6 Unit Root (単位根) and Cointegration (共和分)

6.1 Unit Root (単位根) Test (Dickey-Fuller (DF) Test)

1. Why is a unit root problem important?

   (a) Economic variables increase over time in general.

   One of the assumptions of OLS is stationarity on y_t and x_t.

   This assumption implies that $\frac{1}{T}X'X$ converges to a fixed matrix as T is large.

   That is, asymptotic normality of OLS estimator does not hold.

   (b) In nonstationary time series, the unit root is the most important.

   In the case of unit root, OLSE of the first-order autoregressive coefficient is consistent.

OLSE is $\sqrt{T}$-consistent in the case of stationary AR(1) process, but OLSE is T-consistent in the case of nonstationay AR(1) process.

(c) A lot of economic variables increase over time.

It is important to check an economic variable is trend stationary (i.e., $y_t = a_0 + a_1 t + \epsilon_t$) or difference stationary (i.e., $y_t = b_0 + y_{t-1} + \epsilon_t$).

Consider k-step ahead prediction for both cases.

$$\text{(Trend Stationarity)} \qquad y_{t+k|t} = a_0 + a_1(t+k)$$

$$\text{(Difference Stationarity)} \qquad y_{t+k|t} = b_0 k + y_t$$

2. **The Case of $|\phi_1| < 1$:**

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \qquad y_0 = 0, \qquad t = 1, \cdots, T$$

Then, OLSE of ϕ_1 is:

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}.$$

In the case of $|\phi_1| < 1$,

$$\hat{\phi}_1 = \phi_1 + \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow \phi_1 + \frac{\mathrm{E}(y_{t-1}\epsilon_t)}{\mathrm{E}(y_{t-1}^2)} = \phi_1.$$

Note as follows:

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow \mathrm{E}(y_{t-1}\epsilon_t) = 0.$$

By the central limit theorem,

$$\frac{\overline{y\epsilon} - \mathrm{E}(\overline{y\epsilon})}{\sqrt{\mathrm{V}(\overline{y\epsilon})}} \longrightarrow N(0,1)$$

where

$$\overline{y\epsilon} = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t.$$

$$
\begin{aligned}
\mathrm{E}(\overline{y\epsilon}) &= 0, \\
\mathrm{V}(\overline{y\epsilon}) &= \mathrm{V}(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t) = \mathrm{E}\Big((\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t)^2\Big) \\
&= \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T}\sum_{s=1}^{T} y_{t-1}y_{s-1}\epsilon_t\epsilon_s\Big) = \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T} y_{t-1}^2\epsilon_t^2\Big) = \frac{1}{T}\sigma_\epsilon^2\gamma(0).
\end{aligned}
$$

Therefore,

$$
\frac{\overline{y\epsilon}}{\sqrt{\sigma_\epsilon^2\gamma(0)/T}} = \frac{1}{\sigma_\epsilon\sqrt{\gamma(0)}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0,1),
$$

which is rewritten as:

$$
\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0,\sigma_\epsilon^2\gamma(0)).
$$

Using $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \mathrm{E}(y_{t-1}^2) = \gamma(0)$, we have the following asymptotic distribution:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\Big(0, \frac{\sigma_\epsilon^2}{\gamma(0)}\Big) = N\big(0, 1-\phi_1^2\big).$$

Note that $\gamma(0) = \frac{\sigma_\epsilon^2}{1-\phi_1^2}$.

3. In the case of $\phi_1 = 1$, as expected, we have:

$$\sqrt{T}(\hat{\phi}_1 - 1) \longrightarrow 0.$$

That is, $\hat{\phi}_1$ has the distribution which converges in probability to $\phi_1 = 1$ (i.e., degenerated distribution).

Is this true?

4. **The Case of $\phi_1 = 1$:** $\Longrightarrow$ Random Walk Process

$y_t = y_{t-1} + \epsilon_t$ with $y_0 = 0$ is written as:

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \cdots + \epsilon_1.$$

Therefore, we can obtain:

$$y_t \sim N(0, \sigma_\epsilon^2 t).$$

The variance of y_t depends on time t. $\Longrightarrow$ y_t is nonstationary.

5. Remember that $\hat{\phi}_1 = \phi_1 + \dfrac{\sum y_{t-1}\epsilon_t}{\sum y_{t-1}^2}$.

   (a) First, consider the numerator $\sum y_{t-1}\epsilon_t$.

   We have $y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2$.

Therefore, we obtain:

$$y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2).$$

Taking into account $y_0 = 0$, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2.$$

Divided by $\sigma_\epsilon^2 T$ on both sides, we have the following:

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2.$$

From $y_t \sim N(0, \sigma_\epsilon^2 t)$, we obtain the following result:

$$\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 \sim \chi^2(1).$$

Moreover, the second term is derived from:

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \longrightarrow \mathrm{E}(\epsilon_t^2) = \sigma_\epsilon^2.$$

Therefore,

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \longrightarrow \frac{1}{2}(\chi^2(1) - 1).$$

(b) Next, consider $\sum y_{t-1}^2$.

$$\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) = \sum_{t=1}^{T}\mathrm{E}(y_{t-1}^2) = \sum_{t=1}^{T}\sigma_\epsilon^2(t-1) = \sigma_\epsilon^2\frac{T(T-1)}{2}.$$

Thus, we obtain the following result:

$$\frac{1}{T^2}\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) \longrightarrow \text{a fixed value, i.e., } \frac{\sigma_\epsilon^2}{2}.$$

Therefore,

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \text{ a distribution.}$$

6. Summarizing the results up to now, $T(\hat{\phi}_1 - \phi_1)$, not $\sqrt{T}(\hat{\phi}_1 - \phi_1)$, has limiting distribution in the case of $\phi_1 = 1$.

$$T(\hat{\phi}_1 - \phi_1) = \frac{(1/T)\sum y_{t-1}\epsilon_t}{(1/T^2)\sum y_{t-1}^2} \longrightarrow \text{ a distribution.}$$

7. Basic Concepts of Random Walk Process:

   (a) Model: $y_t = y_{t-1} + \epsilon_t, \qquad y_0 = 0, \qquad \epsilon_t \sim N(0, 1).$

   Then,

$$y_t = \epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1.$$

   Therefore,

$$y_t \sim N(0, t).$$

$\Longrightarrow$ Nonstationary Process (i.e., variance depends on time t.)

Difference between y_s and y_t ($s > t$) is:

$$y_s - y_t = \epsilon_s + \epsilon_{s-1} + \cdots + \epsilon_{t+2} + \epsilon_{t+1}.$$

The distribution of $y_s - y_t$ is:

$$y_s - y_t \sim N(0, s - t).$$

(b) Rewrite as follows:

$$\begin{aligned} y_t &= y_{t-1} + \epsilon_t \\ &= y_{t-1} + e_{1,t} + e_{2,t} + \cdots + e_{N,t}, \end{aligned}$$

where $\epsilon_t = e_{1,t} + e_{2,t} + \cdots + e_{N,t}$.

$e_{1,t}, e_{2,t}, \cdots, e_{N,t}$ are iid with $e_{i,t} \sim N(0, 1/N)$.

That is, suppose that there are N subperiods between time t and time $t+1$.

The limit when $N \to \infty$ is a **continuous time (連続時間)** process known as **standard Brownian motion** or **Wiener process**.

The value of this process at time r is denoted by $W(r)$ for $0 \le r \le 1$.

Definition:

Standard Brownian motion $W(r)$ denotes a continuous-time variable at time r and a stochastic function.

$W(r)$ for $r \in [0, 1]$ satisfies the following:

i. $W(0) = 0$

ii. For any time periods $0 \le r_1 < r_2 < \cdots < r_k \le 1$, $W(r_2) - W(r_1)$, $W(r_3) - W(r_2)$, $\cdots$, $W(r_k) - W(r_{k-1})$ are independently multivariate normal with $W(s) - W(t) \sim N(0, s - t)$ for $s > t$.

iii. $W(r)$ is continuous in r with probability 1.

An example:

$$\sigma W(r) \sim N(0, \sigma^2 r),$$

which denotes the Brownian motion with variance σ^2.

Another example;

$$W(r)^2 \sim r \times \chi^2(1).$$

(c) Assume $\epsilon_t \sim$ iid $(0, \sigma_\epsilon^2)$. Define $X_T(r)$ for $r \in [0, 1]$ as follows:

$$X_T(r) = \begin{cases} 0, & 0 \le r < \dfrac{1}{T} \\ \dfrac{\epsilon_1}{T}, & \dfrac{1}{T} \le r < \dfrac{2}{T} \\ \dfrac{\epsilon_1 + \epsilon_2}{T}, & \dfrac{2}{T} \le r < \dfrac{3}{T} \\ \quad\vdots & \quad\vdots \\ \dfrac{\epsilon_1 + \epsilon_2 + \cdots + \epsilon_T}{T}, & r = 1 \end{cases}$$

Let $[Tr]$ be the largest integer which is less than or equal to $T \times r$.

$$X_T(r) \equiv \frac{1}{T} \sum_{t=1}^{[Tr]} \epsilon_t, \qquad \sqrt{T} X_T(r) \longrightarrow N(0, r\sigma_\epsilon^2).$$

Note that

$$\frac{1}{T} \sum_{t=1}^{[Tr]} \epsilon_t = \frac{[Tr]}{T} \frac{1}{[Tr]} \sum_{t=1}^{[Tr]} \epsilon_t,$$

$$\frac{[Tr]}{T} \longrightarrow r, \qquad \frac{1}{\sqrt{[Tr]}} \sum_{t=1}^{[Tr]} \epsilon_t \longrightarrow N(0, \sigma_\epsilon^2),$$

$$\sqrt{T} X_T(r) = \frac{[Tr]}{T} \sqrt{\frac{T}{[Tr]}} \frac{1}{\sqrt{[Tr]}} \sum_{t=1}^{[Tr]} \epsilon_t, \qquad \sqrt{\frac{T}{[Tr]}} \longrightarrow \frac{1}{\sqrt{r}}.$$

Therefore, we obtain:

$$\sqrt{T} X_T(r) \longrightarrow N(0, r\sigma_\epsilon^2).$$

Moreover, we have the following results:

$$\frac{\sqrt{T} X_T(r)}{\sigma_\epsilon} \longrightarrow N(0, r) = W(r),$$

$$\frac{\sqrt{T}\,(X_T(r_2) - X_T(r_1))}{\sigma_\epsilon} \longrightarrow W(r_2) - W(r_1) = N(0, r_2 - r_1).$$

For example, consider:

$$X_T(1) = \frac{1}{T}\sum_{t=1}^{T}\epsilon_t.$$

Then,

$$\frac{\sqrt{T}X_T(1)}{\sigma_\epsilon} = \frac{1}{\sigma_\epsilon\sqrt{T}}\sum_{t=1}^{T}\epsilon_t \longrightarrow W(1) = N(0,1).$$

(d) Consider $y_t = y_{t-1} + \epsilon_t,\ y_0 = 0$ and $\epsilon_t \sim N(0, \sigma_\epsilon^2)$.

$X_T(r)$ is defined as follows:

$$X_T(r) = \begin{cases} 0, & 0 \le r < \dfrac{1}{T}, \\ \dfrac{y_1}{T}, & \dfrac{1}{T} \le r < \dfrac{2}{T}, \\ \dfrac{y_2}{T}, & \dfrac{2}{T} \le r < \dfrac{3}{T}, \\ \quad\vdots & \quad\vdots \\ \dfrac{y_{T-1}}{T}, & \dfrac{T-1}{T} \le r < 1, \\ \dfrac{y_T}{T}, & r = 1. \end{cases}$$

Define $S_T(r)$ as follows:

$$S_T(r) = \begin{cases} 0, & 0 \le r < \dfrac{1}{T}, \\ \dfrac{y_1^2}{T}, & \dfrac{1}{T} \le r < \dfrac{2}{T}, \\ \dfrac{y_2^2}{T}, & \dfrac{2}{T} \le r < \dfrac{3}{T}, \\ \quad\vdots & \quad\vdots \\ \dfrac{y_{T-1}^2}{T}, & \dfrac{T-1}{T} \le r < 1, \\ \dfrac{y_T^2}{T}, & r = 1. \end{cases}$$

To obtain $\displaystyle\int_0^1 X_T(r)\mathrm{d}r$ and $\displaystyle\int_0^1 S_T(r)\mathrm{d}r$, we compute a sum of rectangulars as follows:

$$\int_0^1 X_T(r)\mathrm{d}r \approx \frac{y_1}{T}\left(\frac{2}{T} - \frac{1}{T}\right) + \frac{y_2}{T}\left(\frac{3}{T} - \frac{2}{T}\right) + \ \cdots \ + \frac{y_{T-1}}{T}\left(1 - \frac{T-1}{T}\right)$$

$$= \frac{y_1}{T^2} + \frac{y_2}{T^2} + \cdots + \frac{y_{T-1}}{T^2} = \frac{1}{T^2} \sum_{t=1}^{T} y_t,$$

$$\int_0^1 S_T(r)\mathrm{d}r \approx \frac{y_1^2}{T}\left(\frac{2}{T} - \frac{1}{T}\right) + \frac{y_2^2}{T}\left(\frac{3}{T} - \frac{2}{T}\right) + \cdots + \frac{y_{T-1}^2}{T}\left(1 - \frac{T-1}{T}\right)$$

$$= \frac{y_1^2}{T^2} + \frac{y_2^2}{T^2} + \cdots + \frac{y_{T-1}^2}{T^2} = \frac{1}{T^2} \sum_{t=1}^{T} y_t^2.$$

We have already known that $\sqrt{T} X_T(r) \longrightarrow \sigma_\epsilon W(r)$.

Therefore,

$$\int_0^1 \sqrt{T} X_T(r)\mathrm{d}r \longrightarrow \sigma_\epsilon \int_0^1 W(r)\mathrm{d}r.$$

That is,

$$\frac{1}{T^{3/2}} \sum_{t=1}^{T} y_t \longrightarrow \sigma_\epsilon \int_0^1 W(r)\mathrm{d}r.$$

From $S_T(r) \equiv \left(\sqrt{T} X_T(r)\right)^2$,

$$S_T(r) \equiv \left(\sqrt{T} X_T(r)\right)^2 \longrightarrow \sigma_\epsilon^2 (W(r))^2,$$

which is called the continuous mapping theorem.

(*) **Continuous Mapping Theorem (連続写像定理):**

if $x_T \longrightarrow x$ (convergence in distribution) and $g(\cdot)$ is a continuous function,

then $g(x_T) \longrightarrow g(x)$ (convergence in distribution).

Threfore, we have the follwoing result:

$$\int_0^1 S_T(r)\mathrm{d}r \longrightarrow \sigma_\epsilon^2 \int_0^1 (W(r))^2 \, \mathrm{d}r.$$

That is,

$$\frac{1}{T^2}\sum_{t=1}^{T} y_t^2 \longrightarrow \sigma_\epsilon^2 \int_0^1 (W(r))^2 \, \mathrm{d}r.$$

8. Asymptotic Distribution of AR(1) Model:

(a) $\boldsymbol{H_0 : y_t = y_{t-1} + \epsilon_t}$ **and** $\boldsymbol{H_1 : y_t = \phi_1 y_{t-1} + \epsilon_t}$ **for** $\boldsymbol{|\phi_1| < 1}$

OLSE of ϕ_1, denoted by $\hat{\phi}_1$, is given by:

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} = \phi_1 + \frac{\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

Using $\phi_1 = 1$ and some formulas shown above, we obtain:

$$T(\hat{\phi}_1 - 1) = \frac{T^{-1}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{T^{-2}\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow \frac{\frac{1}{2}\Big((W(1))^2 - 1\Big)}{\int_0^1 (W(r))^2 \,\mathrm{d}r}$$

Remember that

$$T^{-1}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow \frac{1}{2}\sigma_\epsilon^2\Big((W(1))^2 - 1\Big)$$

and

$$T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \sigma_\epsilon^2 \int_0^1 (W(r))^2 \, \mathrm{d}r,$$

where $(W(1))^2 = \chi^2(1)$.

We say that $\hat{\phi}_1$ is **super-consistent (超一致性)** or ***T*-consistent**.

Remember that when $|\phi_1| < 1$ we have $\sqrt{T}(\hat{\phi}_1 - \phi_1) \longrightarrow N(0, 1 - \phi_1^2)$, and in this case we say that $\hat{\phi}_1$ is **$\sqrt{T}$-consistent**.

Conventional t test statistic is given by:

$$t = \frac{\hat{\phi}_1 - 1}{s_\phi},$$

where

$$s_\phi = \left(s^2 \Big/ \sum_{t=1}^{T} y_{t-1}^2 \right)^{1/2} \quad \text{and} \quad s^2 = \frac{1}{T-1} \sum_{t=1}^{T} (y_t - \hat{\phi}_1 y_{t-1})^2.$$

Next, consider t statistic.

The t test statistic, denoted by t, is represented as follows:

$$t = \frac{\hat{\phi}_1 - 1}{s_\phi} = \frac{T(\hat{\phi}_1 - 1)}{T s_\phi}$$

The denominator is:

$$T s_\phi = \left(s^2 \Big/ \frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}^2 \right)^{1/2}$$
$$\longrightarrow \left(\sigma_\epsilon^2 \Big/ \left(\sigma_\epsilon^2 \int_0^1 (W(r))^2 \, \mathrm{d}r \right) \right)^{1/2} = \left(\int_0^1 (W(r))^2 \, \mathrm{d}r \right)^{-1/2},$$

where $s^2 \longrightarrow \sigma_\epsilon^2$ is utilized.

Therefore, we have the following asymptotic distribution:

$$
\begin{aligned}
t = \frac{\hat{\phi}_1 - 1}{s_\phi} \longrightarrow & \frac{\frac{1}{2}\left((W(1))^2 - 1\right)}{\displaystyle\int_0^1 (W(r))^2 \, \mathrm{d}r} \Bigg/ \left(\int_0^1 (W(r))^2 \, \mathrm{d}r\right)^{-1/2} \\
= & \frac{\frac{1}{2}\left((W(1))^2 - 1\right)}{\left(\displaystyle\int_0^1 (W(r))^2 \, \mathrm{d}r\right)^{1/2}}.
\end{aligned}
$$

Therefore, the distribution of the t statistic shown above is different from the t distribution.

(b) $\boldsymbol{H_0 : y_t = y_{t-1} + \epsilon_t}$ **and** $\boldsymbol{H_1 : y_t = \alpha_0 + \phi_1 y_{t-1} + \epsilon_t}$ **for** $\boldsymbol{|\phi_1| < 1}$

$$\begin{pmatrix} \hat{\alpha}_0 \\ \hat{\phi}_1 \end{pmatrix} = \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_t \\ \sum y_{t-1} y_t \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_0 \\ \phi_1 \end{pmatrix} + \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1} \epsilon_t \end{pmatrix}$$

In the true model, $\alpha_0 = 0$ and $\phi_1 = 1$.

$$\begin{pmatrix} \hat{\alpha}_0 \\ \hat{\phi}_1 - 1 \end{pmatrix} = \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1} \epsilon_t \end{pmatrix}$$

$$= \begin{pmatrix} O_p(T) & O_p(T^{3/2}) \\ O_p(T^{3/2}) & O_p(T^2) \end{pmatrix}^{-1} \begin{pmatrix} O_p(T^{1/2}) \\ O_p(T) \end{pmatrix}$$

(*) For random variable x and constant k, $x = O_p(k)$ implies that x/k converges in distribution.

To change each element of the matrices to $O_p(1)$, we use the following matrix:

$$\Gamma = \begin{pmatrix} T^{1/2} & 0 \\ 0 & T \end{pmatrix}.$$

Multiplying the above matrix from the left, we obtain the following:

$$\Gamma\begin{pmatrix} \hat{\alpha}_0 \\ \hat{\phi}_1 - 1 \end{pmatrix} = \begin{pmatrix} T^{1/2}\hat{\alpha}_0 \\ T(\hat{\phi}_1 - 1) \end{pmatrix} = \Gamma\begin{pmatrix} O_p(T) & O_p(T^{3/2}) \\ O_p(T^{3/2}) & O_p(T^2) \end{pmatrix}^{-1} \Gamma\Gamma^{-1}\begin{pmatrix} O_p(T^{1/2}) \\ O_p(T) \end{pmatrix}$$

$$= \left(\Gamma^{-1}\begin{pmatrix} O_p(T) & O_p(T^{3/2}) \\ O_p(T^{3/2}) & O_p(T^2) \end{pmatrix}\Gamma^{-1}\right)^{-1} \Gamma^{-1}\begin{pmatrix} O_p(T^{1/2}) \\ O_p(T) \end{pmatrix}$$

$$= \left(\Gamma^{-1}\begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}\Gamma^{-1}\right)^{-1} \Gamma^{-1}\begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1}\epsilon_t \end{pmatrix}$$

$$= \begin{pmatrix} 1 & T^{-3/2}\sum y_{t-1} \\ T^{-3/2}\sum y_{t-1} & T^{-2}\sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2}\sum \epsilon_t \\ T^{-1}\sum y_{t-1}\epsilon_t \end{pmatrix}.$$

Each matrix converges in distribution as follows:

$$\begin{pmatrix} 1 & T^{-3/2}\sum y_{t-1} \\ T^{-3/2}\sum y_{t-1} & T^{-2}\sum y_{t-1}^2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \sigma_\epsilon \int_0^1 W(r)\mathrm{d}r \\ \sigma_\epsilon \int_0^1 W(r)\mathrm{d}r & \sigma_\epsilon^2 \int_0^1 (W(r))^2\,\mathrm{d}r \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r)\mathrm{d}r \\ \int_0^1 W(r)\mathrm{d}r & \int_0^1 (W(r))^2\,\mathrm{d}r \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix},$$

$$\begin{pmatrix} T^{-1/2}\sum \epsilon_t \\ T^{-1}\sum y_{t-1}\epsilon_t \end{pmatrix} \longrightarrow \begin{pmatrix} \sigma_\epsilon W(1) \\ \frac{1}{2}\sigma_\epsilon^2\left((W(1))^2 - 1\right) \end{pmatrix} = \sigma_\epsilon \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \begin{pmatrix} W(1) \\ \frac{1}{2}\left((W(1))^2 - 1\right) \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} T^{1/2}\hat{\alpha}_0 \\ T(\hat{\phi}_1 - 1) \end{pmatrix} \longrightarrow \left[\begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r)\mathrm{d}r \\ \int_0^1 W(r)\mathrm{d}r & \int_0^1 (W(r))^2\,\mathrm{d}r \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix}\right]^{-1}$$

$$\times \sigma_\epsilon \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \begin{pmatrix} W(1) \\ \frac{1}{2}\left((W(1))^2 - 1\right) \end{pmatrix}.$$

Finally, $T(\hat{\phi}_1 - 1)$ converges to the following distribution:

$$T(\hat{\phi}_1 - 1) \longrightarrow \frac{\frac{1}{2}\left((W(1))^2 - 1\right) - W(1) \int_0^1 W(r)\mathrm{d}r}{\int_0^1 (W(r))^2\,\mathrm{d}r - \left(\int_0^1 W(r)\mathrm{d}r\right)^2}.$$

The t test statistic is:

$$t = \frac{\hat{\phi}_1 - 1}{\left(s_\phi^2\right)^{1/2}} = \frac{T(\hat{\phi}_1 - 1)}{\left(T^2 s_\phi^2\right)^{1/2}},$$

where

$$s_\phi^2 = s^2 \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$s^2 = \frac{1}{T-2} \sum_{t=1}^{T} (y_t - \hat{\alpha}_0 - \hat{\phi}_1 y_{t-1})^2.$$

The denominator $T^2 s_\phi^2$ converges in distribution as follows:

$$T^2 s_\phi^2 \longrightarrow \sigma_\epsilon^2 \begin{pmatrix} 0 & 1 \end{pmatrix} \left(\begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r)\mathrm{d}r \\ \int_0^1 W(r)\mathrm{d}r & \int_0^1 (W(r))^2 \,\mathrm{d}r \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \right)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \frac{1}{\int_0^1 (W(r))^2 \,\mathrm{d}r - \left(\int_0^1 W(r)\mathrm{d}r \right)^2}$$

Thus, the t test statistic converges to the following distribution:

$$t \longrightarrow \frac{\frac{1}{2}\left((W(1))^2 - 1\right) - W(1)\int_0^1 W(r)\mathrm{d}r}{\left(\int_0^1 (W(r))^2\,\mathrm{d}r - \left(\int_0^1 W(r)\mathrm{d}r\right)^2\right)^{1/2}}.$$

(c) $\boldsymbol{H_0 : y_t = \alpha_0 + y_{t-1} + \epsilon_t}$ **and** $\boldsymbol{H_1 : y_t = \alpha_0 + \phi_1 y_{t-1} + \epsilon_t}$ **for** $|\boldsymbol{\phi_1}| < 1$

$$\begin{pmatrix} T^{1/2}(\hat{\alpha}_0 - \alpha_0) \\ T^{3/2}(\hat{\phi}_1 - 1) \end{pmatrix} \longrightarrow N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_\epsilon^2 \begin{pmatrix} 1 & \frac{\alpha_0}{2} \\ \frac{\alpha_0}{2} & \frac{\alpha_0^2}{3} \end{pmatrix}\right).$$

(abbr.)

(d) $\boldsymbol{H_0 : y_t = \alpha_0 + y_{t-1} + \epsilon_t}$ **and**

$\boldsymbol{H_1 : y_t = \alpha_0 + \alpha_1 t + \phi_1 y_{t-1} + \epsilon_t}$ **for** $|\boldsymbol{\phi_1}| < 1$

(abbr.)

9. The distributions of the t statistic: $\dfrac{\hat{\phi}_1 - 1}{s_{\phi}}$

t Distribution

T	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
25	−2.49	−2.06	−1.71	−1.32	1.32	1.71	2.06	2.49
50	−2.40	−2.01	−1.68	−1.30	1.30	1.68	2.01	2.40
100	−2.36	−1.98	−1.66	−1.29	1.29	1.66	1.98	2.36
250	−2.34	−1.97	−1.65	−1.28	1.28	1.65	1.97	2.34
500	−2.33	−1.96	−1.65	−1.28	1.28	1.65	1.96	2.33
∞	−2.33	−1.96	−1.64	−1.28	1.28	1.64	1.96	2.33

(a) $H_0: \; y_t = y_{t-1} + \epsilon_t$

$H_1: \; y_t = \phi_1 y_{t-1} + \epsilon_t$ **for** $\phi_1 < 1$ **or** $-1 < \phi_1$

T	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
25	−2.66	−2.26	−1.95	−1.60	0.92	1.33	1.70	2.16
50	−2.62	−2.25	−1.95	−1.61	0.91	1.31	1.66	2.08
100	−2.60	−2.24	−1.95	−1.61	0.90	1.29	1.64	2.03
250	−2.58	−2.23	−1.95	−1.62	0.89	1.29	1.63	2.01
500	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00
∞	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00

(b) $\boldsymbol{H_0 : y_t = y_{t-1} + \epsilon_t}$

$\boldsymbol{H_1 : y_t = \alpha_0 + \phi_1 y_{t-1} + \epsilon_t}$ **for** $\boldsymbol{\phi_1 < 1}$ **or** $\boldsymbol{-1 < \phi_1}$

T	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
25	−3.75	−3.33	−3.00	−2.63	−0.37	0.00	0.34	0.72
50	−3.58	−3.22	−2.93	−2.60	−0.40	−0.03	0.29	0.66
100	−3.51	−3.17	−2.89	−2.58	−0.42	−0.05	0.26	0.63
250	−3.46	−3.14	−2.88	−2.57	−0.42	−0.06	0.24	0.62
500	−3.44	−3.13	−2.87	−2.57	−0.43	−0.07	0.24	0.61
∞	−3.43	−3.12	−2.86	−2.57	−0.44	−0.07	0.23	0.60

(d) $H_0: \; y_t = \alpha_0 + y_{t-1} + \epsilon_t$

$H_1: \; y_t = \alpha_0 + \alpha_1 t + \phi_1 y_{t-1} + \epsilon_t$ **for** $\phi_1 < 1$ **or** $-1 < \phi_1$

T	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
25	−4.38	−3.95	−3.60	−3.24	−1.14	−0.80	−0.50	−0.15
50	−4.15	−3.80	−3.50	−3.18	−1.19	−0.87	−0.58	−0.24
100	−4.04	−3.73	−3.45	−3.15	−1.22	−0.90	−0.62	−0.28
250	−3.99	−3.69	−3.43	−3.13	−1.23	−0.92	−0.64	−0.31
500	−3.98	−3.68	−3.42	−3.13	−1.24	−0.93	−0.65	−0.32
∞	−3.96	−3.66	−3.41	−3.12	−1.25	−0.94	−0.66	−0.33

6.2 Serially Correlated Errors

Consider the case where the error term is serially correlated.

6.2.1 Augmented Dickey-Fuller (ADF) Test

Consider the following AR(p) model:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t, \qquad \epsilon_t \sim \text{iid}(0, \sigma_\epsilon^2),$$

which is rewritten as:

$$\phi(L) y_t = \epsilon_t.$$

When the above model has a unit root, we have $\phi(1) = 0$, i.e., $\phi_1 + \phi_2 + \cdots + \phi_p = 1$.

The above AR(p) model is written as:

$$y_t = \rho y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots + +\delta_{p-1} \Delta y_{t-p+1} + \epsilon_t,$$

where $\rho = \phi_1 + \phi_2 + \cdots + \phi_p$ and $\delta_j = -(\phi_{j+1} + \phi_{j+2} + \cdots + \phi_p)$.

The null and alternative hypotheses are:

$$H_0: \ \rho = 1 \text{ (Unit root)},$$

$$H_1 : \rho < 1 \text{ (Stationary)}.$$

Use the t test, where we have the same asymptotic distributions.

We can utilize the same tables as before.

Choose p by AIC or SBIC.

Use $N(0,1)$ to test $H_0: \ \delta_j = 0$ against $H_1: \ \delta_j \neq 0$ for $j = 1, 2, \cdots, p-1$.

Reference

Kurozumi (2008) "Economic Time Series Analysis and Unit Root Tests: Development and Perspective," *Japan Statistical Society*, Vol.38, Series J, No.1, pp.39 – 57.

Download the above paper from:

`http://ci.nii.ac.jp/vol_issue/nels/AA11989749/ISS0000426576_ja.html`