- **True Model $y_t = y_{t-1} + \epsilon_t$ vs Estimated Model $y_t = \alpha + \phi y_{t-1} + \epsilon_t$:** Under $\alpha = 0$ and $\phi = 1$, we estimate α and ϕ in the regression model:

$$y_t = \alpha + \phi y_{t-1} + \epsilon_t$$

OLSEs of α and ϕ are:

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

$$= \begin{pmatrix} \alpha \\ \phi \end{pmatrix} + \frac{1}{T \sum y_{t-1}^2 - (\sum y_{t-1})^2} \begin{pmatrix} \sum y_{t-1}^2 & -\sum y_{t-1} \\ -\sum y_{t-1} & T \end{pmatrix} \begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1} \epsilon_t \end{pmatrix}$$

$$= \begin{pmatrix} \alpha \\ \phi \end{pmatrix} + \frac{1}{T \sum y_{t-1}^2 - (\sum y_{t-1})^2} \begin{pmatrix} (\sum y_{t-1}^2)(\sum \epsilon_t) - (\sum y_{t-1})(\sum y_{t-1} \epsilon_t) \\ -(\sum y_{t-1})(\sum \epsilon_t) + T(\sum y_{t-1} \epsilon_t) \end{pmatrix}$$

$$
= \begin{pmatrix} \alpha \\ \phi \end{pmatrix} + \begin{pmatrix} \dfrac{(\sum y_t^2)(\frac{1}{T}\sum \epsilon_t) - \overline{y}\sum y_{t-1}\epsilon_t}{\sum y_t^2 - T\overline{y}^2} \\ \dfrac{-\overline{y}\sum \epsilon_t + \sum y_{t-1}\epsilon_t}{\sum y_t^2 - T\overline{y}^2} \end{pmatrix}
$$

Note that $\frac{1}{T}\sum y_{t-1} \approx \frac{1}{T}\sum y_t = \overline{y}$ and $\sum y_{t-1}^2 \approx \sum y_t^2$ for large T.

In the true model, $\alpha = 0$ and $\phi = 1$.

$$
\begin{pmatrix} \hat{\alpha} \\ \hat{\phi} - 1 \end{pmatrix} = \begin{pmatrix} \dfrac{(\sum y_t^2)(\frac{1}{T}\sum \epsilon_t) - \overline{y}\sum y_{t-1}\epsilon_t}{\sum y_t^2 - T\overline{y}^2} \\ \dfrac{-\overline{y}\sum \epsilon_t + \sum y_{t-1}\epsilon_t}{\sum y_t^2 - T\overline{y}^2} \end{pmatrix}
$$

For each element of the vector, we consider each term in the numerator and denominator.

- $\sum_{t=1}^{T} y_{t-1}\epsilon_t$:

Taking the square of $y_t = y_{t-1} + \epsilon_t$ on both sides, we obtain: $y_t^2 = y_{t-1}^2 + y_{t-1}\epsilon_t + \epsilon_t^2$.

Then, we can rewrite as: $y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2)$ for $y_0 = 0$.

Taking a sum from $t = 1$ to T, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\sum_{t=1}^{T}(y_t^2 - y_{t-1}^2 - \epsilon_t^2) = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2,$$

which is divided by $T\sigma^2$ on both sides, then we obtain:

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

Note that $\frac{y_T}{\sqrt{T}\sigma} = W(1)$ and $\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \longrightarrow \mathrm{E}(\epsilon_t^2) = \sigma^2$.

- $\overline{y}$:

Note that $\dfrac{1}{T}\sum_{t=1}^{T} y_t = \overline{y}$ and $\dfrac{1}{T}\sum_{t=1}^{T} y_{t-1} = \overline{y}$. We can rewrite $\overline{y}$ as follows:

$$\overline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t = \sqrt{T}\sigma\frac{1}{T}\sum_{t=1}^{T}\frac{y_t}{\sqrt{T}\sigma}$$

which is rewritten as:

$$\frac{\overline{y}}{\sqrt{T}\sigma} = \frac{1}{T}\sum_{t=1}^{T}\frac{y_t}{\sqrt{T}\sigma} \longrightarrow \int_0^1 W(r)\mathrm{d}r$$

Note that $\dfrac{y_t}{\sqrt{T}\sigma} \longrightarrow W(r)$ as $\dfrac{t}{T} \longrightarrow r$.

- $\sum_{t=1}^{T} \epsilon_t$:

From $y_T = \sum_{t=1}^{T} \epsilon_t$, we have: $\dfrac{1}{\sqrt{T}\sigma} \sum_{t=1}^{T} \epsilon_t = \dfrac{y_T}{\sqrt{T}\sigma} = W(1)$.

- $\sum_{t=1}^{T} y_{t-1}^2$:

From $\sum_{t=1}^{T} y_{t-1}^2 \approx \sum_{t=1}^{T} y_t^2$, we obtain: $\dfrac{1}{T} \sum_{t=1}^{T} \Big(\dfrac{y_t}{\sqrt{T}\sigma}\Big)^2 \longrightarrow \int_0^1 W(r)^2 \mathrm{d}r$.

Note that $\dfrac{y_t}{\sqrt{T}\sigma} \longrightarrow W(r)$ as $\dfrac{t}{T} \longrightarrow r$.

Thus, $\hat{\phi} - 1 = \dfrac{\sum y_{t-1}\epsilon_t - \bar{y}\sum \epsilon_t}{\sum y_t^2 - T\bar{y}^2}$ is rerwitten as:

$$T(\hat{\phi} - 1) = \frac{\frac{1}{T\sigma^2}(\sum y_{t-1}\epsilon_t - \bar{y}\sum \epsilon_t)}{\frac{1}{T^2\sigma^2}(\sum y_t^2 - T\bar{y}^2)} = \frac{\frac{1}{T\sigma^2}\sum y_{t-1}\epsilon_t - \Big(\frac{\bar{y}}{\sqrt{T}\sigma}\Big)\Big(\frac{1}{\sqrt{T}\sigma}\sum \epsilon_t\Big)}{\frac{1}{T}\sum\Big(\frac{y_t}{\sqrt{T}\sigma}\Big)^2 - \Big(\frac{\bar{y}}{\sqrt{T}\sigma}\Big)^2}$$

$$\longrightarrow \frac{\frac{1}{2}\Big(W(1)^2 - 1\Big) - W(1)\int_0^1 W(r)\mathrm{d}r}{\int_0^1 W(r)^2\mathrm{d}r - \Big(\int_0^1 W(r)\mathrm{d}r\Big)^2}$$

Remember that OLSE of α is given by:

$$\hat{\alpha} = \alpha + \frac{(\sum y_t^2)(\frac{1}{T}\sum \epsilon_t) - \bar{y}\sum y_{t-1}\epsilon_t}{\sum y_t^2 - T\bar{y}^2}$$

Under $\alpha = 0$, $\hat{\alpha}$ is rewritten as follows:

$$\sqrt{T}\hat{\alpha} = \frac{\sigma\Big(\dfrac{1}{T}\sum\Big(\dfrac{y_t}{\sqrt{T}\sigma}\Big)^2\Big)\Big(\dfrac{1}{\sqrt{T}\sigma}\sum\epsilon_t\Big) - \sigma\Big(\dfrac{\bar{y}}{\sqrt{T}\sigma}\Big)\Big(\dfrac{1}{T\sigma^2}\sum y_{t-1}\epsilon_t\Big)}{\dfrac{1}{T}\sum\Big(\dfrac{y_t}{\sqrt{T}\sigma}\Big)^2 - \Big(\dfrac{\bar{y}}{\sqrt{T}\sigma}\Big)^2}$$

$$\longrightarrow \frac{\sigma W(1)\int_0^1 W(r)^2 \mathrm{d}r - \sigma\frac{1}{2}(W(1)^2 - 1)\int_0^1 W(r)\mathrm{d}r}{\int_0^1 W(r)^2 \mathrm{d}r - \Big(\int_0^1 W(r)\mathrm{d}r\Big)^2}$$

Thus, convergence speed of $\hat{\phi}$ is different from that of $\hat{\alpha}$.

Neither $\sqrt{T}\hat{\alpha}$ nor $T(\hat{\phi} - 1)$ are normal.

8.3 Serially Correlated Errors

Consider the case where the error term is serially correlated.

8.3.1 Augmented Dickey-Fuller (ADF) Test

Consider the following AR(p) model:

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

which is rewritten as: $\qquad \phi(L) y_t = \epsilon_t.$

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

Example of ADF Test

```
. gen time=_n

. tsset time
        time variable:  time, 1 to 516
                delta:  1 unit

. gen sexpend=expend-l12.expend
(12 missing values generated)

. corrgram sexpend

                                          -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        0.7177   0.7184   261.14  0.0000          |-----              |-----
2        0.7036   0.3895    512.6  0.0000          |-----              |---
3        0.7031   0.2817   764.23  0.0000          |-----              |--
4        0.6366   0.0456   970.94  0.0000          |-----              |
5        0.6413   0.1116   1181.1  0.0000          |-----              |
6        0.6267   0.0815   1382.2  0.0000          |-----              |
7        0.6208   0.0972     1580  0.0000          |----               |
8        0.6384   0.1286   1789.5  0.0000          |-----              |-
9        0.5926  -0.0205   1970.5  0.0000          |----               |
10       0.5847  -0.0014   2146.9  0.0000          |----               |
11       0.5658  -0.0185   2312.6  0.0000          |----               |
12       0.4529  -0.2570   2418.9  0.0000          |---              --|
13       0.5601   0.2318   2581.8  0.0000          |----               |-
14       0.5393   0.1095   2733.2  0.0000          |----               |
15       0.5277   0.0850   2878.4  0.0000          |----               |
```

```
. varsoc d.sexpend, exo(l.sexpend) maxlag(25)
```

Selection-order criteria
Sample: 39 - 516 Number of obs = 478

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-4917.7				5.1e+07	20.5845	20.5914	20.6019
1	-4878.69	78.013	1	0.000	4.3e+07	20.4255	20.4358	20.4516
2	-4858.95	39.481	1	0.000	4.0e+07	20.3471	20.3608	20.382
3	-4858.46	.97673	1	0.323	4.0e+07	20.3492	20.3664	20.3928
4	-4855.44	6.0461	1	0.014	4.0e+07	20.3407	20.3613	20.3931
5	-4853.84	3.1904	1	0.074	4.0e+07	20.3383	20.3623	20.3993
6	-4851.58	4.5304	1	0.033	4.0e+07	20.333	20.3604	20.4027
7	-4847.61	7.942	1	0.005	3.9e+07	20.3205	20.3514	20.399
8	-4847.51	.20154	1	0.653	3.9e+07	20.3243	20.3586	20.4115
9	-4847.51	.00096	1	0.975	3.9e+07	20.3285	20.3662	20.4244
10	-4847.43	.16024	1	0.689	4.0e+07	20.3323	20.3735	20.437
11	-4831.38	32.094	1	0.000	3.7e+07	20.2694	20.3139	20.3828
12	-4818.46	25.834	1	0.000	3.5e+07	20.2195	20.2675	20.3416*
13	-4815.64	5.6341	1	0.018	3.5e+07	20.2119	20.2633	20.3427
14	-4813.98	3.321	1	0.068	3.5e+07	20.2091	20.264	20.3487
15	-4813.38	1.2007	1	0.273	3.5e+07	20.2108	20.2691	20.3591
16	-4810.57	5.6184	1	0.018	3.5e+07	20.2032	20.265	20.3603
17	-4808.7	3.7539	1	0.053	3.5e+07	20.1996	20.2647	20.3653
18	-4806.12	5.1557	1	0.023	3.4e+07	20.193	20.2616	20.3674
19	-4804.6	3.0319	1	0.082	3.4e+07	20.1908	20.2628	20.374
20	-4804.6	2.7e-05	1	0.996	3.5e+07	20.195	20.2704	20.3869
21	-4797.33	14.542	1	0.000	3.4e+07	20.1688	20.2476	20.3694
22	-4794.2	6.2571*	1	0.012	3.3e+07*	20.1598*	20.2422*	20.3692
23	-4793.42	1.5626	1	0.211	3.3e+07	20.1608	20.2465	20.3788
24	-4792.85	1.1533	1	0.283	3.3e+07	20.1625	20.2517	20.3893

```
  | 25 | -4792.78  .13518    1  0.713  3.4e+07   20.1664    20.259    20.402  |
  +---------------------------------------------------------------------------+
   Endogenous:  D.sexpend
    Exogenous:  L.sexpend  _cons

. dfuller sexpend, lags(22)

Augmented Dickey-Fuller test for unit root         Number of obs   =       481

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.627            -3.442            -2.871            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.4689

. dfuller sexpend, lags(12)

Augmented Dickey-Fuller test for unit root         Number of obs   =       491

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.399            -3.441            -2.870            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.1420
```

$\Longrightarrow$ Unit root is detected.

8.4 Cointegration (共和分)

1. For a scalar y_t, when $\Delta y_t = y_t - y_{t-1}$ is a white noise (i.e., iid), we write $\Delta y_t \sim I(1)$.

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.

   a is called the cointegrating vector.

3. **Example:**

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

$$
\begin{aligned}
y_{1,t} &= \phi_1 y_{2,t} + \epsilon_{1,t}, \\
y_{2,t} &= y_{2,t-1} + \epsilon_{2,t}.
\end{aligned}
$$

   Then,

$$
\Delta y_{1,t} = \phi_1 \epsilon_{2,t} + \epsilon_{1,t} - \epsilon_{1,t-1}, \quad \text{(MA(1) process)},
$$