Lecture 15: Unit Root and Unit-Root Testing

Bus 41910, Time Series Analysis, Mr. R. Tsay

Recall that a random walk model is

$$Z_t = Z_{t-1} + a_t, \tag{1}$$

where $\{a_t\}$ is a sequence of martigale differences, that is, $E(a_t|F_{t-1}) = 0$, $\text{Var}(a_t|F_{t-1})$ is finite, and $E(|a_t|^{2+\delta}|F_{t-1}) < \infty$ for some $\delta > 0$, where F_{t-1} is the σ-field generated by $\{a_{t-1}, a_{t-2}, \ldots\}$. For simplicity, one often assumes that $Z_0 = 0$. It will be seen later that this assumption has no effect on the limiting distributions of unit-root test statistics. For now, we assume that Z_0 is a random variable so that

$$Z_t = \sum_{i=1}^{t} a_i + Z_0 = S_t + Z_0,$$

where $S_t = \sum_{i=1}^{t} a_i$ is the partial sum of $\{a_t\}_{t=1}^{\infty}$.
Consider the ordinary least squares (OLS) autoregression

$$Z_t = \phi Z_{t-1} + e_t, \quad t = 1, 2, \ldots, T \tag{2}$$

where e_t denotes the error-term and T is the sample size. The LS eatimate of ϕ is

$$\hat{\phi} = \frac{\sum_{t=1}^{T} Z_{t-1} Z_t}{\sum_{t=1}^{T} Z_{t-1}^2}.$$

A simple way to test the random-walk model is to consider the null hypothesis $H_o : \phi = 1$ versus the alternative hypothesis $H_a : \phi < 1$. This is the unit-root testing problem. Dickey and Fuller (1979, JASA) studied this problem that led to the development of unit-root theory in the 1980s and 1990s. The literature on unit-root testing is enormous. However, there are two articles that provide the fundamental results. They are (a) Chan and Wei (1988, Annals of Statistics) and (b) Phillips (1987, Econometrica). The former studies all types of characteristic roots on the unit cirlce (e.g. 1, -1, and all complex roots) for autoregressive processes whereas the latter focuses on a single unit root for general time series models.
The assumption used in Eq. (1) is based on Chan and Wei (1988). The a_t series is allowed to have weak dependence in Phillips (1987). Specifically, Phillips assumes the following:

Assumption A: (a) $E(a_t) = 0$ for all t; (b) $\sup_t E|a_t|^{\beta} < \infty$ for some $\beta > 2$; (c) $\sigma^2 = \lim_{T\to\infty} E(T^{-1}S_T^2)$ exists and $\sigma^2 > 0$, where $S_T = \sum_{t=1}^{T} a_t$; (d) $\{a_t\}$ is strong mixing wiht mixing coefficients α_m that satisfy $\sum_{t=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$.

These conditions are weak and allow for both temporal dependence and heteroscedasticity in the process $\{a_t\}_{t=1}^{\infty}$. Because different assumptions are used, Chan and Wei (1988)

and Phillips (1987) employ different functional central limit theorems (FCLT) to obtain the limiting result for the unit-root statistics. The limiting distributions are the same as expected.
Let C be the space of all real valued continuous on [0,1]. Define the partial sum of $\{a_t\}$ as $S_n = \sum_{t=1}^{n} a_t$. For any $r \in [0,1]$ define

$$X_T(r) = \frac{1}{\sqrt{T}\sigma} S_{[Tr]} \tag{3}$$

where σ is defined in Assumption A and $[Tr]$ denotes the integer part of $T \times r$. The following two results are useful in unit-root study.

Lemma 1. [FCLT, Herndorf (1983, ZW)] If $\{a_t\}$ satisfies the Assumption A, then $X_T \rightarrow_D W$ on C as $T \rightarrow \infty$, where $\rightarrow_D$ denotes convergence in distribution and W is a standard Brownian motion on C.

Lemma 2: [Continuous mapping theorem, e.g., Billingsley (1968)] If $X_T \rightarrow_D W$ as $T \rightarrow \infty$ and h is an y continuous functional on D, then $h(X_T) \rightarrow_D h(W)$ as $T \rightarrow \infty$, where D is the space of all real valued functions on [0,1] that are right continuous at each point of [0,1] and have finite left limits.

Remark: Chan and Wei (1988) used a FCLT by Helland (1982, Scand. J. Statistics) for martingales and establish a generalization of the continuous mapping theorem that is useful for all types of characteristic roots on the unit cirlce. See Theorem 2.2 of Chan and Wei (1988).

Remark: For multivariate ARMA processes with all types of characteristic roots on the unit cirlce, see Tsay and Tiao (1990, Annals of Statistics).

Some limiting Results
For the OLS estimate $\hat{\phi}$, we have

$$T(\hat{\phi} - 1) = \frac{T^{-1}\sum_{t=1}^{T} Z_{t-1}(Z_t - Z_{t-1})}{T^{-2}\sum_{t=1}^{T} Z_{t-1}^2},$$

and the t-ratio of $\hat{\phi}$ is

$$t_\phi = \left[\sum_{t=1}^{T} Z_{t-1}^2\right]^{1/2} (\hat{\phi} - 1)/s, \tag{4}$$

where

$$s^2 = \frac{\sum_{t=1}^{T}(Z_t - \hat{\phi} Z_{t-1})^2}{T}.$$

The following basic results are from Theorem 3.1 of Phillips (1987).

Theorem: If $\{a_t\}$ satisfies Assumption A and if $\sup_t E|a_t|^{\beta+\eta} < \infty$ for some $\eta > 0$, and Z_t follows the random-walk model of Eq. (1), then as $T \to \infty$:

1. $T^{-2}\sum_{t=1}^T Z_{t-1}^2 \to_D \sigma^2 \int_0^1 W(r)^2 dr$;
2. $T^{-1}\sum_{t=1}^T Z_{t-1}(Z_t - Z_{t-1}) \to_D (\sigma^2/2)[W(1)^2 - \sigma_a^2/\sigma^2]$;
3. $T(\hat{\phi} - 1) \to_D \frac{(1/2)[W(1)^2 - \sigma_a^2/\sigma^2]}{\int_0^1 W(r)^2 dr}$;
4. $\hat{\phi} \to_p 1$; (convergence in probability)
5. $t_\phi \to_D \frac{(\sigma/2\sigma_a)[W(1)^2 - \sigma_a^2/\sigma^2]}{\left[\int_0^1 W(r)^2 dr\right]}$;

where $\sigma_a^2 = \lim_{T\to\infty} T^{-1}\sum_{t=1}^T E(a_t^2)$ and $\sigma^2 = \lim_{T\to\infty} E(T^{-1}S_T^2)$.

Proof. The key in proof is to make usre of Lemmas 1 and 2. For Part 1, using $Z_{t-1} = S_{t-1} + Z_0$, we have

$$
\begin{aligned}
T^{-2}\sum_{t=1}^T Z_{t-1}^2 &= T^{-2}\sum_{t=1}^T (S_{t-1}^2 + 2S_{t-1}Z_0 + Z_0^2) \\
&= T^{-1}\sum_{t=1}^T \int_{(t-1)/T}^{t/T} S_{t-1}^2 dr + 2Z_0 T^{-1}\sum_{t=1}^T \int_{(t-1)/T}^{t/T} S_{t-1} dr + T^{-1}Z_0^2 \\
&= \sigma^2 \sum_{t=1}^T \int_{(t-1)/T}^{t/T} \frac{1}{T\sigma^2} S_{[Tr]}^2 dr + 2Z_0\sigma T^{-1/2}\sum_{t=1}^T \int_{(t-1)/T}^{t/T} \frac{1}{\sigma\sqrt{T}} S_{[Tr]} dr + T^{-1}Z_0^2 \\
&= \sigma^2 \int_0^1 X_T^2(r) dr + \frac{\sigma}{\sqrt{T}}\int_0^1 X_T(r) dr + \frac{Z_0^2}{T} \\
&\to_D \sigma^2 \int_0^1 W(r)^2 dr \quad \text{as} \quad T \to \infty.
\end{aligned}
$$

where, in Step 2, we use the fact that S_{t-1} is a constant in the interval $[\frac{t-1}{T}, \frac{t}{T})$ so that the two integrals are simply S_{t-1}^2/T and S_{t-1}/T, respectively, and in Step 3, we use $[Tr] = t-1$ for $r \in [(t-1)/T, t/T)$. From the above proof, Z_0 can be any constant or a random variable as it does not affect the limiting distribution.

For Part 2, we have

$$
\begin{aligned}
T^{-1}\sum_{t=1}^T Z_{t-1}(Z_t - Z_{t-1}) &= T^{-1}\sum_{t-1}^T (S_{t-1} + Z_0)(a_t) \\
&= T^{-1}\sum_{t=1}^T S_{t-1}a_t + Z_0 T^{-1}\sum_{t=1}^T a_t \\
&= (2T)^{-1}\sum_{t=1}^T (S_t^2 - S_{t-1}^2 - a_t^2) + Z_0\bar{a}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\sigma^2}{2}\sum_{t=1}^{T}\frac{1}{T\sigma^2}[S^2_{[T(t/T)]} - S^2_{[T((t-1)/T)]}] - \frac{1}{2T}\sum_{t=1}^{T} a_t^2 + Z_0\bar{a} \\
&= \frac{\sigma^2}{2}\sum_{t=1}^{T}[X_T(t/T)^2 - X_T((t-1)/T)^2] - \frac{1}{2T}\sum_{t=1}^{T} a_t^2 + Z_0\bar{a} \\
&= \frac{\sigma^2}{2}X_T(1)^2 - \frac{1}{2T}\sum_{t=1}^{T} a_t^2 + Z_0\bar{a}.
\end{aligned}
$$

where, in Step 3, we use $S_t^2 = (S_{t-1} + a_t)^2$ so that $S_{t-1}a_t = \frac{1}{2}(S_t^2 - S_{t-1}^2 - a_t^2)$ and $bara = T^1\sum_{t=1}^T a_t$. Under Assumption A, $T^{-1}\sum_{t=1}^T a_t^2$ converges almost surely to the variance of a_t denoted by σ_a^2 and $\bar{a}$ converges almost surely to zero; see, for instance, McLeish (1975, Theorem 2.10). Applying the above results and Lemma 2, we obtain

$$
T^{-1}\sum_{t=1}^{T} Z_{t-1}a_t \to_D \frac{\sigma^2}{2}W(1)^2 - \frac{\sigma_a^2}{2} = \frac{\sigma^2}{2}\left[W(1)^2 - \frac{\sigma_a^2}{\sigma^2}\right].
$$

Part 3 follows directly the continuous mapping theorem and Part 4 follows Part 3. For Part 5, we let

$$
s^2 = T^{-1}\sum_{t=1}^{T}(Z_t - \hat{\phi}Z_{t-1})^2 = T^{-1}\sum_{t=1}^{T} a_t^2 - 2(\hat{\phi}-1)T^{-1}\sum_{t=1}^{T} Z_{t-1}a_t + (\hat{\phi}-1)^2 T^{-1}\sum_{t=1}^{T} Z_{t-1}^2.
$$

Using Parts 1, 2, and 4, it can be shwon that $s^2 \to \sigma_a^2$ in probability as $T \to \infty$. Part 5 then follows the continuous mapping theorem. The proof is complete.

If $\{a_t\}$ is a sequence of $iid(0, \sigma_a^2)$ random variables such that $\sigma_a^2 = \sigma^2$, then

$$
T(\hat{\phi} - 1) \to_D \frac{(1/2)[W(1)^2 - 1]}{\int_0^1 W(r)^2 dr}.
$$

This result was obtained by White (1958, Annals of Math. Statist.).

Exercise. If $a_t = \epsilon_t - \theta\epsilon_{t-1}$ with $\{\epsilon_t\}$ being a sequence of $iid(0, \sigma_e^2)$, what is the limiting distribution of $T(\hat{\phi} - 1)$ as $T \to \infty$?

A useful result in studying unit-root properties is

$$
\int_0^1 W dW = (1/2)[W(1)^2 - 1],
$$

which can be obtained by Ito Lemma. [Apply the Ito Lemma to $G(W, t) = W(t)^2$ to obtain $dW(t)^2 = dt + 2W(t)dW(t)$. Integrate out the equation for $t \in [0, 1]$.]]
To apply the prior results for testing, one need critical values of the test statistics. The critical values are obtained via simulation.

For the random walk Z_t, if one entertains the model

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + e_t,$$

then the OLS estimate of ϕ is

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{T}(Z_{t-1} - \bar{Z}_1)(Z_t - \bar{Z}_2)}{\sum_{t=1}^{T}(Z_{t-1} - \bar{Z}_1)^2},$$

where $\bar{Z}_1 = T^{-1}\sum_{t=1}^{T} Z_{t-1}$ and $\bar{Z}_2 = T^{-1}\sum_{t=1}^{T} Z_t$. The limiting distribution of the t-ratio of $\hat{\phi}_1$ will be different from that of $\hat{\phi}$, but remains a function of the standard Brownian motion. Critical values, again, are obtained by simulation.

Similarly, if one entertains the model

$$Z_t = \alpha_0 + \alpha_1 t + \phi Z_{t-1} + e_t,$$

the limiting distribution of the t-ratio of $\hat{\phi}$ will be another function of the standard Brownian motion.
In sum, the limiting distributions of unit-root statistics depend on (a) the underlying data generating process and (b) the entertained model. For finite samples, one may encounter some difference in the power of test statistics.

Remark: If the Z_t process follows the model

$$Z_t = \phi_0 + Z_{t-1} + a_t,$$

where $\{a_t\}$ is defined as before and $\phi_0 \neq 0$, then the limiting distribution of the OLS $\hat{\phi}_1$ becomes normal. The reason is that all statistics involved are dominated by the time trend (slope ϕ_0) so that the usual central limit theorem applies. The convergence rate, however, is rather slow.

Augmented Dickey-Fuller Statistics
Suppose that the data generating process is an AR(p) model with a unit root. That is,

$$(1 - B)\phi(B)Z_t = a_t,$$

where $\{a_t\}$ is a sequence of $iid(0, \sigma_a^2)$ random variables and $\phi(B) = 1 - \gamma_1 B - \ldots - \gamma_p B^p$ with all zeros outside the unit circle. In theory, the limiting results discussed above hold if we entertain the model

$$Z_t = \phi Z_{t-1} + e_t,$$

in unit-root testing. However, the power of the t-ratio is reduced by the existence of the stationary AR component, which becomes a nuisance component in the testing. Dickey and Fuller suggested an *Augmented test statistic*

$$Z_t = \phi Z_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Z_{t-i} + e_t,$$

where $\Delta Z_{t-i} = Z_{t-i} - Z_{t-i-1}$ is the first differenced series of Z_t. One then employs the t-ratio of ϕ to perform the unit-root test. The limiting distribution of the t-ratio is the same as before. In other words, the stationary AR component has no effect on the limiting distribution of the t-ratio.
Alternatively, one can also emply the model

$$\Delta Z_t = \beta Z_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Z_{t-i} + e_t,$$

and consider the null hypothesis $H_o : \beta = 0$ versus the alternative hypothesis $H_a : \beta < 0$. It is easy to see that $\beta = \phi - 1$.

References

Chan, N. H. and C. Z. Wei (1988). Limiting Distribution of Least Squares Estimates of Unstable Autoregressive Processes. *Annals of Statistics*, 16, 367-401.

McLeish, D. L. (1975). A Maximal Inequality and Dependent Strong Laws. *Annals of Probability*, 3, 829-839.

Phillips, P.C.B. (1987). Time Series Regression with a Unit Root. *Economtrica*, 55, 277-301.

Demonstration: Consider the U.S. quarterly real GDP series from 1947.I to 2005.III. In R, the package "fSeries" contains several unit-root test statistics (based on the menu). But I only managed to get some results for the augmented Dickey-Fuller test. The name is **adfTest**. You may consult R for further information.

Splus output:

```
> x=read.table("gdp.txt")
> gdp=x[,4]
> x=log(gdp)

> mm=unitroot(x,lags=5)
> mm

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
```

```
 Test Statistic: -0.5703
        P-value: 0.8732

Coefficients:
    lag1    lag2    lag3    lag4    lag5 constant
 -0.0003  0.4163  0.1963 -0.1346 -0.0535  0.0120

Degrees of freedom: 230 total; 224 residual
Residual standard error: 0.009789

> mu=unitroot(x)
> mu

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -0.9976
        P-value: 0.7544

Coefficients:
    lag1    lag2    lag3    lag4    lag5    lag6    lag7
 -0.0005  0.3649  0.2031 -0.1116  0.0171 -0.1381  0.0791

    lag8    lag9   lag10   lag11 constant
  0.0725 -0.1129  0.1428  0.0847  0.0109

Degrees of freedom: 224 total; 212 residual
Residual standard error: 0.009349

> mu=unitroot(x,lags=4,trend="nc")
> mu

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: 5.918
        P-value: 1

Coefficients:
    lag1    lag2    lag3    lag4
```

```
  0.0010  0.4473  0.2060 -0.1319

Degrees of freedom: 231 total; 227 residual
Residual standard error: 0.009887

> mu=unitroot(x,lags=4,trend='ct')
> mu

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -0.8615
        P-value: 0.9573

Coefficients:
    lag1    lag2    lag3    lag4 constant    time
 -0.0051  0.4223  0.1914 -0.1495  0.0371   0.0204

Degrees of freedom: 231 total; 225 residual
Residual standard error: 0.009771

=== First differenced series ***

> y=diff(x)
>
> mu=unitroot(y,lags=4)
> mu

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -6.967
        P-value: 2.425e-9

Coefficients:
    lag1    lag2    lag3    lag4 constant
 -0.5721 -0.0106  0.1865  0.0524  0.0096

Degrees of freedom: 230 total; 225 residual
Residual standard error: 0.009774
```