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

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

From $y_t \sim N(0, \sigma t)$, we obtain the following result:

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

Moreover, the second term is derived from:

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

Therefore,

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

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

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

Thus, we obtain the following result:

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

Therefore,

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

5. 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 $\phi_1 = 1$.

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

6. 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.)

Let T be the sample size.

$$\frac{y_t}{\sqrt{T}} \sim N(0, \frac{t}{T})$$

Let us define $r = \dfrac{t}{T}$.

Keeping $r = \dfrac{t}{T}$, as $T \longrightarrow \infty$,

$$\frac{y_t}{\sqrt{T}} \sim N(0, \frac{t}{T}) \longrightarrow N(0, r) = W(r)$$

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

$W(r)$ is a normal random variable with mean zero and variance r and it is a continuous function in r.

(b) **Examples:**

$W(0) = 0$

$W(1) = N(0, 1)$

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

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

$W(r)^2 = \sqrt{r} W(1)^2 \sim r \times \chi^2(1)$

(c) Suppose that $y_t = y_{t-1} + \epsilon_t, \quad y_0 = 0, \quad \epsilon_t \sim N(0, \sigma^2)$.

$y_t \sim N(0,\ \sigma^2 t)$, i.e.,

$\dfrac{y_t}{\sqrt{T}} \sim N(0,\ \sigma^2 \dfrac{t}{T})$, i.e.,

$\dfrac{y_t}{\sigma\sqrt{T}} \sim N(0,\ \dfrac{t}{T}) \longrightarrow N(0, r) = W(r)$.

- $\displaystyle\sum_{t=1}^{T} \frac{1}{T} \frac{y_t}{\sigma\sqrt{T}} \longrightarrow \int_0^1 W(r)\,\mathrm{d}r.$

$$\Big(\frac{y_t}{\sigma\sqrt{T}}\Big)^2 \longrightarrow W(r)^2.$$

- $\displaystyle\sum_{t=1}^{T} \frac{1}{T} \Big(\frac{y_t}{\sigma\sqrt{T}}\Big)^2 \longrightarrow \int_0^1 W(r)^2\,\mathrm{d}r.$

(*) **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:

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

7. 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:

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

Remember that

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

and

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

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

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

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:

$$T s_\phi = \left(s^2 \Big/ \frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}^2 \right)^{1/2}$$

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

where $s^2 \longrightarrow \sigma^2$ is utilized.

Therefore, we have the following asymptotic distribution:

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

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

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

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

$H_1: \; y_t = \alpha_0 + \phi_1 y_{t-1} + \epsilon_t$ **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	−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) $\boldsymbol{H_0 : \; y_t = \alpha_0 + y_{t-1} + \epsilon_t}$

$\boldsymbol{H_1 : \; y_t = \alpha_0 + \alpha_1 t + \phi_1 y_{t-1} + \epsilon_t}$ **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	−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

5.4 AR(p) model: Augmented Dickey-Fuller (ADF) Test

Consider the case where the error term is serially correlated.

Consider the following AR(p) model:

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

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: