Time Series Econometrics, 2ST111

Lecture 9. Univariate and Multivariate Processes with Unit Roots

Yukai Yang

Department of Statistics, Uppsala University

Outline of Today's Lecture

- Univariate Processes with Unit Roots (pp.475-543 in Hamilton)
  - Introduction
  - Brownian Motion
  - The Functional CLT
  - The Continuous Mapping Theorem
  - Inference for the Simple Random Walk

Introduction

Consider OLS estimation of the parameter ρ in the simple Gaussian AR(1) process

$$y_t = \rho y_{t-1} + u_t, \quad t = 1, 2, \ldots \tag{1}$$

where $u_t \overset{iid}{\sim} N(0, \sigma^2)$ and the initial value $y_0 = 0$.

The OLS estimator for ρ is given by

$$\hat{\rho}_T = \left(\sum_{t=1}^{T} y_{t-1}^2 \right)^{-1} \left(\sum_{t=1}^{T} y_{t-1} y_t \right) = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{2}$$

and we have

$$\hat{\rho}_T - \rho = \left(\sum_{t=1}^{T} y_{t-1}^2 \right)^{-1} \left(\sum_{t=1}^{T} y_{t-1} u_t \right) = \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{3}$$

Introduction

If $|\rho| < 1$, then definitely

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{d} N(0, 1 - \rho) \tag{4}$$

What if $\rho = 1$? that is,

$$(1 - L)y_t = u_t \tag{5}$$

Check (4) and you will find immediately that the variance is $1 - \rho = 0$. The limiting distribution is degenerate and collapses to a point mass.

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{p} 0 \tag{6}$$

Introduction

To obtain a non-degenerate limiting distribution for $\hat{\rho}_T$ in the unit root case, it turns out that we have to multiply (or scale) $\hat{\rho}_T$ by T rather than $\sqrt{T}$.

To get a better sense of why scaling by T is necessary when $\rho = 1$, note that $T(\hat{\rho}_T - 1)$ can be written as

$$T(\hat{\rho}_T - 1) = T \times \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{T^{-1} \sum_{t=1}^{T} y_{t-1} u_t}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2} \tag{7}$$

Introduction

First, consider the numerator of (7). It can be shown that

$$\frac{1}{\sigma^2} \times \frac{1}{T} \sum_{t=1}^{T} y_{t-1} u_t \xrightarrow{d} \frac{1}{2}(X - 1), \tag{8}$$

where $X \sim \chi^2(1)$.

Second, consider the denominator of (7):

$$T^{-2} \sum_{t=1}^{T} y_{t-1}^2. \tag{9}$$

Let us consider the expectation $\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right)$. Since

$$y_{t-1} \sim N(0, \sigma^2(t-1)),$$

we get

$$\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) = \sum_{t=1}^{T} \mathrm{E}(y_{t-1}^2) = \sigma^2 \sum_{t=1}^{T}(t-1) = \frac{\sigma^2(T-1)T}{2} \tag{10}$$

Now we see that, if we divide $\sum_{t=1}^{T} y_{t-1}^2$ by T^2, its expectation will converge to $\sigma^2/2$ without T.

Introduction

Summary:

- If $\rho = 1$ (random walk process), $\hat{\rho}_T - 1$ should be multiplied by T instead of $\sqrt{T}$ to obtain a non-degenerate limiting distribution.
- This limiting distribution is not the usual Gaussian distribution. It is a ratio involving a $\chi^2(1)$ distribution in the numerator and another nonstandard distribution in the denominator.
- We would like to describe this limiting distribution. This can be done in terms of functionals of Brownian motion.

Brownian Motion

Suppose that

$$y_t = y_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots \tag{11}$$

where $\varepsilon_t \overset{iid}{\sim} N(0,1)$ and $y_0 = 0$.

Then if we expand all the $y_{t-1} \ldots$, we get $y_t = y_0 + \varepsilon_1 + \ldots + \varepsilon_t \sim N(0,t)$.

Letting $s > t$, we have

$$\begin{aligned} y_s - y_t &= (y_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t + \varepsilon_{t+1} + \ldots + \varepsilon_s) - (y_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \\ &= \varepsilon_{t+1} + \ldots + \varepsilon_s \end{aligned}$$

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

For any $t < s < r < q$, the two random variables $y_s - y_t$ and $y_q - y_r$ are independent.

Brownian Motion

In particular, consider the change between y_{t-1} and y_t

$$y_t - y_{t-1} = \varepsilon_t$$

Suppose that we view ε_t as the sum of two independent Gaussian random variables

$$\varepsilon_t = e_{1t} + e_{2t}$$

where $e_{it} \overset{iid}{\sim} N(0, 1/2)$.

We might associate e_{1t} with the change between y_{t-1} and $y_{t-1/2}$

$$y_{t-1/2} - y_{t-1} = e_{1t}, \tag{12}$$

and e_{2t} with the change between $y_{t-1/2}$ and y_t

$$y_t - y_{t-1/2} = e_{2t}. \tag{13}$$

Note that (12) added to (13) implies (11)

$$y_t - y_{t-1} = e_{1t} + e_{2t},$$

where $e_{1t} + e_{2t} \overset{iid}{\sim} N(0,1)$.

That is, sampled at $t = 1, 2, \ldots$ the stochastic process defined by (12) and (13) are equivalent to (11) except the frequency.

In addition, (12) and (13) describe a stochastic process defined not only for $t = 1, 2, \ldots$ but also for $t = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$

For both integer and noninteger values of s and t, $(s > t)$

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

$y_s - y_t$ and $y_q - y_r$ are independent for any $t < s < r < q$.

Brownian Motion

Similarly, we could view the change between y_{t-1} and y_t

$$y_t - y_{t-1} = \varepsilon_t$$

as the sum of n independent Gaussian variables

$$\varepsilon_t = e_{1t} + e_{2t} + \ldots + e_{nt},$$

where $e_{it} \overset{iid}{\sim} N(0, 1/n)$ and

$$\begin{aligned}
y_{t-(n-1)/n} - y_{t-1} &= e_{1t} \\
y_{t-(n-2)/n} - y_{t-(n-1)/n} &= e_{2t} \\
&\vdots \\
y_{t-1/n} - y_{t-2/n} &= e_{(n-1)t} \\
y_t - y_{t-1/n} &= e_{nt}
\end{aligned}$$

Remarks:

- The result would be a stochastic process with the same properties as the process in (11), defined at a finer and finer grid of time points as we increase the value of n.
- The limit as n tends to infinity is a continuous-time stochastic process known as standard Brownian motion. The value of this process at time t is denoted by $W(t)$.
- Brownian motion is named after the biologist Robert Brown whose research dates to the 1820s. A standard Brownian motion is also known as a Wiener process.
- A continuous-time process is a random variable that takes on a value for any nonnegative real number t, this is in contrast to a discrete-time process, which is only defined for integer values of t.

Brownian Motion

Definition (Standard Brownian Motion)

Standard Brownian motion $W(\cdot)$ is a continuous time process, associating each $0 \leq t \leq 1$ with the scalar random variable $W(t)$ such that:

1. $W(0) = 0$
2. For any $0 \leq t_1 < t_2 < ... < t_k \leq 1$ then random variables

$$W(t_2) - W(t_1),\ W(t_3) - W(t_2),\ ...,\ W(t_k) - W(t_{k-1})$$

   are independent and jointly multivariate Gaussian distributed, with

$$W(s) - W(t) \sim N(0, s - t)$$

   for any $0 \leq t < s \leq 1$.
3. $W(t)$ has continuous sample paths.

Brownian Motion

Remarks:

- Though $W(t)$ has continuous sample paths, it can be shown that its sample paths are nowhere differentiable.
- Other continuous time processes can be generated from standard Brownian motion.
- Since, by definition, $W(t) \sim N(0, t)$, it follows that

$$W_\sigma(t) = \sigma W(t) \sim N(0, \sigma^2 t).$$

- Similarly, it is readily seen that $W^2(t)$ is $t \times \chi^2(1)$ distributed. In particular, $W^2(1) \sim \chi^2(1)$.

Functional Central Limit Theorem

One of the uses of Brownian motion is to allow for more general statements of the CLT.

Recall the classical CLT.

Theorem

Let $\bar{y}_T = T^{-1} \sum_{t=1}^{T} y_t$, where $y_1, ..., y_T$ is a sequence of i.i.d. random variables with finite mean μ and variance σ^2. Then

$$\sqrt{T}(\bar{y}_T - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Functional Central Limit Theorem

Suppose that $u_1, ..., u_T$ is an *i.i.d.* sequence with mean zero and variance σ^2, and consider the estimator

$$\bar{u}_{\lfloor T/2 \rfloor} = \frac{1}{\lfloor T/2 \rfloor} \sum_{t=1}^{\lfloor T/2 \rfloor} u_t, \tag{14}$$

where $\lfloor \cdot \rfloor$ denotes the floor function (integer part for positive numbers).

Given a sample of size T, for an even T, this estimator uses only the first half of the sample and discards the other half.

Clearly, this estimator also satisfies the classical CLT

$$\sqrt{\lfloor T/2 \rfloor} \times \bar{u}_{\lfloor T/2 \rfloor} \xrightarrow{d} N(0, \sigma^2), \quad \text{as } T \to \infty. \tag{15}$$

Moreover, $\bar{u}_{\lfloor T/2 \rfloor}$ would be independent of an estimator that uses only the second half of the sample.

Functional Central Limit Theorem

More generally, we can construct a random variable $X_T(r)$ that uses only the first rth fraction of the sample $u_1, ..., u_T$

$$X_T(r) = \frac{1}{T} \sum_{t=1}^{\lfloor rT \rfloor} u_t, \tag{16}$$

where $0 \leq r \leq 1$.

Thus, by construction

$$X_T(r) = \begin{cases} 0 & \text{for} \quad 0 \leq r < 1/T \\ u_1/T & \text{for} \quad 1/T \leq r < 2/T \\ (u_1 + u_2)/T & \text{for} \quad 2/T \leq r < 3/T \\ \vdots & \\ (u_1 + u_2 + ... + u_n)/T & \text{for} \quad r = 1 \end{cases}$$

Functional Central Limit Theorem

For $r > 0$, it can be shown that

$$\sqrt{T}\, X_T(r) \xrightarrow{d} N(0, r\sigma^2).$$

Hence

$$\frac{1}{\sigma}\sqrt{T}\, X_T(r) \xrightarrow{d} N(0, r).$$

This implies that

$$\frac{1}{\sigma}\sqrt{T}\, [X_T(r_2) - X_T(r_1)] \xrightarrow{d} N(0, r_2 - r_1)$$

for any $0 \le r_1 \le r_2 \le 1$.

In addition, note that the random variable

$$X_T(r_2) - X_T(r_1) = \frac{1}{T}\sum_{t=1}^{\lfloor r_2 T\rfloor} u_t - \frac{1}{T}\sum_{t=1}^{\lfloor r_1 T\rfloor} u_t = \frac{1}{T}\sum_{t=\lfloor r_1 T\rfloor+1}^{\lfloor r_2 T\rfloor} u_t$$

is independent of $X_T(r)$ for any $0 \leq r \leq r_1$.

Functional Central Limit Theorem

So it should not be surprising that

$$\frac{1}{\sigma}\sqrt{T}X_T(\cdot) \xrightarrow{d} W(\cdot) \tag{17}$$

for $0 \leq r \leq 1$. This is the functional central limit theorem.

For example, when the functions in (17) are evaluated at $r = 1$, we have

$$\frac{1}{\sigma}\sqrt{T}X_T(1) \xrightarrow{d} W(1)$$

where $X_T(1) = T^{-1}\sum_{t=1}^{T} u_t$ and $W(1)$ is the standard normal distribution.

The classical CLT is a special case of the functional CLT.

Functional Central Limit Theorem

Remarks:

- The expression $X_T(\cdot)$ denotes a function, while $X_T(r)$ denotes the value that function assumes at time r (a random variable).
- In previous lectures, we defined the convergence in distribution for (a sequence of) random variables. Now this definition can be extended to (a sequence of) random functions.

FCLT Again

- Suppose that ε_t is *i.i.d* $(0, \sigma^2)$, but not necessarily normally distributed!
- The functional central limit theorem (FCLT) (or the invariance principle) tells us that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{s} \varepsilon_t \quad \xrightarrow{d} \quad \sigma W(r) \sim N(0, r\sigma^2)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \quad \xrightarrow{d} \quad \sigma W(1) \sim N(0, \sigma^2)$$

as both T and s go to infinity and $s/T \to r$. Remember this!

Continuous Mapping Theorem

Proposition 7.3 (c) on pp.184 in Hamilton says that

- if the sequence of random variables $\{x_t\}_{t=1}^{\infty}$ converges in distribution to x, i.e., $x_t \xrightarrow{d} x$, and
- if the function $g : \mathbb{R} \mapsto \mathbb{R}$ is a continuous function, then

$$g(x_t) \xrightarrow{d} g(x)$$

A similar result holds for a sequence of random functions. The analog to the function $g(\cdot)$ is a continuous functional.

Continuous Mapping Theorem

The continuous mapping theorem states that if

$$S_T(\cdot) \xrightarrow{d} S(\cdot)$$

and $g(\cdot)$ is continuous functional, then

$$g(S_T(\cdot)) \xrightarrow{d} g(S(\cdot)) \tag{18}$$

For example,

$$\sqrt{T}X_T(\cdot) \xrightarrow{d} \sigma W(\cdot) \quad \text{and} \quad [\sqrt{T}X_T(\cdot)]^2 \xrightarrow{d} \sigma^2[W(\cdot)]^2$$

and even

$$\int_0^1 \sqrt{T}X_T(x)\mathrm{d}x \xrightarrow{d} \int_0^1 \sigma W(x)\mathrm{d}x \quad \text{and...}$$

Consider the simple random walk

$$y_t = y_{t-1} + u_t, \quad t = 1, 2, \ldots$$

where $u_1, \ldots, u_T$ is an *i.i.d.* with mean zero and variance σ^2 and $y_0 = 0$.

By recursion,

$$y_t = u_1 + u_2 + \ldots + u_t$$

Note that

$$X_T(r) = \begin{cases} 0 = y_0/T & \text{for} \quad 0 \leq r < 1/T \\ u_1/T = y_1/T & \text{for} \quad 1/T \leq r < 2/T \\ (u_1 + u_2)/T = y_2/T & \text{for} \quad 2/T \leq r < 3/T \\ \vdots & \\ (u_1 + u_2 + \ldots + u_n)/T = y_T/T & \text{for} \quad r = 1 \end{cases}$$

Inference for the Simple Random Walk

Please see Figure 17.1 on pp.484 in Hamilton!

A simple geometrical argument shows that

$$\int_0^1 X_T(r)\mathrm{d}r = \frac{y_1}{T^2} + \frac{y_2}{T^2} + \ldots + \frac{y_{T-1}}{T^2}$$

or

$$\int_0^1 \sqrt{T} X_T(r)\mathrm{d}r = T^{-3/2} \sum_{t=1}^{T} y_{t-1}$$

Inference for the Simple Random Walk

Recall that

$$\sqrt{T}X_T(\cdot) \xrightarrow{d} \sigma W(\cdot).$$

Therefore, by the continuous mapping theorem

$$\int_0^1 \sqrt{T}X_T(r)\mathrm{d}r \xrightarrow{d} \int_0^1 \sigma W(r)\mathrm{d}r$$

which implies that

$$T^{-3/2}\sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \sigma \int_0^1 W(r)\mathrm{d}r,$$

as the sample size T tends to infinity.

Inference for the Simple Random Walk

A similar approach can be used to describe the limiting distribution of

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

Let

$$S_T(r) = T[X_T(r)]^2$$

Then $S_T(r)$ can be written as

$$S_T(r) = \begin{cases} 0 & \text{for} \quad 0 \le r < 1/T \\ y_1^2/T & \text{for} \quad 1/T \le r < 2/T \\ y_2^2/T & \text{for} \quad 2/T \le r < 3/T \\ \vdots & \\ y_T^2/T & \text{for} \quad r = 1 \end{cases}$$

A simple geometrical argument shows that

$$\int_0^1 S_T(r)\mathrm{d}r = \frac{y_1^2}{T^2} + \frac{y_2^2}{T^2} + \ldots + \frac{y_{T-1}^2}{T^2}$$

or equivalently

$$\int_0^1 S_T(r)\mathrm{d}r = \frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2$$

Inference for the Simple Random Walk

Recall that

$$S_T(\cdot) \xrightarrow{d} \sigma^2[W(\cdot)]^2.$$

Therefore, by the continuous mapping theorem

$$\int_0^1 S_T(r)\mathrm{d}r \xrightarrow{d} \int_0^1 \sigma^2[W(r)]^2\mathrm{d}r$$

which implies that

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 W(r)^2\mathrm{d}r, \tag{19}$$

as the sample size T tends to infinity.

Inference for the Simple Random Walk

Consider now the limiting distribution of the test statistic

$$T(\hat{\rho}_T - 1) = \frac{T^{-1}\sum_{t=1}^{T} y_{t-1}u_t}{T^{-2}\sum_{t=1}^{T} y_{t-1}^2}. \tag{20}$$

In (7) and (8), we see that the numerator of (20) converges in distribution to $\frac{\sigma^2}{2}(X-1)$, where $X \sim \chi^2(1)$ as $T \to \infty$.

Recall that the random variable $W(1)^2$ is $\chi^2(1)$ distributed.

Hence, another way to describe the limiting distribution of the numerator of (20) is using a functional of Brownian motion

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} \frac{\sigma^2}{2}\left(W(1)^2 - 1\right) \tag{21}$$

Inference for the Simple Random Walk

Since (20) is a continuous function of the LHSs of (21) and (19), it follows that, under the null hypothesis that $\rho = 1$, the OLS estimator $\hat{\rho}_T$ is characterized by

$$T(\hat{\rho}_T - 1) \xrightarrow{d} \frac{\frac{1}{2}\left(W(1)^2 - 1\right)}{\int_0^1 W(r)^2 \mathrm{d}r} \tag{22}$$

Remark:
In practice, critical values for the test statistic in (22) are found by calculating the exact finite-sample distribution of $T(\hat{\rho}_T - 1)$ for a given sample size T, under the assumption that u_t are Gaussian distributed. This can be doen either by Monte Carlo simulation, or by using exact numerical procedures. Read pp.488 in Hamilton.

You can use the results in Proposition 17.1 on pp.486 in Hamilton. Note that $\xi_t = y_t$.

Dickey-Fuller Tests

- Now consider the somewhat general model

$$y_t = \rho y_{t-1} + \alpha + \delta t + \varepsilon_t \tag{23}$$

where ε_t is *i.i.d.* with zero mean and finite variance σ^2.

- We are interested in whether $\rho = 1$ (unit root), and we test it based on the observations y_t.
- Dickey-Fuller tests are several unit root tests for different situations (different assumptions), but they all assume that there is not autocorrelation in the errors ε_t.

Dickey-Fuller Test for Case 1

- The regression model
$$y_t = \rho y_{t-1} + \varepsilon_t \tag{24}$$
- Assumptions: $\alpha = 0$ and $\delta = 0$
- Null hypothesis $H_0 : \rho = 1$
- The alternative $H_1 : |\rho| < 1$
- The test has been given in (22)
- There are two versions for the test, ρ version in (22), and t-ratio version
$$t_T \xrightarrow{d} \frac{\frac{1}{2}\left(W(1)^2 - 1\right)}{\sqrt{\int_0^1 W(r)^2 \mathrm{d}r}} \tag{25}$$

Dickey-Fuller Test for Case 2

- The regression model

$$y_t = \rho y_{t-1} + \alpha + \varepsilon_t \tag{26}$$

- Assumptions: $\delta = 0$
- Null hypothesis $H_0 : \rho = 1$ and $\alpha = 0$
- The alternative $H_1 : |\rho| < 1$ or $\alpha \neq 0$
- The joint test for the null hypothesis is in [17.4.25] on pp.492.
- The tests for $\rho = 1$ are given in [17.4.28] on pp.492 (ρ) and [17.4.36] on pp.494 (t-ratio).
- If the null is true, the model is simply a random walk.

Dickey-Fuller Test for Case 3

- The regression model

$$y_t = \rho y_{t-1} + \alpha + \varepsilon_t \tag{27}$$

- Assumptions: $\delta = 0$ and $\alpha \neq 0$
- Null hypothesis $H_0 : \rho = 1$
- The alternative $H_1 : |\rho| < 1$
- The test is given in [17.4.46] on pp.492 (ρ). Note that it is the marginal distribution of $\hat{\rho}_T$
- Both $\hat{\alpha}_T$ and $\hat{\rho}_T$ converge to Gaussian, but with different rates of convergence.
- If the null is true, the model is $y_t = y_0 + \alpha t + \sum_{s=1}^{t} \varepsilon_s$. Random walk with drift αt.
- We see that, from cases 2 and 3, the asymptotic distributions of ρ are different based on different beliefs about the true value of α.

Dickey-Fuller Test for Case 4

- The regression model

$$y_t = \rho y_{t-1} + \alpha + \delta t + \varepsilon_t \tag{28}$$

- Assumptions: α can be anything
- Null hypothesis $H_0 : \rho = 1$, $\delta = 0$ and $\alpha = \alpha_0$
- The alternative $H_1 : |\rho| < 1$ or $\delta \neq 0$ or $\alpha \neq \alpha_0$
- The model can be reparameterized as follows

$$y_t = \alpha^* + \rho^* \xi_{t-1} + \delta^* t + \varepsilon_t \tag{29}$$

  where $\alpha^* = (1-\rho)\alpha$, $\rho^* = \rho$, $\delta^* = \delta + \rho\alpha$ and $\xi_t = y_{t-1} - \alpha(t-1)$. Moreover, $\xi_t = y_0 + \varepsilon_1 + \varepsilon_2 + ...\varepsilon_t$.
- The new and equivalent null hypothesis is $H_0 : \rho^* = 1$, $\alpha^* = 0$ and $\delta^* = \alpha_0$.
- The joint test is given in [17.4.53] on pp.499 (ρ). The t-ration of ρ is given in [17.4.55].

Remarks

If $|\rho| < 1$,

- $\alpha \neq 0$ is simply an intercept
- $\delta \neq 0$ is a linear trend.

If $\rho = 1$,

- $\alpha \neq 0$ will become a drift term (linear trend)
- $\delta \neq 0$ will become a quadratic trend.

In practice, you may assume that either $|\rho| < 1$ or $\rho = 1$,

- take 1st and 2nd order differences, see whether they look stationary
- if the 2nd order difference shows strong stationarity, you may skip case 4.
- you have to plot the data, see whether there is a clear linear trend; the linear trend may come from either ($|\rho| < 1$ and $\delta \neq 0$) or ($\rho = 1$ and $\alpha \neq 0$).
- choose one or several DF tests and analyse.

Serial correlation

- In all cases, it was assumed that the error term is independent (hence not serially correlated). We can test for serial correlation by using, for example, the Breusch-Godfrey autocorrelation test.
- If we find serial correlation, we should take it into account. An easy strategy for this is to respecify the estimation equation by adding lagged first differences. The corresponding unit root tests are called augmented Dickey-Fuller (ADF) tests.
- Another strategy is to estimate the autocovariances (nuisance parameters) γ_i of the errors and construct new tests similar to the DF tests. The resulting unit root tests are called Phillips-Perron tests but their finite-sample performance are poor in contrast to the ADF tests.

To be continued! Thank you!