Time Series Analysis

Lecture 6

Review

1. The Autocovariance Generating Function
2. Sums of ARMA processes
3. Wold's Decomposition and the Box-Jenkins Modeling Philosophy
4. Maximum Likelihood Estimation

Today's Topics

1. Review of AR(1) model
2. Brownian motion and Functional central limit theorem
3. Asymptotic properties of unit root processes and tests for unit root
4. Generalization to processes with serial correlation

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation

Gaussian AR(1) process

Consider a Gaussian AR(1) process,

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

where $u_t \sim i.i.d.N(0, \sigma^2)$, and $y_0 = 0$.

- if $|\rho| < 1$, y_t is called a stationary process;
- if $|\rho| = 1$, y_t is called the Unit Root (random walk) process;
- if $|\rho| > 1$, y_t is called an explosive(unstable) process.

The OLS estimate of ρ is given by

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}. \tag{2}$$

Asymptotics for stationary AR(1)

We can show that if $|\rho| < 1$, i.e., y_t is stationary, then

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{L} N(0, (1-\rho^2)). \tag{3}$$

To see this, note that

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

But by the Ergodic theorem, we have

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} E[y_{t-1}^2] = \sigma^2/(1-\rho^2).$$

By the CLT for MDS, we obtain

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1} u_t \xrightarrow{L} N(0, Var(y_{t-1} u_t)),$$

where $Var(y_{t-1}u_t) = \sigma^2 E[y_{t-1}^2] = \sigma^4/(1-\rho^2)$. Therefore, an application of the Slutsky Theorem leads to (3).

Asymptotics for Unit Root AR(1)

If (3) were also valid for $\rho = 1$, it would seem to claim that

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{p} 0. \tag{4}$$

This is indeed true for unit root processes, but it obviously is not very helpful for hypothesis tests.

Q: How to obtain a nondegenerate asymptotic distribution for $\hat{\rho}_T$ in the unit root case?

It turns out that we need to look at

$$T(\hat{\rho}_T - 1) = \frac{(1/T)\sum_{t=1}^{T} y_{t-1}u_t}{(1/T^2)\sum_{t=1}^{T} y_{t-1}^2}. \tag{5}$$

What is the asymptotic distribution of $T(\hat{\rho}_T - 1)$?

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation

Today's Topics

1. Rewiew of AR(1)
2. Brownian Motion and Functional CLT
   - Brownian Motion
   - Functional CLT
3. Asymptotic properties of Unit Root processes and tests for Unit Root
4. Generalization to processes with serial correlation

Random walk

Consider a random walk,

$$y_t = y_{t-1} + \varepsilon_t, \varepsilon_t \sim i.i.d.N(0,1). \tag{6}$$

If $y_0 = 0$, then it follows that

$$\begin{aligned} y_t &= \varepsilon_1 + \ldots + \varepsilon_t, \\ y_t &\sim N(0,t), \end{aligned}$$

and, for $s < t$,

$$\begin{aligned} y_s - y_t &= \varepsilon_{t+1} + \ldots + \varepsilon_s, \\ y_s - y_t &\sim N(0, s-t), \end{aligned}$$

which is independent with $y_q - y_r$, where $t < s < r < q$.

Suppose we view ε_t as:

$$\varepsilon_t = e_{1t} + e_{2t}, \tag{7}$$

with $e_{it} \overset{i.i.d}{\sim} N(0, 1/2)$. We might define some interim point $y_{t-(1/2)}$, s.t.,

$$\begin{aligned} y_{t-(1/2)} - y_{t-1} &= e_{1t}, \\ y_t - y_{t-(1/2)} &= e_{2t}. \end{aligned}$$

Similarly, we could imagine partitioning the change between $t-1$ and t into N separate subperiods:

$$y_t - y_{t-1} = e_{1t} + e_{2t} + \cdots + e_{Nt}, \tag{8}$$

with $e_{it} \overset{i.i.d}{\sim} N(0, 1/N)$.
The result would be a process with all the same properties as (6), defined at a finer and finer grid of dates as we increase N.

The limit as $N \to \infty$ is a continuous-time process known as standard Brownian motion. The value of this process at time t is denoted as $W(t)$.

Brownian motion

Definition: Standard Brownian motion $W(\cdot)$ is a continuous-time stochastic process, associating each date $t \in [0, 1]$ with the scalar $W(t)$ such that:

(a) $W(0) = 0$;

(b) For any dates $0 \le t_1 < t_2 < \cdots < t_k \le 1$, the changes $[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \cdots, [W(t_k) - W(t_{k-1})]$ are independent multivariate Gaussian with $[W(s) - W(t)] \sim N(0, s - t)$;

(c) For any given realization, $W(t)$ is continuous in t with probability 1.

- Other continuous-time processes can be generated from standard Brownian motion. For example, the process

$$Z(t) = \sigma \cdot W(t)$$

has independent increment and is distributed $N(0, \sigma^2 t)$ across realizations. Such a process is described as Brownian motion with variance σ^2.

- As another example,

$$Z(t) = [W(t)]^2$$

would be distributed as t times a $\chi^2(1)$ variable.

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT
   - Brownian Motion
   - Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation

Review of conventional CLT

If $u_t \sim i.i.d.$ with zero mean and variance σ^2, then the sample mean $\bar{u}_T = (1/T)\sum_{t=1}^{T} u_t$ satisfies

$$\sqrt{T}\bar{u}_T \xrightarrow{L} N(0, \sigma^2). \tag{9}$$

Consider now an estimator based on only the first half of the sample,

$$\bar{u}_{[T/2]^*} = (1/[T/2]^*) \sum_{t=1}^{[T/2]^*} u_t.$$

Here $[T/2]^*$ denotes the largest integer that is less than or equal to $T/2$. This strange estimator would also satisfy the CLT:

$$\sqrt{[T/2]^*}\bar{u}_{[T/2]^*} \xrightarrow[T\to\infty]{L} N(0, \sigma^2).$$

More generally, we can construct a variable $X_T(r)$ from the sample mean of the first rth fraction of observation, $r \in [0, 1]$, defined by

$$X_T(r) = (1/T) \sum_{t=1}^{[Tr]^*} u_t. \tag{10}$$

For any given realization, $X_T(r)$ is a step function in r, with

$$X_T(r) = \begin{cases} 0 & \text{for } 0 \leq r < 1/T \\ u_1/T & \text{for } 1/T \leq r < 2/T \\ \vdots & \\ (u_1 + u_2 + \cdots + u_T)/T & \text{for } r = 1. \end{cases} \tag{11}$$

Then

$$\sqrt{T}\cdot X_T(r) = (1/\sqrt{T})\sum_{t=1}^{[Tr]^*} u_t = (\sqrt{[Tr]^*}/\sqrt{T})(1/\sqrt{[Tr]^*})\sum_{t=1}^{[Tr]^*} u_t.$$

But

$$(1/\sqrt{[Tr]^*})\sum_{t=1}^{[Tr]^*} u_t \xrightarrow{L} N(0,\sigma^2),$$

by the CLT, while $(\sqrt{[Tr]^*}/\sqrt{T}) \to \sqrt{r}$. Hence,

$$\sqrt{T}\cdot X_T(r) \xrightarrow{L} \sqrt{r}N(0,\sigma^2) = N(0,r\sigma^2),$$

and

$$\sqrt{T}\cdot [X_T(r)/\sigma] \xrightarrow{L} N(0,r).$$

Similarly, for $r_2 > r_1$, we have

$$\sqrt{T} \cdot [X_T(r_2) - X_T(r_1)]/\sigma \xrightarrow{L} N(0, r_2 - r_1).$$

The stochastic functions $\{\sqrt{T} \cdot [X_T(\cdot)/\sigma]\}_{T=1}^{\infty}$ has an asymptotic probability law that is described by standard Brownian motion $W(\cdot)$:

$$\sqrt{T} \cdot [X_T(\cdot)/\sigma] \xrightarrow{L} W(\cdot). \tag{12}$$

Result (12) is known as the functional CLT.

Note that $X_T(\cdot)$ is a random function while $X_T(r)$ is a random variable. The conventional CLT is a special case of functional CLT:

$$\sqrt{T} X_T(1)/\sigma = [1/(\sigma\sqrt{T})] \sum_{t=1}^{T} u_t \xrightarrow{L} W(1) \sim N(0, 1).$$

Continuous Mapping Theorem

- CMT: if $\{x_T\}_{T=1}^{\infty}$ is a sequence of random variables with $x_T \xrightarrow{L} x$ and if $g : \mathbb{R}^1 \to \mathbb{R}^1$ is a continuous function, then $g(x_T) \xrightarrow{L} g(x)$. A similar result holds for sequence of random functions.
- If $\{X_T(\cdot)\}_{T=1}^{\infty}$ is a sequence of stochastic functions with

$$X_T(\cdot) \xrightarrow{L} X(\cdot)$$

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

$$g(X_T(\cdot)) \xrightarrow{L} g(X(\cdot))$$

.
- **Example:** Since $\sqrt{T} \cdot X_T(\cdot) \xrightarrow{L} \sigma \cdot W(\cdot)$, define $S_T(r) = [\sqrt{T} \cdot X_T(r)]^2$, it follows that

$$S_T(\cdot) \xrightarrow{L} \sigma^2[W(\cdot)]^2. \tag{13}$$

Applications to Unit Root Processes

Example 1

Consider

$$y_t = y_{t-1} + u_t, \tag{14}$$

where $\{u_t\}$ is an i.i.d. sequence with mean zero and variance σ^2. If $y_0 = 0$, then

$$y_t = u_1 + u_2 + \cdots + u_t. \tag{15}$$

This can be used to express the stochastic function $X_T(r)$ defined in (11) as

$$X_T(r) = \begin{cases} 0 & \textit{for } 0 \le r < 1/T \\ y_1/T & \textit{for } 1/T \le r < 2/T \\ y_2/T & \textit{for } 2/T \le r < 3/T \\ \vdots & \\ y_T/T & \textit{for } r = 1. \end{cases} \tag{16}$$

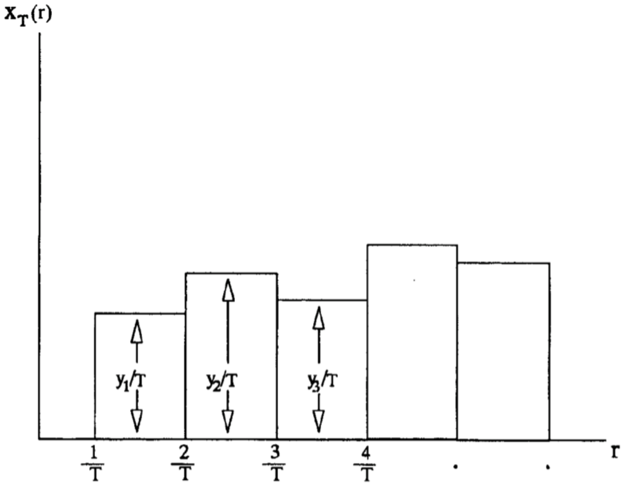

FIGURE 17.1 Plot of $X_T(r)$ as a function of r.

The integral of $X_T(r)$ is thus equivalent to

$$\int_0^1 X_T(r)dr = y_1/T^2 + y_2/T^2 + \cdots + y_{T-1}/T^2. \qquad (17)$$

Multiplying both sides of (17) by $\sqrt{T}$ establishes that

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

We know from the FCLT and the continuous mapping theorem that as $T \to \infty$,

$$\int_0^1 \sqrt{T} \cdot X_T(r)dr \xrightarrow{L} \sigma \cdot \int_0^1 W(r)dr,$$

then we get

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{L} \sigma \cdot \int_0^1 W(r)dr. \tag{18}$$

Example 2

Further consider

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

which can be written as

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

It follows that

$$\int_0^1 S_T(r)dr = y_1^2/T^2 + y_2^2/T^2 + \cdots + y_{T-1}^2/T^2.$$

Thus from the FCLT and the CMT,

$$T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 [W(r)]^2 dr. \tag{19}$$

Example 3

Note that for a random walk,

$$y_t^2 = (y_{t-1} + u_t)^2 = y_{t-1}^2 + 2y_{t-1}u_t + u_t^2,$$

implying that

$$y_{t-1}u_t = (1/2)\{y_t^2 - y_{t-1}^2 - u_t^2\}.$$

Then

$$\sum_{t=1}^{T} y_{t-1}u_t = (1/2)\{y_T^2 - y_0^2\} - (1/2)\sum_{t=1}^{T} u_t^2.$$

Recalling that $y_0 = 0$, then

$$\begin{aligned}
(1/T)\sum_{t=1}^{T} y_{t-1}u_t &= (1/2) \cdot y_T^2/T - (1/2) \cdot \sum_{t=1}^{T} u_t^2/T \\
&= (1/2)S_T(1) - (1/2) \cdot \sum_{t=1}^{T} u_t^2/T \\
&\xrightarrow{L} (1/2)\sigma^2[W(1)]^2 - (1/2)\sigma^2.
\end{aligned}$$

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root
   - Asymptotics for unit root processes
   - Dickey-Fuller tests

4. Generalization to processes with serial correlation

Asymptotics for unit root processes

Proposition 1 Suppose that y_t follows a random walk without drift,

$$y_t = y_{t-1} + u_t,$$

where $y_0 = 1$ and $\{u_t\}$ is an $i.i.d.$ sequence with mean zero and variance σ^2. Then

(a) $T^{-1/2} \sum_{t=1}^{T} u_t \xrightarrow{L} \sigma \cdot W(1)$;

(b) $T^{-1} \sum_{t=1}^{T} y_{t-1} u_t \xrightarrow{L} (1/2)\sigma^2\{[W(1)]^2 - 1\}$; **(Example 3)**

(c) $T^{-3/2} \sum_{t=1}^{T} t u_t \xrightarrow{L} \sigma \cdot W(1) - \sigma \cdot \int_0^1 W(r)dr$;

(d) $T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{L} \sigma \cdot \int_0^1 W(r)dr$; **(Example 1)**

(e) $T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 [W(r)]^2 dr$; **(Example 2)**

(f) $T^{-5/2} \sum_{t=1}^{T} t y_{t-1} \xrightarrow{L} \sigma \cdot \int_0^1 rW(r)dr$;

(g) $T^{-3} \sum_{t=1}^{T} t y_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 r \cdot [W(r)]^2 dr$;

(h) $T^{(\nu+1)} \sum_{t=1}^{T} t^{\nu} \rightarrow 1/(\nu + 1)$ for $\nu = 0, 1, \cdots$.

The asymptotics of $T(\hat{\rho}_T - 1)$

Now we can apply Proposition 1 (b) and (e) to show that

$$\begin{aligned} T(\hat{\rho}_T - 1) &= \frac{(1/T)\sum_{t=1}^{T} y_{t-1}u_t}{(1/T^2)\sum_{t=1}^{T} y_{t-1}^2} \\ &\xrightarrow{L} \frac{(1/2)\{[W(1)]^2 - 1\}}{\int_0^1 [W(r)]^2 dr}. \end{aligned}$$

This answers the question raised at the beginning of this lecture. It is worth noting that such a limiting result involves:

(i) a convergence rate T (super-consistent), and

(ii) a nonstandard limiting distribution.

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root
    - Asymptotics for unit root processes
    - Dickey-Fuller tests

4. Generalization to processes with serial correlation

Dickey-Fuller tests of Unit Root

We shall consider four cases in sequence,

Case 1 Estimated regression: $y_t = \rho y_{t-1} + u_t$
True process: $y_t = y_{t-1} + u_t \ u_t \sim i.i.d.N(0, \sigma^2)$

Case 2 Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
True process: $y_t = y_{t-1} + u_t \ u_t \sim i.i.d.N(0, \sigma^2)$

Case 3 Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
True process: $y_t = \alpha + y_{t-1} + u_t \ \alpha \neq 0, u_t \sim i.i.d.(0, \sigma^2)$

Case 4 Estimated regression: $y_t = \alpha + \delta t + \rho y_{t-1} + u_t$
True process: $y_t = \alpha + y_{t-1} + u_t \ \alpha$ any, $u_t \sim i.i.d.N(0, \sigma^2)$

Dickey-Fuller tests of Unit Root: Case 1

We first consider Case 1

Case 1 Estimated regression: $y_t = \rho y_{t-1} + u_t$

True process: $y_t = y_{t-1} + u_t \; u_t \sim i.i.d.N(0, \sigma^2)$

Dickey-Fuller ρ test for Case 1

Under the null hypothesis that $\rho = 1$, the Dickey-Fuller ρ statistic

$$T(\hat{\rho}_T - 1) \xrightarrow{L} \frac{(1/2)\{[W(1)]^2 - 1\}}{\int_0^1 [W(r)]^2 dr} \overset{\triangle}{=\!=} DF_{\rho,case1}. \qquad (20)$$

Sample size T	Probability that $T(\hat{\rho} - 1)$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
25	−11.9	−9.3	−7.3	−5.3	1.01	1.40	1.79	2.28
50	−12.9	−9.9	−7.7	−5.5	0.97	1.35	1.70	2.16
100	−13.3	−10.2	−7.9	−5.6	0.95	1.31	1.65	2.09
250	−13.6	−10.3	−8.0	−5.7	0.93	1.28	1.62	2.04
500	−13.7	−10.4	−8.0	−5.7	0.93	1.28	1.61	2.04
∞	−13.8	−10.5	−8.1	−5.7	0.93	1.28	1.60	2.03

Figure 1 : Table 1 Critical values for the Dickey-Fuller ρ test

Skewed to the left! For finite T, these are exact only under the assumption of Gaussian innovations. As T becomes large, these values are also valid for non-Gaussian innovations.

Example 1 (Nominal Interest Rate)

Data: nominal three-month U.S. Treasury bill rate, quarterly, from 1947:2 to 1989:1, $T = 168$
Model: AR(1) by OLS estimation

$$\begin{array}{ccc} i_t = & 0.99694 & i_{t-1}, \\ & (0.010592) & \end{array}$$

The Dickey-Fuller ρ test of $\rho = 1$ is

$$T(\hat{\rho}_T - 1) = 168(0.99694 - 1) = -0.51,$$

This is well above the critical value -7.9 (T=100). So the null is accepted at the 5% level.

Dickey-Fuller t test for Case 1

Another popular statistics for testing the null hypothesis that $\rho = 1$ is based on the usual OLS t test of this hypothesis,

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}}, \tag{21}$$

where $\hat{\sigma}_{\hat{\rho}_T}$ is the usual OLS standard error for the estimated coefficient,

$$\hat{\sigma}_{\hat{\rho}_T} = \left\{ s_T^2 \div \sum_{t=1}^{T} y_{t-1}^2 \right\}^{1/2},$$

and s_T^2 denotes the OLS estimate of the residual variance:

$$s_T^2 = \sum_{t=1}^{T} (y_t - \hat{\rho}_T y_{t-1})^2 / (T - 1).$$

As $T \to \infty$,

$$t_T \xrightarrow{L} \frac{(1/2)\{[W(1)]^2 - 1\}}{\left\{\int_0^1 [W(r)]^2 dr\right\}^{1/2}} \triangleq DF_{t,case1}. \tag{22}$$

Sample size T	Probability that $(\hat{\rho} - 1)/\hat{\sigma}_{\hat{\rho}}$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
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

Figure 2 : Table 2 Critical values for the Dickey-Fuller t test

Example 1 (Nominal Interest Rate)

$$i_t = \underset{(0.010592)}{0.99694} \quad i_{t-1},$$

The Dickey-Fuller t test of $\rho = 1$ is

$$t = (0.99694 - 1)/0.010592 = -0.29,$$

This is well above the 5% critical value of -1.95 (T=100). So the null is again accepted.

Dickey-Fuller tests of Unit Root: Case 2

We now consider Case 2

Case 2 Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
True process: $y_t = y_{t-1} + u_t$ $u_t \sim i.i.d.N(0, \sigma^2)$

Dickey-Fuller ρ test for Case 2

Let us first describe the properties of the OLS estimates:

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T \end{bmatrix} = \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum y_{t-1} y_t \end{bmatrix}, \tag{23}$$

under the null hypothesis that $\alpha = 0$ and $\rho = 1$ (here $\sum$ indicates summation over $t = 1, 2, \cdots, T$). We have

$$\begin{bmatrix} T^{1/2}\hat{\alpha}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \sigma & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int [W(r)]^2 dr \end{bmatrix}^{-1} \times \begin{bmatrix} W(1) \\ (1/2)\left\{[W(1)]^2 - 1\right\} \end{bmatrix}$$

The second element in the above vector expression states that

$$T(\hat{\rho}_T - 1) \xrightarrow{L} \frac{\frac{1}{2}\{[W(1)]^2 - 1\} - W(1) \cdot \int W(r)dr}{\int [W(r)]^2 dr - \left[\int W(r)dr\right]^2} \triangleq DF_{\rho,case2}. \tag{24}$$

This distribution is even more strongly skewed than $DF_{\rho,case1}$.

Sample size T	Probability that $T(\hat{\rho} - 1)$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
25	−11.9	−9.3	−7.3	−5.3	1.01	1.40	1.79	2.28
50	−12.9	−9.9	−7.7	−5.5	0.97	1.35	1.70	2.16
100	−13.3	−10.2	−7.9	−5.6	0.95	1.31	1.65	2.09
250	−13.6	−10.3	−8.0	−5.7	0.93	1.28	1.62	2.04
500	−13.7	−10.4	−8.0	−5.7	0.93	1.28	1.61	2.04
∞	−13.8	−10.5	−8.1	−5.7	0.93	1.28	1.60	2.03
				Case 2				
25	−17.2	−14.6	−12.5	−10.2	−0.76	0.01	0.65	1.40
50	−18.9	−15.7	−13.3	−10.7	−0.81	−0.07	0.53	1.22
100	−19.8	−16.3	−13.7	−11.0	−0.83	−0.10	0.47	1.14
250	−20.3	−16.6	−14.0	−11.2	−0.84	−0.12	0.43	1.09
500	−20.5	−16.8	−14.0	−11.2	−0.84	−0.13	0.42	1.06
∞	−20.7	−16.9	−14.1	−11.3	−0.85	−0.13	0.41	1.04

Figure 3 : Table 1 Critical values for the Dickey-Fuller ρ test

Dickey-Fuller t test for Case 2

An alternative test based on the OLS t test of the null hypothesis that $\rho = 1$:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\hat{\rho}_T}}$$

where

$$\hat{\sigma}^2_{\hat{\rho}_T} = s_T^2 [0 \ 1] \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$s_T^2 = (T-2)^{-1} \sum_{t=1}^{T} (y_t - \hat{\alpha}_T - \hat{\rho}_T y_{t-1})^2.$$

The asymptotic distribution is

$$t_T \xrightarrow{L} \frac{\frac{1}{2}\{[W(1)]^2 - 1\} - W(1) \cdot \int W(r)dr}{\left\{ \int [W(r)]^2 dr - \left[\int W(r)dr \right]^2 \right\}^{1/2}} \triangleq DF_{t,case2}. \quad (25)$$

Sample size T	Probability that $(\hat{\rho} - 1)/\hat{\sigma}_{\hat{\rho}}$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
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
				Case 2				
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

Figure 4 : Table 2 Critical values for the Dickey-Fuller t test

Dickey-Fuller F test for case 2

We are also concerned with the joint hypothesis that $\alpha = 0$ and $\rho = 1$. This null can be represented as $R\theta = r$, with $R = I_2$, $\theta = (\alpha, \rho)'$ and $r = (0, 1)'$.

The F test is then

$$F_T = (\hat{\theta}_T - \theta)' R' \left\{ s_T^2 \cdot R(\sum x_t x_t')^{-1} R' \right\}^{-1} R(\hat{\theta}_T - \theta)/2,$$

where $x_t = (1, y_{t-1})'$. We have

$$\begin{aligned} F_T \quad \xrightarrow{L} \quad & \frac{1}{2} \left[W(1) \ \frac{1}{2}\{[W(1)]^2 - 1\} \right] \\ & \times \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int [W(r)]^2 dr \end{bmatrix}^{-1} \begin{bmatrix} W(1) \\ \frac{1}{2}\{[W(1)]^2 - 1\} \end{bmatrix}. \end{aligned}$$

Sample size T	Probability that F test is greater than entry							
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
				Case 2				
	(F test of $\alpha = 0$, $\rho = 1$ in regression $y_t = \alpha + \rho y_{t-1} + u_t$)							
25	0.29	0.38	0.49	0.65	4.12	5.18	6.30	7.88
50	0.29	0.39	0.50	0.66	3.94	4.86	5.80	7.06
100	0.29	0.39	0.50	0.67	3.86	4.71	5.57	6.70
250	0.30	0.39	0.51	0.67	3.81	4.63	5.45	6.52
500	0.30	0.39	0.51	0.67	3.79	4.61	5.41	6.47
∞	0.30	0.40	0.51	0.67	3.78	4.59	5.38	6.43

Figure 5 : Table 3 Critical values for the Dickey-Fuller F test

Example 1 (nominal interest rate)

The estimated regression now is

$$\begin{array}{rcc} i_t = & 0.211 + 0.96691 & i_{t-1}, \\ & (0.112)\ (0.019133) & \end{array}$$

The Dickey-Fuller ρ test is

$$T(\hat{\rho}_T - 1) = 168(0.96691 - 1) = -5.56,$$

The critical value is found by interpolation to be -13.8. Since $-5.56 > -13.8$, the null of a unit root is accepted at the 5% level.

The Dickey-Fuller t test is

$$(0.96691 - 1)/0.019133 = -1.73. \tag{26}$$

Since $-1.73 > -2.89$ (the cv when T=100), the null of $\rho = 1$ is again accepted.

Example 1 (nominal interest rate)

We further consider the joint hypothesis that $\alpha = 0$ and $\rho = 1$.

The Dickey Fuller F statistic is 1.81. This would have an nonstandard distribution. The 5% critical value is found by interpolation to be 4.67. Since $1.81 < 4.67$, the joint null is accepted.

Dickey-Fuller tests of Unit Root: Case 3

We now consider Case 3

Case 3 Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
True process: $y_t = \alpha + y_{t-1} + u_t \ \alpha \neq 0, u_t \sim i.i.d.(0, \sigma^2)$

Standard Inference for Case 3

Note that under the null, $y_t = \alpha + y_{t-1} + u_t$, we have

$$y_t = y_0 + \alpha t + (u_1 + \ldots + u_t) = y_0 + \alpha t + \xi_t,$$

where $\xi_t = u_1 + \ldots + u_t$ with $\xi_0 = 0$. This leads to

$$\sum_{t=1}^{T} y_{t-1} = \sum_{t=1}^{T} [y_0 + \alpha(t-1) + \xi_{t-1}],$$

where the second term is $O(T^2)$, while the last term is $O(T^{3/2})$. Thus, y_{t-1} behaves approximately as $\alpha(t-1)$, a time trend. The asymptotic distribution of the OLS estimate is

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T^{3/2}(\hat{\rho}_T - 1) \end{bmatrix} \xrightarrow{L} N(0, \sigma^2 Q^{-1}),$$

where $Q = \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix}$. Therefore,

$$(\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T} \xrightarrow{L} N(0, 1).$$

The normal and F critical values are applicable (West, 1988).

Dickey-Fuller tests of Unit Root: Case 4

We now consider Case 4

Case 4 Estimated regression: $y_t = \alpha + \delta t + \rho y_{t-1} + u_t$
True process: $y_t = \alpha + y_{t-1} + u_t$ α any, $u_t \sim i.i.d.N(0, \sigma^2)$

Dickey-Fuller ρ test for Case 4

For this case, the true value of α turns out not to matter for the asymptotic distribution.

Note that the estimated regression can be equivalently be written as

$$\begin{aligned} y_t &= (1-\rho)\alpha + \rho[y_{t-1} - \alpha(t-1)] + (\delta + \rho\alpha)t + u_t, \\ &= \alpha^* + \rho^* \xi_{t-1} + \delta^* t + u_t, \end{aligned} \tag{27}$$

where $\alpha^* = (1-\rho)\alpha, \rho^* = \rho, \delta^* = (\delta + \rho\alpha)$, and $\xi_t = y_t - \alpha t$.

The maintained hypothesis is that $\alpha = \alpha_0, \rho = 1$, and $\delta = 0$, which in the transformed system would mean $\alpha^* = 0, \rho^* = 1$, and $\delta^* = \alpha_0$.

The asymptotic distribution is

$$\begin{bmatrix} T^{1/2}\hat{\alpha}_T^* \\ T(\hat{\rho}_T^* - 1) \\ T^{3/2}(\hat{\delta}_T^* - \alpha_0) \end{bmatrix}$$

$$\xrightarrow{L} \begin{bmatrix} \sigma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sigma \end{bmatrix} \begin{bmatrix} 1 & \int W(r)dr & \frac{1}{2} \\ \int W(r)dr & \int [W(r)]^2 dr & \int rW(r)dr \\ \frac{1}{2} & \int rW(r)dr & \frac{1}{3} \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} W(1) \\ \frac{1}{2}\{[W(1)]^2 - 1\} \\ w(1) - \int W(r)dr \end{bmatrix}. \tag{28}$$

The asymptotic distribution of $\hat{\rho}_T$ is identical to $\hat{\rho}_T^*$, denoted as $DF_{\rho,case4}$.

Sample size T	Probability that $T(\hat{\rho} - 1)$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
25	−11.9	−9.3	−7.3	−5.3	1.01	1.40	1.79	2.28
50	−12.9	−9.9	−7.7	−5.5	0.97	1.35	1.70	2.16
100	−13.3	−10.2	−7.9	−5.6	0.95	1.31	1.65	2.09
250	−13.6	−10.3	−8.0	−5.7	0.93	1.28	1.62	2.04
500	−13.7	−10.4	−8.0	−5.7	0.93	1.28	1.61	2.04
∞	−13.8	−10.5	−8.1	−5.7	0.93	1.28	1.60	2.03
				Case 2				
25	−17.2	−14.6	−12.5	−10.2	−0.76	0.01	0.65	1.40
50	−18.9	−15.7	−13.3	−10.7	−0.81	−0.07	0.53	1.22
100	−19.8	−16.3	−13.7	−11.0	−0.83	−0.10	0.47	1.14
250	−20.3	−16.6	−14.0	−11.2	−0.84	−0.12	0.43	1.09
500	−20.5	−16.8	−14.0	−11.2	−0.84	−0.13	0.42	1.06
∞	−20.7	−16.9	−14.1	−11.3	−0.85	−0.13	0.41	1.04
				Case 4				
25	−22.5	−19.9	−17.9	−15.6	−3.66	−2.51	−1.53	−0.43
50	−25.7	−22.4	−19.8	−16.8	−3.71	−2.60	−1.66	−0.65
100	−27.4	−23.6	−20.7	−17.5	−3.74	−2.62	−1.73	−0.75
250	−28.4	−24.4	−21.3	−18.0	−3.75	−2.64	−1.78	−0.82
500	−28.9	−24.8	−21.5	−18.1	−3.76	−2.65	−1.78	−0.84
∞	−29.5	−25.1	−21.8	−18.3	−3.77	−2.66	−1.79	−0.87

The probability shown at the head of the column is the area in the left-hand tail.

Figure 6 : Table 1 Critical values for the Dickey-Fuller ρ test

Dickey-Fuller t test for Case 4

The asymptotic distribution of the OLS t test of the hypothesis that $\rho = 1$ is given by

$$t_T = T(\hat{\rho}_T - 1) \div (T^2 \cdot \hat{\sigma}^2_{\hat{\rho}_T})^{1/2} \xrightarrow{p} T(\hat{\rho}_T - 1) \div \sqrt{Q} \stackrel{\triangle}{=\!=} DF_{t,case4},$$

where

$$Q = [0\ 1\ 0] \begin{bmatrix} 1 & \int W(r)dr & \frac{1}{2} \\ \int W(r)dr & \int [W(r)]^2 dr & \int rW(r)dr \\ \frac{1}{2} & \int rW(r)dr & \frac{1}{3} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Sample size T	Probability that $(\hat{\rho} - 1)/\hat{\sigma}_{\hat{\rho}}$ is less than entry							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				Case 1				
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
				Case 2				
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
				Case 4				
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

The probability shown at the head of the column is the area in the left-hand tail.

Figure 7 : Table 2 Critical values for the Dickey-Fuller t test

Dickey-Fuller F test for case 4

We are concerned with the hypothesis that $\rho = 1$ and $\delta = 0$. The F statistic can be derived in a similar way to that of case 2.

Sample size T	Probability that F test is greater than entry							
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
	Case 2							
	(F test of $\alpha = 0$, $\rho = 1$ in regression $y_t = \alpha + \rho y_{t-1} + u_t$)							
25	0.29	0.38	0.49	0.65	4.12	5.18	6.30	7.88
50	0.29	0.39	0.50	0.66	3.94	4.86	5.80	7.06
100	0.29	0.39	0.50	0.67	3.86	4.71	5.57	6.70
250	0.30	0.39	0.51	0.67	3.81	4.63	5.45	6.52
500	0.30	0.39	0.51	0.67	3.79	4.61	5.41	6.47
∞	0.30	0.40	0.51	0.67	3.78	4.59	5.38	6.43
	Case 4							
	(F test of $\delta = 0$, $\rho = 1$ in regression $y_t = \alpha + \delta t + \rho y_{t-1} + u_t$)							
25	0.74	0.90	1.08	1.33	5.91	7.24	8.65	10.61
50	0.76	0.93	1.11	1.37	5.61	6.73	7.81	9.31
100	0.76	0.94	1.12	1.38	5.47	6.49	7.44	8.73
250	0.76	0.94	1.13	1.39	5.39	6.34	7.25	8.43
500	0.76	0.94	1.13	1.39	5.36	6.30	7.20	8.34
∞	0.77	0.94	1.13	1.39	5.34	6.25	7.16	8.27

The probability shown at the head of the column is the area in the right-hand tail.

Figure 8 : Table 3 Critical values for the Dickey-Fuller F test

Summary of Dickey-Fuller tests for unit roots

- Case 1:
  - Estimated regression: $y_t = \rho y_{t-1} + u_t$
  - True process: $y_t = y_{t-1} + u_t \; u_t \sim i.i.d.N(0, \sigma^2)$
  - $T(\hat{\rho}_T - 1)$ has the distribution $DF_{\rho,case1}$ (Case 1 in Table 1)
  - $(\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T}$ has the distribution $DF_{t,case1}$ (Case 1 in Table 2)
- Case 2:
  - Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
  - True process: $y_t = y_{t-1} + u_t \; u_t \sim i.i.d.N(0, \sigma^2)$
  - $T(\hat{\rho}_T - 1)$ has the distribution $DF_{\rho,case2}$ (Case 2 in Table 1)
  - $(\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T}$ has the distribution $DF_{t,case2}$ (Case 2 in Table 2)
  - OLS F test of the joint hypothesis that $\alpha = 0$ and $\rho = 1$ has nonstandard distribution (Dickey and Fuller, 1981), see Case 2 in Table 3

- Case 3:
  - Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
  - True process: $y_t = \alpha + y_{t-1} + u_t \ \alpha \neq 0, \ u_t \sim i.i.d.(0, \sigma^2)$
  - $(\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T} \xrightarrow{L} N(0,1)$
- Case 4:
  - Estimated regression: $y_t = \alpha + \rho y_{t-1} + \delta t + u_t$
  - True process: $y_t = \alpha + y_{t-1} + u_t \ \alpha$ any, $u_t \sim i.i.d.N(0, \sigma^2)$
  - $T(\hat{\rho}_T - 1)$ has the distribution $DF_{\rho,case4}$ (Case 4 in Table 1)
  - $(\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T}$ has the distribution $DF_{t,case4}$ (Case 4 in Table 2)
  - OLS F test of joint hypothesis that $\rho = 1$ and $\delta = 0$ has nonstandard distribution (Case 4 in Table 3)

Which is the 'correct' case to use?

1. If the analyst has a specific null hypothesis about the process that generated the data, then obviously this would guide the choice of test.
2. In the absence of such guidance, one general principle would be to fit a specification that is a plausible description of the data under both the null hypothesis and the alternative. Thus principle would suggest using the case 4 test for a series with an obvious trend and the case 2 test for series without a significant trend.

Example 1 (nominal interest rate)

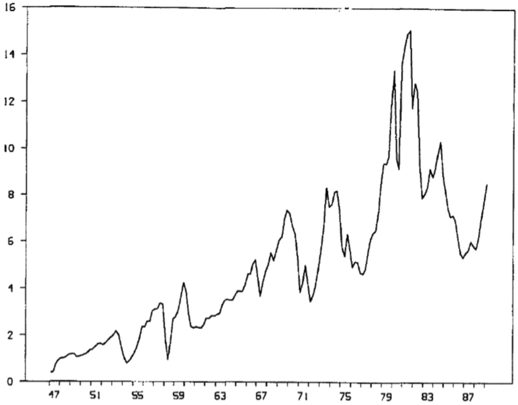

FIGURE 17.2 U.S. nominal interest rate on 3-month Treasury bills, data sampled quarterly but quoted at an annual rate, 1947:I to 1989:I.

Although the time series has tended upward over this sample period, there is nothing in economic theory to suggest that nominal interest rates should exhibit a deterministic time trend, so a natural null hypothesis is that the true process is a random walk without trend. In terms of framing a plausible alternative, it is difficult to maintain that these data could have been generated by $i_t = \rho i_{t-1} + u_t$ with $|\rho|$ significantly less than 1. If these data were to be described by a stationary process, surely the process would have a positive mean. This argues for including a constant term in the estimated regression, even though under the null hypothesis the true process does not contain a constant term. Thus, case 2 is a sensible approach for these data.

Example 2 (U.S. real GNP)

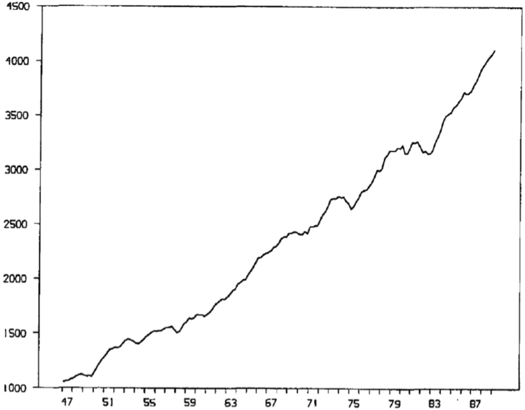

FIGURE 17.3 U.S. real GNP, data sampled quarterly but quoted at an annual rate in billions of 1982 dollars, 1947:I to 1989:I.

Given a growing population and technological improvements, such a series would certainly be expected to exhibit a persistent upward trend, and this trend is unmistakable in the figure. The question is whether this trend arises from positive drift term of a random walk:

$$H_0 : y_t = \alpha + y_{t-1} + u_t, \alpha > 0, \tag{29}$$

or from a deterministic time trend added to a stationary AR(1):

$$H_A : y_t = \alpha + \delta t + \rho y_{t-1} + u_t, |\rho| < 1. \tag{30}$$

Thus the recommended test statistics for this case are those described in case 4.

The following model for 100 times the log of real GNP (denoted y_t) was estimated by OLS regression:

$$y_t = \underset{(13.53)}{27.24} + \underset{(0.019304)}{0.96252}\, y_{t-1} + \underset{(0.01521)}{0.02753}\, t.$$

The sample size is $T = 168$. The Dickey-Fuller ρ test is

$$T(\hat{\rho} - 1) = 168(0.96252 - 1.0) = -6.3. \qquad (31)$$

Since $-6.3 > -21.0$, the null hypothesis that GNP is characterized by a random walk with possible drift is accepted at the 5% level. The Dickey-Fuller t test,

$$t = \frac{0.96252 - 1.0}{0.019304} = -1.94, \qquad (32)$$

exceeds the 5% critical value of -3.44, so that the null hypothesis of a unit root is accepted by this test as well. Finally the F test of the joint null hypothesis that $\delta = 0$ and $\rho = 1$ is 2.44. Since this is less than the 5% critical value of 6.42 from Table 3, this null hypothesis is again accepted.

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation
   - Asymptotic Results for unit root processes with general serial correlation
   - Phillips-Perron Unit Root test for Case 2
   - Augmented Dickey-Fuller tests for unit roots

Serially correlated disturbances

Proposition 2 (Beveridge and Nelson (1981) Decomposition)

$$u_t = \phi(L)\varepsilon_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j}$$

where

$$\begin{aligned} E(\varepsilon_t) &= 0 \\ E(\varepsilon_t \varepsilon_\tau) &= \begin{cases} \sigma^2 & \text{for } t = \tau \\ 0 & \text{otherwise} \end{cases} \\ \sum_{j=0}^{\infty} j \cdot |\phi_j| &< \infty. \end{aligned}$$

Then

$$u_1 + \ldots + u_t = \phi(1) \cdot (\varepsilon_1 + \ldots + \varepsilon_t) + \eta_t - \eta_0,$$

where $\phi(1) = \sum \phi_j$, $\eta_t = \sum \alpha_j \varepsilon_{t-j}$, $\alpha_j = -(\phi_{j+1} + \phi_{j+2} + \ldots)$, and $\sum |\alpha_j| < \infty$.

Proposition 3 Let $u_t = \phi(L)\varepsilon_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j}$, where $\sum_{j=0}^{\infty} j \cdot |\phi_j| < \infty$ and $\{\varepsilon_t\}$ is an $i.i.d.$ sequence with mean zero, variance σ^2, and finite fourth moment. Define

$$\begin{aligned}
\gamma_j &= E(u_t u_{t-j}) = \sigma^2 \sum_{s=0}^{\infty} \phi_s \phi_{s+j} \ \textit{for}\ j = 0, 1, 2, \cdots \\
\lambda &= \sigma \sum_{j=0}^{\infty} \phi_j = \sigma \cdot \phi(1) \\
y_t &= u_1 + u_2 + \cdots + u_t \ \textit{for}\ t = 1, 2, \cdots, T
\end{aligned}$$

with $y_0 = 0$, Then

(a) $T^{-1/2}\sum_{t=1}^{T} u_t \xrightarrow{L} \lambda \cdot W(1)$;

(b) $T^{-1/2}\sum_{t=1}^{T} u_{t-j}\varepsilon_t \xrightarrow{L} N(0, \sigma^2\gamma_0)$ for $j = 1, 2, \cdots$;

(c) $T^{-1}\sum_{t=1}^{T} u_t u_{t-j} \xrightarrow{p} \gamma_j$ for $j = 0, 1, 2, \cdots$;

(d) $T^{-1}\sum_{t=1}^{T} y_{t-1}\varepsilon_t \xrightarrow{L} (1/2)\sigma \cdot \lambda \cdot \{[W(1)]^2 - 1\}$;

(e) $T^{-1}\sum_{t=1}^{T} y_{t-1}u_{t-j}$

$$\xrightarrow{L} \begin{cases} (1/2)\{\lambda^2 \cdot [W(1)]^2 - \gamma_0\} & \textit{for} \quad j = 0 \\ (1/2)\{\lambda^2 \cdot [W(1)]^2 - \gamma_0\} + \gamma_0 + \gamma_1 + \cdots + \gamma_{j-1} \\ \qquad \textit{for} \quad j = 1, 2, \cdots ; \end{cases}$$

(f) $T^{-3/2}\sum_{t=1}^{T} y_{t-1} \xrightarrow{L} \lambda \cdot \int_0^1 W(r)dr$;

(g) $T^{-3/2}\sum_{t=1}^{T} tu_{t-j} \xrightarrow{L} \lambda \cdot \left\{W(1) - \int_0^1 W(r)dr\right\}$ for $j = 0, 1, 2, \cdots$;

(h) $T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{L} \lambda^2 \cdot \int_0^1 [W(r)]^2 dr$;

(i) $T^{-5/2}\sum_{t=1}^{T} ty_{t-1} \xrightarrow{L} \lambda \cdot \int_0^1 rW(r)dr$;

(j) $T^{-3}\sum_{t=1}^{T} ty_{t-1}^2 \xrightarrow{L} \lambda^2 \cdot \int_0^1 r \cdot [W(r)]^2 dr$;

(k) $T^{(\nu+1)}\sum_{t=1}^{T} t^{\nu} \to 1/(\nu + 1)$ for $\nu = 0, 1, \cdots$.

Unit Root tests with serial correlation

Two approaches:

- Phillips-Perron Tests: Phillips (1987) and Phillips and Perron (1988) still use OLS estimation, but modify the statistics
- Augmented Dickey-Fuller Tests: Said and Dickey (1984) add lagged changes of y as explanatory variables in the regression to deal with serial correlation

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation
   - Asymptotic Results for unit root processes with general serial correlation
   - Phillips-Perron Unit Root test for Case 2
   - Augmented Dickey-Fuller tests for unit roots

Phillips-Perron tests for unit roots

Case 2:
Estimated regression: $y_t = \alpha + \rho y_{t-1} + u_t$
True process: $y_t = y_{t-1} + u_t$
where $u_t = \phi(L)\varepsilon_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j}$ satisfies Proposition 3.

We have

$$T(\hat{\rho}_T - 1) \xrightarrow{L} DF_{\rho,case2} + \frac{(1/2)\cdot(\lambda^2 - \gamma_0)}{\lambda^2 \left\{ \int [W(r)]^2 dr - \left[\int W(r) dr \right]^2 \right\}},$$

where the second term in the right hand side is a correction for serial correlation.

By modifying the test statistic, we have

$$
\begin{aligned}
Z_\rho &= T(\hat{\rho}_T - 1) - (1/2)\{T^2 \cdot \hat{\sigma}^2_{\hat{\rho}_T} \div s^2_T\}(\hat{\lambda}^2_T - \hat{\gamma}_{0,T}) \\
&\xrightarrow{L} DF_{\rho,case2}
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\gamma}_{j,T} &= T^{-1} \sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j} \\
\hat{u}_t &= \textit{OLS sample residual from the estimated regression} \\
\hat{\lambda}^2_T &= \hat{\gamma}_{0,T} + 2 \cdot \sum_{j=1}^{q} [1 - j/(q+1)] \hat{\gamma}_{j,T} \\
s^2_T &= (T-k)^{-1} \sum_{t=1}^{T} \hat{u}^2_t \\
k &= \textit{number of parameters in estimated regression} \\
\hat{\sigma}_{\hat{\rho}_T} &= \textit{OLS standard error for } \hat{\rho}_T
\end{aligned}
$$

We also have

$$
\begin{aligned}
Z_t &= (\hat{\gamma}_{0,T}/\hat{\lambda}_T^2)^{1/2} \cdot (\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T} \\
&\quad -(1/2)(\hat{\lambda}_T^2 - \hat{\gamma}_{0,T})(1/\hat{\lambda}_T)\{T \cdot \hat{\sigma}_{\hat{\rho}_T} \div s_T\} \\
&\xrightarrow{L} DF_{t,case2}.
\end{aligned}
$$

For Case 1, we have $Z_\rho \xrightarrow{L} DF_{\rho,case1}$ and $Z_t \xrightarrow{L} DF_{t,case1}$.
For Case 4, we have $Z_\rho \xrightarrow{L} DF_{\rho,case4}$ and $Z_t \xrightarrow{L} DF_{t,case4}$.

Then Table 1 and 2 can be used to find out the critical values for the Phillips-Perron test.

Today's Topics

1. Rewiew of AR(1)

2. Brownian Motion and Functional CLT

3. Asymptotic properties of Unit Root processes and tests for Unit Root

4. Generalization to processes with serial correlation
   - Asymptotic Results for unit root processes with general serial correlation
   - Phillips-Perron Unit Root test for Case 2
   - Augmented Dickey-Fuller tests for unit roots

Augmented Dickey-Fuller tests for unit roots

Suppose y_t were really generated from an $AR(p)$ process,

$$(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p)y_t=\varepsilon_t, \tag{33}$$

where $\varepsilon_t \sim i.i.d.(0,\sigma^2)$, and finite fourth moment.

It is easy to verify that (33) can equivalently be written as

$$y_t=\rho y_{t-1}+\zeta_1\Delta y_{t-1}+\zeta_2\Delta y_{t-2}+\cdots+\zeta_{p-1}\Delta y_{t-p+1}+\varepsilon_t, \tag{34}$$

with

$$\begin{aligned}\rho &= \phi_1+\phi_2+\cdots+\phi_p,\\ \zeta_j &= -[\phi_{j+1}+\phi_{j+2}+\cdots+\phi_p] for\ j=1,2,\cdots,p-1.\end{aligned}$$

Under the null hypothesis that $\rho = 1$, expression (34) can be written as

$$(1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})\Delta y_t = \varepsilon_t$$

or

$$\Delta y_t = u_t,$$

where

$$u_t = (1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})^{-1}\varepsilon_t.$$

This indicates that y_t behaves like the variable y_t described in Proposition 3, with

$$\phi(L) = (1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})^{-1}.$$

Case 1:

- Estimated Regression:

$$y_t = \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t.$$

- True process: same specification as estimated regression with $\rho = 1$
- Any t or F test involving $\zeta_1, \zeta_2, \cdots, \zeta_{p-1}$ can be compared with the usual t or F tables for an asymptotically valid test.
- The augmented Dickey-Fuller ρ test

$$Z_{DF} = \frac{T \cdot (\hat{\rho}_T - 1)}{1 - \hat{\zeta}_{1,T} - \hat{\zeta}_{2,T} - \cdots - \hat{\zeta}_{p-1,T}} \xrightarrow{L} DF_{\rho,case1}$$

- OLS t test of $\rho = 1 \xrightarrow{L} DF_{t,case1}$

Case 2:

- Estimated Regression:

$$y_t = \alpha + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t. \tag{35}$$

- True process: same specification as estimated regression with $\alpha = 0$ and $\rho = 1$
- Any t or F test involving $\zeta_1, \zeta_2, \cdots, \zeta_{p-1}$ can be compared with the usual t or F tables for an asymptotically valid test.
- The augmented Dickey-Fuller ρ test $Z_{DF} \xrightarrow{L} DF_{\rho,case2}$
- OLS t test of $\rho = 1 \xrightarrow{L} DF_{t,case2}$
- OLS F test of joint hypothesis that $\alpha = 0$ and $\rho = 1$ (Case 2 in Table 3).

Example 1 (nominal interest rate)

The following model was estimated by OLS for the interest rate data described in Example 17.3 (standard errors in parentheses):

$$\begin{aligned} i_t = &\underset{(0.0788)}{0.335}\ \Delta i_{t-1} - \underset{(0.0808)}{0.388}\ \Delta i_{t-2} + \underset{(0.0800)}{0.276}\ \Delta i_{t-3} \\ &- \underset{(0.0794)}{0.107}\ \Delta i_{t-4} + \underset{(0.109)}{0.195} + \underset{(0.018604)}{0.96904}\ i_{t-1}. \end{aligned}$$

Dates t = 1948:II through 1989:I were used for estimation, so in this case the sample size is $T = 164$. For these estimates, the augmented Dickey-Fuller ρ test [17.7.35] would be

$$\frac{164}{1 - 0.335 + 0.388 - 0.276 + 0.107}(0.96904 - 1) = -5.74.$$

Since $-5.74 > -13.8$, the null hypothesis that the Treasury bill rate follows a fifth-order autoregression with no constant term, and a single unit root, is accepted at the 5% level. The OLS t test for this same hypothesis is

$$(0.96904 - 1)/(0.018604) = -1.66.$$

Since $-1.66 > -2.89$, the null hypothesis of a unit root is accepted by the augmented Dickey-Fuller t test as well. Finally, the OLS F test of the joint null hypothesis that $\rho = 1$ and $\alpha = 0$ is 1.65. Since this is less than 4.68, the null hypothesis is again accepted.

Case 3:

- Estimated Regression:

$$y_t = \alpha + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t.$$

- True process: same specification as estimated regression with $\alpha \neq 0$ and $\rho = 1$
- $\hat{\rho}_T$ converges at rate $T^{3/2}$ to a Gaussian variable; all other estimated coefficients converge at rate $T^{1/2}$ to Gaussian variables.
- Any t or F test involving any coefficients from the regression can be compared with the usual t or F tables for an asymptotically valid test.

Case 4:

- Estimated Regression:

$$y_t = \alpha + \rho y_{t-1} + \delta t + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t.$$

- True process: same specification as estimated regression with α any value, $\rho = 1$, and $\delta = 0$
- Any t or F test involving $\zeta_1, \zeta_2, \cdots, \zeta_{p-1}$ can be compared with the usual t or F tables for an asymptotically valid test.
- The augmented Dickey-Fuller ρ test $Z_{DF} \xrightarrow{L} DF_{\rho,case4}$
- OLS t test of $\rho = 1 \xrightarrow{L} DF_{t,case4}$
- OLS F test of joint hypothesis that $\rho = 1$ and $\delta = 0$ (Case 4 in Table 3).

Example 2 (U.S. real GNP)

The following autoregression was estimated by *OLS* for the GNP data in Figure 17.3 (standard errors in parentheses):

$$
\begin{aligned}
y_t = {} & \underset{(0.0777)}{0.329}\ \Delta y_{t-1} + \underset{(0.0813)}{0.209}\ \Delta y_{t-2} - \underset{(0.0818)}{0.084}\ \Delta y_{t-3} \\
& - \underset{(0.0788)}{0.075}\ \Delta y_{t-4} + \underset{(13.57)}{35.92} + \underset{(0.019386)}{0.94969}\ y_{t-1} + \underset{(0.0152)}{0.0378}\ t.
\end{aligned}
$$

Here, $T = 164$ and the augmented Dickey-Fuller ρ test is

$$\frac{164}{1 - 0.329 - 0.209 + 0.084 + 0.075}(0.94969 - 1) = -13.3.$$

Since $-13.3 > -21.0$, the null hypothesis that the log of GNP is $ARIMA(4, 1, 0)$ with possible drift is accepted at the 5% level. The augmented Dickey-Fuller t test also accepts this hypothesis:

$$(0.94969 - 1)/(0.019386) = -2.60 > -3.44.$$

The *OLS* F test of the joint null hypothesis that $\rho = 1$ and $\delta = 0$ is $3.74 < 6.42$, and so the augmented Dickey-Fuller F test is also consistent with the unit root specification.

Choice of p in ADF test

Take Case 2 as an example. Ng and Perron (1995) suggest estimate (35) with p taken to be some prespecified upper bound $\bar{p}$.

First consider the OLS t test of $\zeta_{\bar{p}-1} = 0$. If this null is accepted, then consider OLS F test of $\zeta_{\bar{p}-1} = 0$ and $\zeta_{\bar{p}-2} = 0$ ($F(2, T-k)$). The procedure continued sequentially until the joint null hypothesis that $\zeta_{\bar{p}-1} = 0, \zeta_{\bar{p}-2} = 0, \cdots, \zeta_{\bar{p}-\ell} = 0$ is rejected for some ℓ. The recommended regression is then

$$y_t = \alpha + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{\bar{p}-\ell} \Delta y_{t-\bar{p}+\ell} + \varepsilon_t.$$

If no value of ℓ leads to the rejection, the simple Dickey-Fuller test is used.

See Hall (1994) for a variety of alternative approaches for estimating p.

For other issues with unit roots, see Choi (2015), *Almost All About Unit Roots*, Cambridge University Press.