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 $\int_0^1 X_T(r)\mathrm{d}r$ and $\int_0^1 S_T(r)\mathrm{d}r$, 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:

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

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 $\boldsymbol{\sqrt{T}}$**-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:

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

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

Therefore, we have the following asymptotic distribution:

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

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

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:

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

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: $\frac{\hat{\phi}_1 - 1}{s_\phi}$

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

6.3 Cointegration (共和分)

1. For a scalar y_t, when $(1-L)^d y_t$ is stationary, we write $y_t \sim I(d)$.

   When $\Delta y_t = y_t - y_{t-1}$ is stationary, we write $\Delta y_t \sim I(0)$ or $y_t \sim I(1)$.

2. **Definition of Cointegration:**

   Suppose that each series in a $g \times 1$ vector y_t is $I(1)$, i.e., each series has unit root, and that a linear combination of each series (i.e, $a'y_t$ for a nonzero vector a) is $I(0)$, i.e., stationary.

   Then, we say that y_t has a cointegration.

3. **Example:**

   Suppose that $y_t = (y_{1,t},\ y_{2,t})'$ is the following vector autoregressive process:

$$y_{1,t} = \gamma y_{2,t} + \epsilon_{1,t},$$

$$y_{2,t} = y_{2,t-1} + \epsilon_{2,t}.$$

Then,

$$\Delta y_{1,t} = \gamma\epsilon_{2,t} + \epsilon_{1,t} - \epsilon_{1,t-1}, \quad \text{(MA(1) process)},$$

$$\Delta y_{2,t} = \epsilon_{2,t},$$

where both $y_{1,t}$ and $y_{2,t}$ are $I(1)$ processes.

The linear combination $y_{1,t} - \gamma y_{2,t}$ is $I(0)$.

In this case, we say that $y_t = (y_{1,t},\ y_{2,t})'$ is cointegrated with $a = (1,\ -\gamma)$.

$a = (1,\ -\gamma)$ is called the **cointegrating vector (共和分ベクトル)**, which is not unique. Therefore, the first element of a is set to be one.

4. Suppose that $y_t \sim I(1)$ and $x_t \sim I(1)$.

For the regression model $y_t = x_t\beta + u_t$, OLS does not work well if we do not have the β which satisfies $u_t \sim I(0)$.

$\Longrightarrow$ **Spurious regression (見せかけの回帰)**

5. Suppose that $y_t \sim I(1)$, y_t is a $g \times 1$ vector and $y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}$.

   $y_{2,t}$ is a $k \times 1$ vector, where $k = g - 1$.

   Consider the following regression model:

$$y_{1,t} = \alpha + \gamma' y_{2,t} + u_t, \qquad t = 1, 2, \cdots, T.$$

   OLSE is given by:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2,t}' \\ \sum y_{2,t} & \sum y_{2,t} y_{2,t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1,t} \\ \sum y_{1,t} y_{2,t} \end{pmatrix}.$$

   Next, consider testing the null hypothesis $H_0: \ R\gamma = d$, where R is a $G \times k$

matrix ($G \leq k$) and r is a $G \times 1$ vector. G denotes the number of the linear restrictions.

The F statistic, denoted by F, is given by:

$$F = \frac{1}{G}(R\hat{\gamma} - d)' \left(s^2 \begin{pmatrix} 0 & R \end{pmatrix} \begin{pmatrix} T & \sum y'_{2,t} \\ \sum y_{2,t} & \sum y_{2,t} y'_{2,t} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ R' \end{pmatrix} \right)^{-1} (R\hat{\gamma} - d),$$

where

$$s^2 = \frac{1}{T - g} \sum_{t=1}^{T} (y_{1,t} - \hat{\alpha} - \hat{\gamma}' y_{2,t})^2.$$

When we have the γ such that $y_{1,t} - \gamma y_{2,t}$ is stationary, OLSE of γ, i.e., $\hat{\gamma}$, is not statistically equal to zero.

When the sample size T is large enough, H_0 is rejected by the F test.

6. Phillips, P.C.B. (1986) "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, Vol.33, pp.95 – 131.

Consider a $g \times 1$ vector y_t whose first difference is described by:

$$\Delta y_t = \Psi(L)\epsilon_t = \sum_{s=0}^{\infty} \Psi_s \epsilon_{t-s},$$

for ϵ_t an i.i.d. $g \times 1$ vector with mean zero , variance $\mathrm{E}(\epsilon_t \epsilon_t') = PP'$, and finite fourth moments and where $\{s\Psi_s\}_{s=0}^{\infty}$ is absolutely summable.

Let $k = g - 1$ and $\Lambda = \Psi(1)P$.

Partition y_t as $y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}$ and $\Lambda\Lambda'$ as $\Lambda\Lambda' = \begin{pmatrix} \Sigma_{11} & \Sigma_{21}' \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where $y_{1,t}$ and Σ_{11} are scalars, $y_{2,t}$ and Σ_{21} are $k \times 1$ vectors, and Σ_{22} is a $k \times k$ matrix.

Suppose that $\Lambda\Lambda'$ is nonsingular,and define $\sigma_1^2 = \Sigma_{11} - \Sigma_{21}'\Sigma_{22}^{-1}\Sigma_{21}$.

Let L_{22} denote the Cholesky factor of Σ_{22}^{-1}, i.e., L_{22} is the lower triangular matrix satisfying $\Sigma_{22}^{-1} = L_{22}L_{22}'$.

Then, (a) – (c) hold.

(a) OLSEs of α and γ in the regression model $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$, denoted by $\hat{\alpha}_T$ and $\hat{\gamma}_T$, are characterized by:

$$\begin{pmatrix} T^{-1/2}\hat{\alpha}_T \\ \hat{\gamma}_T - \Sigma_{22}^{-1}\Sigma_{21} \end{pmatrix} \longrightarrow \begin{pmatrix} \sigma_1 h_1 \\ \sigma_1 L_{22} h_2 \end{pmatrix},$$

where

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} 1 & \int_0^1 W_2(r)'\mathrm{d}r \\ \int_0^1 W_2(r)\mathrm{d}r & \int_0^1 W_2(r)W_2(r)'\mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} \int_0^1 W_1(r)\mathrm{d}r \\ \int_0^1 W_2(r)W_1(r)\mathrm{d}r \end{pmatrix},$$

where $W_1(r)$ and $W_2(r)$ denote scalar and g-dimensional standard Brownian motions, and $W_1(r)$ is independent of $W_2(r)$.

(b) The sum of squared residuals, denoted by $\mathrm{RSS}_T = \sum_{t=1}^{T} \hat{u}_t^2$, satisfies

$$T^{-2}\mathrm{RSS}_T \longrightarrow \sigma_1^2 H,$$

where

$$H = \int_0^1 (W_1(r))^2 \mathrm{d}r - \left(\begin{pmatrix} \int_0^1 W_1(r)\mathrm{d}r \\ \int_0^1 W_2(r)W_1(r)\mathrm{d}r \end{pmatrix}' \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right)^{-1}.$$

(c) The F test satisfies:

$$
\begin{aligned}
T^{-1}F \longrightarrow\ & \frac{1}{G}(\sigma_1 R^* h_2 - d^*)' \\
& \times\left(\sigma_1^2 H \begin{pmatrix} 0 & R^* \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W_2(r)'\mathrm{d}r \\ \int_0^1 W_2(r)\mathrm{d}r & \int_0^1 W_2(r)W_2^*(r)'\mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} 0 & R^* \end{pmatrix}'\right)^{-1} \\
& \times(\sigma_1 R^* h_2 - d^*),
\end{aligned}
$$

where $R^* = RL_{22}$ and $d^* = d - R\Sigma_{22}^{-1}\Sigma_{21}$.

(a) indicates that OLSE $\hat{\gamma}_T$ is not consistent.

(b) indicates that $s^2 = \dfrac{1}{T-g}\sum_{t=1}^{T}\hat{u}_t^2$ diverges.

(c) indicates that F diverges.

$\Longrightarrow$ **Spurious regression (見せかけの回帰)**

7. **Resolution for Spurious Regression:**

Suppose that $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ is a spurious regression.

(1) Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + \phi y_{1,t-1} + \delta y_{2,t-1} + u_t$.

Then, $\hat{\gamma}_T$ is $\sqrt{T}$-consistent, and the t test statistic goes to the standard normal distribution under $H_0: \ \gamma = 0$.

(2) Estimate $\Delta y_{1,t} = \alpha + \gamma' \Delta y_{2,t} + u_t$. Then, $\hat{\alpha}_T$ and $\hat{\beta}_T$ are $\sqrt{T}$-consistent, and the t test and F test make sense.

(3) Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ by the Cochrane-Orcutt method, assuming that u_t is the first-order serially correlated error.

Usually, choose (2).

However, there are two exceptions.

(i) The true value of ϕ in (1) above is not one, i.e., less than one.

(ii) $y_{1,t}$ and $y_{2,t}$ are the cointegrated processes.

In these two cases, taking the first difference leads to the misspecified regression.

8. **Cointegrating Vector:**

Suppose that each element of y_t is $I(1)$ and that $a'y_t$ is $I(0)$.

a is called a **cointegrating vector (共和分ベクトル),** which is not unique.

Set $z_t = a'y_t$, where z_t is scalar, and a and y_t are $g \times 1$ vectors.

For $z_t \sim I(0)$ (i.e., stationary),

$$T^{-1}\sum_{t=1}^{T} z_t^2 = T^{-1}\sum_{t=1}^{T}(a'y_t)^2 \longrightarrow \mathrm{E}(z_t^2).$$

For $z_t \sim I(1)$ (i.e., nonstationary, i.e., a is not a cointegrating vector),

$$T^{-2}\sum_{t=1}^{T}(a'y_t)^2 \longrightarrow \lambda^2 \int_0^1 (W(r))^2\,\mathrm{d}r,$$

where $W(r)$ denotes a standard Brownian motion and λ^2 indicates variance of $(1-L)z_t$.

If a is not a cointegrating vector, $T^{-1}\sum_{t=1}^{T} z_t^2$ diverges.

$\Longrightarrow$ We can obtain a consistent estimate of a cointegrating vector by minimizing $\sum_{t=1}^{T} z_t^2$ with respect to a, where a normalization condition on a has to be imposed.

The estimator of the a including the normalization condition is super-consistent (T-consistent).

Stock, J.H. (1987) "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, Vol.55, pp.1035 – 1056.

Proposition:

Let $y_{1,t}$ be a scalar, $y_{2,t}$ be a $k \times 1$ vector, and $(y_{1,t}, y_{2,t}')'$ be a $g \times 1$ vector, where $g = k + 1$.

Consider the following model:

$$y_{1,t} = \alpha + \gamma' y_{2,t} + u_{1,t}$$

$$\Delta y_{2,t} = u_{2,t}$$

$$\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} = \Psi(L)\epsilon_t$$

ϵ_t is a $g \times 1$ i.i.d. vector with $\mathrm{E}(\epsilon_t) = 0$ and $\mathrm{E}(\epsilon_t \epsilon_t') = PP'$.

OLSE is given by:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2,t}' \\ \sum y_{2,t} & \sum y_{2,t} y_{2,t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1,t} \\ \sum y_{1,t} y_{2,t} \end{pmatrix}.$$

Define λ_1, which is a $g \times 1$ vector, and Λ_2, which is a $k \times g$ matrix, as follows:

$$\Psi(1)\, P = \begin{pmatrix} \lambda_1{}' \\ \Lambda_2 \end{pmatrix}.$$

Then, we have the following results:

$$\begin{pmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T(\hat{\gamma} - \gamma) \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \left(\Lambda_2 \int W(r) \mathrm{d}r\right)' \\ \Lambda_2 \int W(r) \mathrm{d}r & \Lambda_2 \left(\int (W(r))\,(W(r))'\, \mathrm{d}r \right) \Lambda_2{}' \end{pmatrix}^{-1} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

where

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \lambda_1{}' W(1) \\ \Lambda_2 \left(\int W(r)\,(\mathrm{d}W(r))' \right) \lambda_1 + \sum_{\tau=0}^{\infty} \mathrm{E}(u_{2,t} u_{1,t+\tau}) \end{pmatrix}.$$

$W(r)$ denotes a g-dimensional standard Brownian motion.

1) OLSE of the cointegrating vector is consistent even though u_t is serially correlated.

2) The consistency of OLSE implies that $T^{-1}\sum \hat{u}_t^2 \longrightarrow \sigma^2$.

3) Because $T^{-1}\sum(y_{1,t} - \bar{y}_1)^2$ goes to infinity, a coefficient of determination, R^2, goes to one.

6.4 Testing Cointegration

6.4.1 Engle-Granger Test

$y_t \sim I(1)$

$y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$

- $u_t \sim I(0) \implies$ Cointegration
- $u_t \sim I(1) \implies$ Spurious Regression

Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ by OLS, and obtain $\hat{u}_t$.

Estimate $\hat{u}_t = \rho \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \delta_2 \Delta \hat{u}_{t-2} + \cdots + \delta_{p-1} \Delta \hat{u}_{t-p+1} + e_t$ by OLS.

ADF Test:

- $H_0 : \rho = 1$ (Sprious Regression)
- $H_1 : \rho < 1$ (Cointegration)

$\Longrightarrow$ **Engle-Granger Test**

For example, see Engle and Granger (1987), Phillips and Ouliaris (1990) and Hansen (1992).

Asymmptotic Distribution of Residual-Based ADF Test for Cointegration

# of Refressors,	(a) Regressors have no drift				(b) Some regressors have drift			
excluding constant	1%	2.5%	5%	10%	1%	2.5%	5%	10%
1	−3.96	−3.64	−3.37	−3.07	−3.96	−3.67	−3.41	−3.13
2	−4.31	−4.02	−3.77	−3.45	−4.36	−4.07	−3.80	−3.52
3	−4.73	−4.37	−4.11	−3.83	−4.65	−4.39	−4.16	−3.84
4	−5.07	−4.71	−4.45	−4.16	−5.04	−4.77	−4.49	−4.20
5	−5.28	−4.98	−4.71	−4.43	−5.36	−5.02	−4.74	−4.46

J.D. Hamilton (1994), *Time Series Analysis*, p.766.