow{pr.} P_k$? $(i = 1, 2)$

Yes! This is sketched as follows.

(i) $\frac{1}{n}(\boldsymbol{w} - \hat{\boldsymbol{w}})'(\boldsymbol{y} - \hat{\boldsymbol{y}}) \xrightarrow{pr.} \mathrm{Cov}[(Z_t - \tilde{Z}_t), (Z_{t+k} - \tilde{Z}_{t+k})]$

(ii) $\frac{1}{n}\|\boldsymbol{w} - \hat{\boldsymbol{w}}\|^2 \xrightarrow{pr.} \mathrm{Var}(Z_t - \tilde{Z}_t)$

(iii) $\frac{1}{n}\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2 \xrightarrow{pr.} \mathrm{Var}(Z_{t+k} - \tilde{Z}_{t+k})$

- WLOG, assume $\bar{Z} = 0$. Then,

$$\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\eta}} = \boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{M})\boldsymbol{y},$$

  and

$$\boldsymbol{w} - \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{w} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{w} = (\boldsymbol{I} - \boldsymbol{M})\boldsymbol{w}.$$

- It is easy to see that $\boldsymbol{M}^2 = \boldsymbol{M}$ and $\boldsymbol{M}' = \boldsymbol{M}$.

Hence,

$$\begin{aligned}
\frac{1}{n}(\boldsymbol{w}-\hat{\boldsymbol{w}})'(\boldsymbol{y}-\hat{\boldsymbol{y}}) &= \frac{1}{n}[\boldsymbol{w}'\boldsymbol{y}-\boldsymbol{w}'\boldsymbol{M}\boldsymbol{y}] \\
&\simeq \hat{\gamma}_k - n^{-1}\boldsymbol{w}'\boldsymbol{X}(n^{-1}\boldsymbol{X}'\boldsymbol{X})^{-1}n^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&\simeq \hat{\gamma}_k - (\hat{\gamma}_{k-1},\ldots,\hat{\gamma}_1)\hat{\boldsymbol{R}}_{k-1}^{-1}\begin{pmatrix}\hat{\gamma}_1\\ \vdots \\ \hat{\gamma}_{k-1}\end{pmatrix} \\
&\xrightarrow{pr.} \gamma_k - (\gamma_{k-1},\ldots,\gamma_1)\boldsymbol{R}_{k-1}^{-1}\begin{pmatrix}\gamma_1\\ \vdots \\ \gamma_{k-1}\end{pmatrix} \\
&= \gamma_k - (\gamma_{k-1},\ldots,\gamma_1)\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_{k-1}\end{pmatrix} \\
&\overset{\text{why?}}{=} \mathrm{Cov}[(Z_t-\tilde{Z}_t),(Z_{t+k}-\tilde{Z}_{t+k})],
\end{aligned}$$

where $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_{k-1})'$ is the minimizer of $E(Z_{t+k}-\tilde{Z}_{t+k})^2$.

- Similarly,

$$\begin{aligned}\frac{1}{n}\|\boldsymbol{w}-\hat{\boldsymbol{w}}\|^2 &= \frac{1}{n}\boldsymbol{w}'(\boldsymbol{I}-\boldsymbol{M})\boldsymbol{w}\\ &= n^{-1}\boldsymbol{w}'\boldsymbol{w}-n^{-1}\boldsymbol{w}'\boldsymbol{X}(n^{-1}\boldsymbol{X}'\boldsymbol{X})^{-1}n^{-1}\boldsymbol{X}'\boldsymbol{w}\\ &\xrightarrow{pr.} \gamma_0-\boldsymbol{\gamma}_{k-1}'\boldsymbol{R}_{k-1}^{-1}\boldsymbol{\gamma}_{k-1}\\ &= \mathrm{Var}(Z_t-\tilde{Z}_t)=C,\end{aligned}$$

and

$$\frac{1}{n}\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2=\mathrm{Var}(Z_{t+k}-\tilde{Z}_{t+k})=C.$$

- Combining these facts yields that

$$\hat{P}_k^{(1)}=\frac{(\boldsymbol{w}-\hat{\boldsymbol{w}})'(\boldsymbol{y}-\hat{\boldsymbol{y}})}{\|\boldsymbol{w}-\hat{\boldsymbol{w}}\|\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|}\xrightarrow{pr.}\frac{\mathrm{Cov}(Z_{t+k}-\tilde{Z}_{t+k},Z_t-\tilde{Z}_t)}{\sqrt{\mathrm{Var}(Z_{t+k}-\tilde{Z}_{t+k})}\sqrt{\mathrm{Var}(Z_t-\tilde{Z}_t)}}=P_k.$$

- Recall that ϕ_k minimizes $E(Z_{t+k} - \boldsymbol{c}'\boldsymbol{Z}_{t+k-1}(k))^2$ over $\boldsymbol{c} \in \mathbb{R}^k$, where $\boldsymbol{Z}_{t+k-1}(k) = (Z_{t+k-1}, \ldots, Z_t)'$. Therefore, a natural estimate of $\boldsymbol{\phi}_k$ is

$$\hat{\boldsymbol{\phi}}_k = \begin{pmatrix} \hat{\phi}_{k1} \\ \vdots \\ \hat{\phi}_{kk} \end{pmatrix} = \hat{\boldsymbol{R}}_k^{-1} \begin{pmatrix} \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_k \end{pmatrix}.$$

- It can be shown that

$$\hat{\phi}_{kk} = \frac{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{w}}{\boldsymbol{w}'(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{w}} = \hat{P}_k^{(2)}.$$

- Moreover, since $n^{-1}\|\boldsymbol{w} - \hat{\boldsymbol{w}}\|^2 \xrightarrow{pr.} C$ and $n^{-1}\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2 \xrightarrow{pr.} C$,

$$\hat{P}_k^{(2)} - \hat{P}_k^{(1)} \xrightarrow{pr.} 0.$$

Therefore, $\hat{P}_k^{(2)}$ is also a consistent estimate of P_k.

Question 14

How to construct a test of the null hypothesis

$$\mathrm{H}_0 : P_1 = P_2 = \cdots = 0 \quad \text{vs} \quad \mathrm{H}_1 : \sim \mathrm{H}_0$$

at the asymptotic 0.05 significance level?

We will show later in this semester that

$$\sqrt{n}(\hat{\boldsymbol{\phi}}_k - \boldsymbol{\phi}_k) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{I}),$$

provided H_0 holds, yielding

$$\sqrt{n}\hat{P}_k^{(2)} \xrightarrow{d} N(0, 1),\ k \geq 1.$$

Therefore, the testing rule:

$$\text{Reject } \mathrm{H}_0 \text{ if } |\hat{P}_k^{(2)}| > \frac{1.96}{\sqrt{n}}$$

has an asymptotic 0.05 significant level.

CLT for $\bar{Z}$

Theorem

If $\{Z_t\}$ is a stationary process satisfying

$$Z_t = \mu + \sum_{j=0}^{\infty} b_j \epsilon_{t-j},$$

where $\epsilon_i \overset{\text{indep.}}{\sim} (0, \sigma^2)$. Moreover, assume that $\{\epsilon_t\}$ obeys Lindeberg's condition and

$$\sum_{j=1}^{\infty} j^{1/2} |b_j| < \infty.$$

Then, $\sqrt{n}(\bar{Z} - \mu) \xrightarrow{d} \mathcal{N}(0, \sum_{j=-\infty}^{\infty} \gamma_j)$.

Proof of Theorem

- Without loss of generality, we assume that $\mu = 0$. Then

$$Z_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j} = \sum_{\ell=-\infty}^{t} b_{t-\ell} \epsilon_\ell,$$

and

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{\ell=-\infty}^{i} b_{i-\ell} \epsilon_\ell \\
&= \frac{1}{\sqrt{n}} \sum_{\ell=-\infty}^{n} \sum_{i=\ell \vee 1}^{n} b_{i-\ell} \epsilon_\ell \\
&= \frac{1}{\sqrt{n}} \sum_{\ell=-\infty}^{0} \left(\sum_{i=1}^{n} b_{i-\ell} \right) \epsilon_\ell + \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \left(\sum_{i=\ell}^{n} b_{i-\ell} \right) \epsilon_\ell := (\mathrm{I}) + (\mathrm{II}).
\end{aligned}$$

- Note that $\sum_{j=1}^{\infty} j^{1/2} |b_j| < \infty$ implies $n^{1/2} \sum_{i=n}^{\infty} |b_i| \to 0$ as $n \to \infty$.

Proof of Theorem

Since for any $\epsilon > 0$,

$$\begin{aligned}
\mathrm{Var}((\mathrm{I})) = \frac{\sigma^2}{n} \sum_{\ell=-\infty}^{0} \left(\sum_{i=1}^{n} b_{i-\ell} \right)^2 &\leq \frac{\sigma^2}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{\ell=-\infty}^{0} |b_{i-\ell}||b_{j-\ell}| \\
&\leq \frac{\sigma^2}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{\ell=-\infty}^{0} b_{i-\ell}^2 \right)^{1/2} \left(\sum_{\ell=-\infty}^{0} b_{j-\ell}^2 \right)^{1/2} \\
&\leq \frac{\sigma^2}{n} \left(\sum_{i=1}^{n} \sum_{\ell=-\infty}^{0} |b_{i-\ell}| \right)^2 = \sigma^2 \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{k=i}^{\infty} |b_k| \right)^2 \\
&\leq \sigma^2 \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\epsilon n} i|b_i| + \frac{1}{\sqrt{n}} \sum_{i=\epsilon n+1}^{n} i|b_i| + \sqrt{n} \sum_{i=n+1}^{\infty} |b_i| \right)^2 \\
&\leq \sigma^2 \left(\sqrt{\epsilon} \sum_{i=1}^{\epsilon n} i^{1/2}|b_i| + \sum_{i=\epsilon n+1}^{n} i^{1/2}|b_i| + o(1) \right)^2,
\end{aligned}$$

it holds that $\mathrm{Var}((\mathrm{I})) \underset{n\to\infty}{\longrightarrow} 0$, and hence $(\mathrm{I}) = o_{\mathrm{p}}(1)$ follows from Chebyshev's inequality.

Proof of Theorem

- Note that

$$\begin{aligned}(\text{II}) &= \frac{1}{\sqrt{n}}[\epsilon_n(b_0) + \epsilon_{n-1}(b_0+b_1) + \cdots + \epsilon_1(b_0 + \cdots + b_{n-1})] \\ &= \frac{S}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_i - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_i R_{in} := (\text{III}) + (\text{IV}),\end{aligned}$$

  where $S = \sum_{i=0}^{\infty} b_i$ and $R_{in} = \sum_{j=n-i+1}^{\infty} b_j$.

- Since $(\text{III}) \xrightarrow{d} \mathcal{N}(0, S^2\sigma^2)$ and

$$\begin{aligned}\text{Var}((\text{IV})) \le \frac{\sigma^2}{n}\sum_{i=1}^{n}\left(\sum_{j=n-i+1}^{\infty}|b_j|\right)^2 &= \frac{\sigma^2}{n}\sum_{i=1}^{n}\left(\sum_{j=i}^{\infty}|b_j|\right)^2 \\ &= \frac{\sigma^2}{n}\sum_{i=1}^{n}\left(i^{-1/2}i^{1/2}\sum_{j=i}^{\infty}|b_j|\right)^2 \\ &\le \frac{C}{n}\sum_{i=1}^{n} i^{-1} \underset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

  where C is a constant, the proof is completed by $S^2\sigma^2 = \sum_{j=-\infty}^{\infty}\gamma_j$ and Slutsky's theorem.

Remark

If $\sum_{j=1}^{\infty} j^{1/2}|b_j| < \infty$ is replaced with $\sum_{j=1}^{\infty} |b_j| < \infty$, then the above theorem still holds.

- To see this, we recall that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i = \frac{1}{\sqrt{n}} \sum_{\ell=-\infty}^{0} \left(\sum_{i=1}^{n} b_{i-\ell} \right) \epsilon_\ell + \frac{S}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i R_{in} := (\mathrm{I}) + (\mathrm{III}) + (\mathrm{IV}),$$

  where $S = \sum_{i=0}^{\infty} b_i$, $R_{in} = \sum_{j=n-i+1}^{\infty} b_j$, and $(\mathrm{III}) \xrightarrow{d} \mathcal{N}(0, \sum_{j=-\infty}^{\infty} \gamma_j)$.

- Now it suffices to show that $\mathrm{Var}((\mathrm{I})) = o(1)$ and $\mathrm{Var}((\mathrm{IV})) = o(1)$.
- Since $\sum_{j=n}^{\infty} |b_j| \to 0$ as $n \to \infty$, we have

$$\mathrm{Var}((\mathrm{I})) \leq \frac{\sigma^2}{n} \sum_{i=1}^{n} \sum_{\ell=0}^{\infty} b_{i+\ell}^2 + \frac{2\sigma^2}{n} \sum_{i=1}^{n} \sum_{j=i+1}^{\infty} \sum_{\ell=0}^{\infty} |b_{i+\ell}||b_{j+\ell}| = \frac{\sigma^2}{n} \sum_{i=1}^{n} \left(\sum_{j=i}^{\infty} |b_j| \right)^2 = o(1).$$

  Similarly, $\mathrm{Var}((\mathrm{IV})) \leq \sigma^2 n^{-1} \sum_{i=1}^{n} \left(\sum_{j=i}^{\infty} |b_j| \right)^2 = o(1)$.

Now I will show that

$$\sum_{j=-\ell_n}^{\ell_n} \hat{\gamma}_j \xrightarrow{pr.} \sum_{j=-\infty}^{\infty} \gamma_j,$$

as $\ell_n \to \infty$ and $\ell_n = o(n^{1/2})$. Note first that

$$\left| \sum_{j=-\ell_n}^{\ell_n} \hat{\gamma}_j - \sum_{j=-\infty}^{\infty} \gamma_j \right| \leq \sum_{j=-\ell_n}^{\ell_n} |\hat{\gamma}_j - \gamma_j| + 2 \sum_{j=\ell_n+1}^{\infty} |\gamma_j| := \text{(V)} + \text{(VI)}.$$

Moreover, since $E[(\sqrt{n}(\hat{\gamma}_j - \gamma_j))^2] \leq C$ where C is some constant, we have

$$E(\text{(V)}) = \sum_{j=-\ell_n}^{\ell_n} E|\hat{\gamma}_j - \gamma_j| \leq \sum_{j=-\ell_n}^{\ell_n} [E((\hat{\gamma}_j - \gamma_j)^2)]^{1/2} \leq C n^{-1/2}(2\ell_n + 1) \underset{n\to\infty}{\longrightarrow} 0,$$

and thus $\text{(V)} = o_p(1)$ is obtained by Markov's inequality. Furthermore,

$$\text{(VI)} \leq 2\sigma^2 \sum_{j=\ell_n+1}^{\infty} \sum_{s=0}^{\infty} |b_s||b_{s+j}| = 2\sigma^2 \sum_{s=0}^{\infty} |b_s| \sum_{j=\ell_n+1}^{\infty} |b_{s+j}| \leq 2\sigma^2 \sum_{s=0}^{\infty} |b_s| \sum_{j=\ell_n+1}^{\infty} |b_j| \underset{n\to\infty}{\longrightarrow} 0,$$

provided $\sum_{j=0}^{\infty} j^{1/2}|b_j| < \infty$. Hence

$$\bar{Z} \pm \frac{z_{1-\alpha/2}\sqrt{\sum_{j=-\ell_n}^{\ell_n} \hat{\gamma}_j}}{\sqrt{n}}$$

is an asymptotic $100(1-\alpha)\%$ confidence interval for μ.