The Unit Root Land

Material used and recommended:

* Hamilton (1994) "Time Series Analysis"
  chapters 15, 16, 17

* Fuller (1996) "Introduction to Statistical Time Series"
  chapter 10

* Billingsley (1999) "Convergence of Probability Measures"
  chapter 3

* Mikosch (2000) "Elementary Stochastic Calculus"
  chapters 1–2

* Phillips – "Lecture Notes"

Reasons to study Unit root Models

$X_t = X_{t-1} + \varepsilon_t$ $\quad \varepsilon_t \sim iid$ Random Walk

$X_t = \mu + X_{t-1} + \varepsilon_t$ $\quad \varepsilon_t \sim iid$ Random Walk with Drift

* Growth
* The effect of a shock
* Trend-Cycle Decomposition
* Forecasting
* Asymptotic Results
* Testing for a Unit Root
  - Dickey-Fuller
  - Phillips-Perron
  - ....
* Problems of testing

Growth :

- $X_t = X_{t-1} + \varepsilon_t$

- $X_t = \mu + X_{t-1} + \varepsilon_t \rightarrow \Delta X_t = \mu + \varepsilon_t$

$$X_t = \underline{\mu t} + \sum_{s=1}^{t} \varepsilon_s$$

- $X_t = a + \underline{bt} + \varepsilon_t$

The effect of a shock:

$$\boxed{\frac{\partial X_{t+h}}{\partial \varepsilon_t}}$$

Transitory shock : $h \rightarrow \infty \quad \frac{\partial X_{t+h}}{\partial \varepsilon_t} = 0$

Permanent shock : $h \rightarrow \infty \quad \frac{\partial X_{t+h}}{\partial \varepsilon_t} \neq 0$

Examples : (1) $X_t = \rho X_{t-1} + \varepsilon_t \quad |\rho| < 1$

$$X_t = \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \cdots$$

$$X_{t+h} = \varepsilon_{t+h} + \rho \varepsilon_{t+h-1} + \rho^2 \varepsilon_{t+h-2} + \cdots + \underline{\rho^h} \varepsilon_t + \cdots$$

$$\frac{\partial X_{t+h}}{\partial \varepsilon_t} = \rho^h \longrightarrow 0 \quad \text{as } h \rightarrow \infty$$

(2) $X_t = X_{t-1} + \varepsilon_t$

$X_t = \varepsilon_t + \varepsilon_{t-1} + \cdots$

$X_{t+h} = \varepsilon_{t+h} + \varepsilon_{t+h-1} + \cdots + \varepsilon_t + \cdots$

$h \to \infty \quad \dfrac{\partial X_{t+h}}{\partial \varepsilon_t} = 1 \neq 0$

(3) $\Delta X_t = C(L)\varepsilon_t$

$= C(1)\varepsilon_t + (1-L)\tilde{C}(L)\varepsilon_t$

$X_t = C(1)\dfrac{\varepsilon_t}{1-L} + \tilde{C}(L)\varepsilon_t$

$X_{t+h} = C(1)\dfrac{\varepsilon_{t+h}}{1-L} + \tilde{C}(L)\varepsilon_{t+h}$

$\dfrac{\partial X_{t+h}}{\partial \varepsilon_t} = C(1) \neq 0$ if X_t has a unit root

(4) Think on $X_t = a + bt + \varepsilon_t$
where $\varepsilon_t = \Psi(L)a_t$ stationary

Trend-Cycle Decomposition

$$\Delta X_t = \mu + C(L)\,\varepsilon_t$$

$$\Delta X_t = \mu + C(1)\,\varepsilon_t + (1-L)\,\tilde{C}(L)\,\varepsilon_t$$

$$X_t = \mu t + C(1)\,\frac{\varepsilon_t}{\Delta} + \tilde{C}(L)\,\varepsilon_t$$

Deterministic Trend

Stochastic Trend

Cycle

or

Permanent Component

Transitory Component

Remember the conditions on the polynomial $C(L)$ in order for $\tilde{C}(L)$ to behave correctly (???)

Forecasting

Comparison between an $I(0)$ and an $I(1)$

$$\boxed{X_t = a + bt + C(L)\varepsilon_t} \quad \text{Trend stationary}$$

$$\sum_{j=1}^{\infty} |c_j| < \infty$$

$$X_{t+h} = a + b(t+h) + \varepsilon_{t+h} + C_1 \varepsilon_{t+h-1} + \dots + C_h \varepsilon_t + \dots$$

$$E[X_{t+h} | I_t] = a + b(t+h) + C_h \varepsilon_t + C_{h+1} \varepsilon_{t-1} + \dots$$

$$E\left[X_{t+h} - E[X_{t+h} | I_t]\right]^2 \longrightarrow \underline{\text{constant}} \text{ as } h \longrightarrow \infty$$

Now

$$\boxed{X_t = \mu + X_{t-1} + C(L)\varepsilon_t} \quad \text{Difference stationary}$$

$$X_{t+h} = h\mu + X_t + \varepsilon_{t+h} + (1 + C_1)\varepsilon_{t+h-1} + \dots + (1 + C_1 + \dots C_{h-1})\varepsilon_{t+1} + \dots$$

$$E[X_{t+h} | I_t] = h\mu + X_t + [1 + C_1 + \dots + C_h]\varepsilon_t + [\dots]\varepsilon_{t-1}$$

$$E\left[X_{t+h} - E[X_{t+h} | I_t]\right]^2 \longrightarrow \underline{\underline{\infty}} \quad \text{as } h \to \infty$$

Asymptotic Results

$$X_t = \rho X_{t-1} + \varepsilon_t \qquad \underset{iid}{\varepsilon_t \sim N(0, \sigma^2)}$$

If $|\rho| < 1$

$$\sqrt{T}\left(\hat{\rho}_T - \rho\right) \sim N(0, 1-\rho^2)$$

$$\sqrt{T}\left(\hat{\rho}_T - 1\right) \xrightarrow{P} 0$$

If $\rho = 1$

$$\boxed{\left(\hat{\rho}_T - 1\right) = \frac{\sum_1^T X_{t-1}\varepsilon_t}{\sum_1^T X_{t-1}^2}}$$

$$\boxed{\frac{1}{T}\sum_1^T X_{t-1}\varepsilon_t}$$

$$X_t^2 = (X_{t-1} + \varepsilon_t)^2 = X_{t-1}^2 + 2X_{t-1}\varepsilon_t + \varepsilon_t^2$$

$$X_{t-1}\varepsilon_t = \frac{1}{2}\left\{X_t^2 - X_{t-1}^2 - \varepsilon_t^2\right\}$$

$$\sum_1^T X_{t-1}\varepsilon_t = \frac{1}{2}\left\{X_T^2 - \underset{\overset{\shortparallel}{0}}{X_0^2}\right\} - \frac{1}{2}\sum \varepsilon_t^2$$

$$\frac{1}{\sigma^2 T}\sum_1^T X_{t-1}\varepsilon_t = \frac{1}{2}\underbrace{\left(\frac{X_T}{\sigma\sqrt{T}}\right)^2}_{\chi^2_{(1)}} - \frac{1}{2\sigma^2}\underbrace{\frac{\sum_1^T \varepsilon_t^2}{T}}_{\to\ \sigma^2}$$

$$\frac{X_T}{\sigma\sqrt{T}} \overset{d}{\sim} N(0,1)$$

So $\frac{1}{\sigma^2}\frac{1}{T}\sum_1^T X_{t-1}\varepsilon_t \xrightarrow{d} \frac{1}{2}\left(\chi^2_{(1)}-1\right)$

$$\boxed{\frac{1}{T^2}\sum_1^T X_{t-1}^2}$$

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

$$E\left(\sum_1^T X_{t-1}^2\right) = E\left\{\sum_{t=1}^{T-1}\left(\sum_{i=1}^{t}\varepsilon_i\right)^2\right\}$$

$$= \frac{1}{2}T(T-1)\sigma^2$$

$$V\left(\sum_1^T X_{t-1}^2\right) = \frac{1}{3}T(T-1)(T^2-T+1)\sigma^4$$

$$\frac{\sum_1^T X_{t-1}^2}{.5\,T(T-1)\sigma^2}\;\begin{cases} E[\quad] = 1 \\ V[\quad] = \dfrac{4(T^2-T+1)}{3(T^2-T)} \xrightarrow[T\to\infty]{} \boxed{\dfrac{4}{3}} \end{cases}$$

So $\frac{1}{T^2}\sum_1^T X_{t-1}^2$ doesn't converge to a constant in mean square.

[Notice the difference with the case $\rho=1$]

Brownian Motion

Def: B is a stochastic process $\{B(t): 0 \le t < \infty\}$ on a probability space (Ω, F, P) with properties:

(a) $B(0, \omega) = 0 \quad \forall \omega$

(b) $B(\cdot, \omega) = 0$ is continuous for each ω

(c) For $0 < t_1 < t_2 < \ldots < t_{n-1} < t_n$, the incrementes $B(t_1), B(t_2) - B(t_1), \ldots B(t_n) - B(t_{n-1})$ are independent and normally distributed, with means 0 and variances $t_1, t_2 - t_1, \ldots, t_n - t_{n-1}$.

Def: A Brownian motion is a gaussian process with $E(B(t)) = 0$ and $Cov(B(s), B(t)) = s \wedge t$

[For a graphical view see the Appleb on my web page]

Construction of a Brownian motion

Let $\{u_t\}_{t=0}^{\infty}$ be a stochastic process

such that:

(1) $E(u_t) = 0$ for all t

(2) $\sup E|u_t|^{\beta} < \infty$ for some $\beta > 2$

(3) $\sigma^2 = \lim_{t \to \infty} E[T^{-1} S_T^2]$ exists and $\sigma^2 > 0$;

$S_T = \sum_{i=1}^{T} u_i$ and σ^2 can also be

written as $\boxed{\sigma^2 = \sigma_u^2 + 2\lambda}$ with

$\sigma_u^2 = E(u_1^2)$ and $\lambda = \sum_{j=2}^{\infty} E(u_1 u_j)$

(4) u_t is strongly-mixing with mixing coefficients α_m such that $\sum_1^{\infty} \alpha_m^{(1-2/\beta)} < \infty$.

Let's consider $X_T = \frac{1}{\sqrt{T}} S_T = \frac{1}{\sqrt{T}} \sum_{i=1}^{T} u_i$

Notice that

$$\frac{1}{\sqrt{T}} S_T \Rightarrow N(0, \sigma^2)$$

$$\frac{1}{\sqrt{[T/2]}} \sum_{i=1}^{[T/2]} u_i \Rightarrow N(0, \sigma^2)$$

where $[T/2]$ denotes the largest integer that is less or equal to $T/2$.

Now consider the following variable (process)

$$X_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} u_t \qquad r \in [0,1]$$

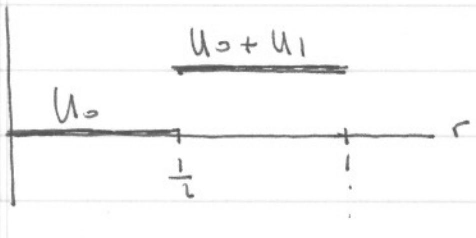
For any given realization, $X_T(r)$ is a step function in r, with

$$
\begin{aligned}
X_T(r) &= 0 && \text{for} \quad 0 \le r < 1/T \\
&= u_1/\sqrt{T} && \text{for} \quad 1/T \le r < 2/T \\
&= (u_1+u_2)/\sqrt{T} && \text{for} \quad 2/T \le r < 3/T \\
&\ \ \vdots && \qquad\quad \vdots \\
&= (u_1+u_2+\cdots+u_T)/\sqrt{T} && \text{for} \quad r = 1
\end{aligned}
$$

$T=1$

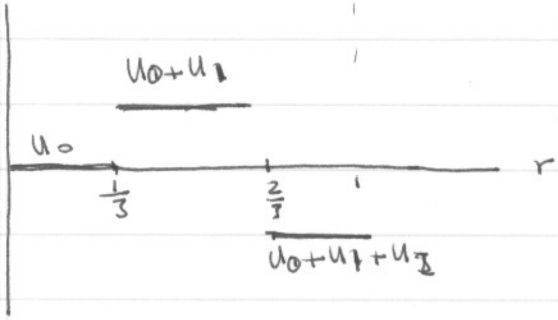
$$X_1(r) = \frac{1}{\sqrt{1}} \sum_{t=1}^{[1r]} u_t = \frac{1}{\sqrt{1}} u_0 = 0$$

$T=2$



$$X_2(r) = \frac{1}{\sqrt{2}} \sum_{t=1}^{[2r]} u_t \begin{cases} \frac{1}{\sqrt{2}} u_0 & 0 \le r < \frac{1}{2} \\ \frac{1}{\sqrt{2}} (u_0+u_1) & \frac{1}{2} \le r < 1 \end{cases}$$

$T=3$



$$X_3(r) = \frac{1}{\sqrt{3}} \sum_{t=1}^{[3r]} u_t = \begin{cases} \frac{1}{\sqrt{3}} u_0 & 0 \le r < \frac{1}{3} \\ \frac{1}{\sqrt{3}} (u_0+u_1) & \frac{1}{3} \le r < \frac{2}{3} \\ \frac{1}{\sqrt{3}} (u_0+u_1+u_2) & \frac{2}{3} \le r < 1 \end{cases}$$

$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$

Then $X_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} u_t$

$$= \left(\sqrt{[Tr]}/\sqrt{T}\right)\left(\frac{1}{\sqrt{[Tr]}}\right) \sum_{t=1}^{[Tr]} u_t$$

But

$$\left(\frac{1}{\sqrt{[Tr]}}\right) \sum_{t=1}^{[Tr]} u_t \Longrightarrow N(0, \sigma^2) \quad \text{by the CLT}$$

$$\sqrt{Tr}/\sqrt{T} \longrightarrow \sqrt{r}$$

Therefore $X_T(r) \Longrightarrow N(0, r\sigma^2)$

and

$$\boxed{\frac{X_T(r)}{\sigma} \Longrightarrow N(0, r)}$$

In similar way $(r_2 > r_1)$

$$\boxed{\left(X_T(r_2) - X_T(r_1)\right)/\sigma \Longrightarrow N(0, r_2 - r_1)}$$

The sequence of stochastic functions $\{X_T(\cdot)/\sigma\}_{T=1}^{\infty}$ has an asymptotic probability law that is described by the Wiener process $W(\cdot)$

$$\boxed{X_T(\cdot)/\sigma \overset{L}{\Longrightarrow} W(\cdot)}$$

or

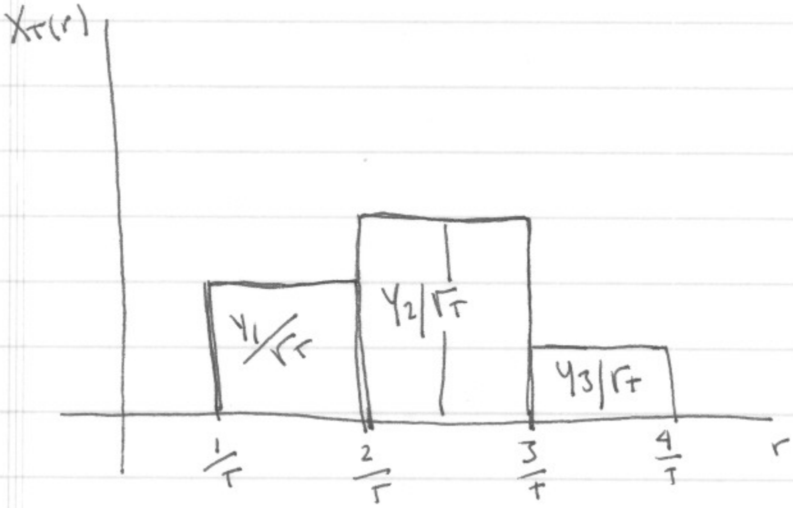
$\boxed{X_T(\cdot) \Longrightarrow \sigma^2 W(\cdot)}$ Functional Central Limit Theorem (see Appendix)

Application to Unit root processes

$$\boxed{Y_t = Y_{t-1} + u_t}$$

$Y_t = u_1 + u_2 + \cdots \; u_t$

$$
\begin{aligned}
X_T(r) &= 0 && \text{for} \quad 0 \le r < 1/T \\
&= Y_1/\sqrt{T} && \text{for} \quad 1/T \le r < 2/T \\
&= Y_2/\sqrt{T} && \text{for} \quad 2/T \le r < 3/T \\
&\;\vdots \\
&= Y_T/\sqrt{T} && \text{for} \quad r = 1
\end{aligned}
$$



$$
\begin{aligned}
\int_0^1 X_T(r)\,dr &= Y_1/T^{3/2} + Y_2/T^{3/2} + \cdots + Y_{T-1}/T^{3/2} \\
&= T^{-3/2} \sum_{t=1}^{T} Y_{t-1} \\
&\Rightarrow \boxed{\sigma \int_0^1 W(r)\,dr}
\end{aligned}
$$

By the Continuous mapping theorem

$$T^{-3/2} \sum_{t=1}^{T} Y_{t-1} \Longrightarrow \sigma \underbrace{\int_0^1 W(r)\, dr}_{\text{Integrated Brownian Motion}}$$

<u>Some Properties:</u>

$$E\left[\int_0^t W(s)\, ds\right] = \int_0^t E(W(s))\, ds = \underline{\underline{0}}$$

$s \le t$

$$Cov\left[\int_0^s W(y)\, dy \int_0^t W(u)\, du\right] =$$

$$= E\left[\int_0^s \int_0^t W(y)\, W(u)\; dy\, du\right]$$

$$= \int_0^s \int_0^t E\left[W(y)\, W(u)\right] dy\, du$$

$$= \int_0^s \int_0^t \min(y, u)\; dy\; du$$

$$= \int_0^s \left(\int_0^u y\, dy + \int_u^t u\, dy\right) du$$

$$= s^2\left(\frac{t}{2} - \frac{s}{6}\right)$$

If $t = s = 1$ then $Cov(\;) = \frac{1}{3}$

$$\underline{\underline{\text{So} \quad \int_0^1 W(r)\, dr \equiv N\left(0, \frac{1}{3}\right)}}$$

Other averages

$$T^{-3/2}\sum_{t=1}^{T} Y_{t-1} = T^{-3/2}\left[u_1 + (u_1+u_2) + (u_1+u_2+u_3) + \dots + (u_1+u_2+\dots+u_{T-1})\right]$$

$$= T^{-3/2}\left[(T-1)u_1 + (T-2)u_2 + (T-3)u_3 + \dots + [T-(T-1)]u_{T-1}\right]$$

$$= T^{-3/2}\sum_{t=1}^{T}(T-t)u_t$$

$$\boxed{= T^{-1/2}\sum_{t=1}^{T} u_t - T^{-3/2}\sum_{t=1}^{T} t u_t}$$

Therefore

$$T^{-3/2}\sum_{t=1}^{T} t u_t = T^{-\frac{1}{2}}\sum_{t=1}^{T} u_t - T^{-3/2}\sum_{t=1}^{T} Y_{t-1}$$

$$\Rightarrow \boxed{\sigma W(1) - \sigma\int_0^1 W(r)\,dr} \equiv N(0, \sigma^2/3)$$

$$\mathrm{Cov}\left(\int_0^t W(r)\,dr,\ W(t)\right) = \frac{t^2}{2}$$

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

Define $S_T(r) = [X_T(r)]^2$;

$$\begin{aligned} S_T(r) &= 0 && \text{for} \quad 0 \le r < 1/T \\ &= Y_1^2/T && \text{for} \quad 1/T \le r < 2/T \\ &= Y_2^2/T && \text{for} \quad 2/T \le r < 3/T \\ &\vdots \\ &= Y_T^2/T && \text{for} \quad r = 1 \end{aligned}$$

It follows that

$$\int_0^1 (X_T(r))^2 \, dr = \frac{Y_1^2}{T^2} + \frac{Y_2^2}{T^2} + \cdots + \frac{Y_{T-1}^2}{T^2}$$

$$\Rightarrow \sigma^2 \int_0^1 (W(r))^2 \, dr$$

So $$\boxed{T^{-2} \sum_{t=1}^{T} Y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 (W(r))^2 \, dr}$$

Other useful results:

$$T^{-\frac{5}{2}} \sum_{t=1}^{T} t Y_{t-1} = \underbrace{T^{-\frac{3}{2}} \sum_{t=1}^{T} (t/T) Y_{t-1}}_{= \xi} \Rightarrow \boxed{\sigma \int_0^1 r \, W(r) \, dr}$$

$$T^{-3} \sum_{t=1}^{T} t Y_{t-1}^2 = T^{-2} \sum_{t=1}^{T} (t/T) Y_{t-1}^2 \Rightarrow \boxed{\sigma^2 \int_0^1 r (W(r))^2 \, dr}$$

$$T^{-1}\sum_{t=1}^{T} Y_{t-1}u_t \overset{(*)}{=} \left(\frac{1}{2}\right)\left(\frac{1}{T}\right)Y_T^2 - \left(1/2\right)\left(\frac{1}{T}\right)\left(u_1^2+u_2^2+\cdots+u_T^2\right)$$

$$= \frac{1}{2}\underbrace{X_T^2(1)}_{\Downarrow\ \sigma^2[W(1)]^2} - \left(\frac{1}{2}\right)\underbrace{\left(\frac{1}{T}\right)\left(u_1^2+u_2^2+\cdots+u_T^2\right)}_{\xrightarrow{LLN}\ \underline{\sigma^2}}$$

Therefore

$$\boxed{\frac{1}{T}\sum_{t=1}^{T} Y_{t-1}u_t \Rightarrow \frac{1}{2}\sigma^2\left[W(1)^2-1\right]}$$

(*) $\frac{1}{T}\sum_{1}^{T} y_{t-1}u_t = \left(\frac{1}{T}\right)\Big\{(u_1)(u_2) + (u_1+u_2)u_3 + (u_1+u_2+u_3)u_4 + \cdots$
$+ (u_1+u_2+u_3+\cdots+u_{T-1})u_T\Big\}$

Notice that

$$\left(\frac{1}{T}\right)y_T^2 = \left(\frac{1}{T}\right)\left(u_1+u_2+\cdots+u_T\right)^2$$

$$=\left(\frac{1}{T}\right)\Big\{\left(u_1^2+u_2^2+\cdots+u_T^2\right)+2(u_1u_2)+2(u_1u_3)+\cdots+2(u_1u_T)$$
$$+2(u_2u_3)+2(u_2u_4)+\cdots+2(u_{T-1}u_T)\Big\}$$

Therefor

$$\left(\frac{1}{T}\right)y_T^2 = \frac{1}{T}\left(u_1^2+u_2^2+\cdots+u_T^2\right)+2\left(\frac{1}{T}\right)\sum_{1}^{T} y_{t-1}u_t$$

$\frac{y_T}{\sqrt{T}} \sim N(0,1)$

Summary of important results

$$\boxed{X_t = X_{t-1} + u_t} \quad \text{where} \quad u_t \sim iid \quad \begin{array}{l} E(u_t) = 0 \\ V(u_t) = \sigma^2 \end{array}$$

(a) $T^{-\frac{1}{2}} \sum_{t=1}^{T} u_t \Rightarrow \sigma W(1)$

(b) $T^{-1} \sum_{t=1}^{T} X_{t-1} u_t \Rightarrow \frac{1}{2} \sigma^2 \{[W(1)]^2 - 1\} = \boxed{\int_0^1 W \, dW}$

(c) $T^{-\frac{3}{2}} \sum_{t=1}^{T} t \, u_t \Rightarrow \sigma W(1) - \sigma \int_0^1 W(r) \, dr \equiv N(0, \sigma^2/3)$

(d) $T^{-3/2} \sum_{t=1}^{T} X_{t-1} \Rightarrow \sigma \int_0^1 W(r) \, dr$

(e) $T^{-2} \sum_{t=1}^{T} X_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 (W(r))^2 \, dr$

(f) $T^{-5/2} \sum_{t=1}^{T} t \, X_{t-1} \Rightarrow \sigma \int_0^1 r \, W(r) \, dr$

(g) $T^{-3} \sum_{t=1}^{T} t \, X_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 r (W(r))^2 \, dr$

(h) $T^{-(v+1)} \sum_{t=1}^{T} t^v \xrightarrow{p} 1/(v+1)$ for $v = 0, 1, \ldots$

DICKEY-FULLER TESTS for UNIT ROOT

Case (a) DGP $X_t = X_{t-1} + u_t$

Regression $X_t = \rho X_{t-1} + u_t$ $\boxed{t_{\hat{\rho}}}$

Case (b) DGP $X_t = X_{t-1} + u_t$

Regression $X_t = a + \rho X_{t-1} + u_t$ $\boxed{t_{\hat{\rho}}}$

Case (b') DGP $X_t = a + X_{t-1} + u_t$

Regression $X_t = a + \rho X_{t-1} + u_t$ $\boxed{t_{\hat{\rho}}}$

Case (c) DGP $X_t = X_{t-1} + u_t$

Regression $X_t = a + bt + \rho X_{t-1} + u_t$ $\boxed{t_{\hat{\rho}}}$

Case (c') DGP $X_t = a + X_{t-1} + u_t$

Regression $X_t = a + bt + \rho X_{t-1} + u_t$ $\boxed{t_{\hat{\rho}}}$

Source: Fuller (1976)

Unit Root Dist^n Tables

$T(\hat\rho - 1)$

Table 8.5.1. Empirical cumulative distribution of $n(\hat\rho - 1)$ for $\rho = 1$

Sample Size n	Probability of a Smaller Value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				$\hat\rho$				
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
				$\hat\rho_\mu$				
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
				$\hat\rho_\tau$				
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

NOTE. This table was constructed by David A. Dickey using the Monte Carlo method. Details are given in Dickey (1975). Standard errors of the estimates vary, but most are less than 0.15 for entries in the left half of the table and less than 0.03 for entries in the right half of the table.

Although the sign of e_{t-j} in the weighted sum $\sum_{j=0}^{t-1}(-1)^j e_{t-j}$ is not the same for all t, the sign is always opposite of that for e_{t-j-1} and e_{t-j+1}, and it follows that

$$\sum_{t=1}^{n-1} X_t^2 = \sum_{t=1}^{n-1}\left(\sum_{j=1}^{t}(-1)^j e_j\right)^2.$$

The distribution of e_t, $t = 1, 2, \ldots$, is symmetric and hence the distributional properties of the sequence $-e_1, e_2, -e_3, e_4, \ldots$, are precisely the same as the

$\hat\tau$

Table 8.5.2. Empirical cumulative distribution of $\hat\tau$ for $\rho = 1$

Sample Size n	Probability of a Smaller Value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				$\hat\tau$	Random walk			
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
				$\hat\tau_\mu$	Random walk + drift			
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
				$\hat\tau_\tau$	Random walk with drift and time-trend			
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

This table was constructed by David A. Dickey using the Monte Carlo method. Details are given in Dickey (1975). Standard errors of the estimates vary, but most are less than 0.02.

To extend the results for the first order process with $\rho = 1$ to the pth order autoregressive process, we consider the time series

$$Y_t = \sum_{j=1}^{t} Z_j, \qquad t = 1, 2, \ldots, \tag{8.5.11}$$

where $\{Z_t: t \in (0, \pm 1, \pm 2, \ldots)\}$ is a $(p-1)$ order autoregressive time series with the representation

$$Z_t + \sum_{i=2}^{p} a_i Z_{t-i+1} = e_t. \tag{8.5.12}$$

Case (a) DGP : $X_t = X_{t-1} + u_t$ Random Walk

Regression: $X_t = \rho X_{t-1} + u_t$ $u_t \sim iid$ $E(u_t) = 0$

$V(u_t) = \sigma^2$

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} X_{t-1} X_t}{\sum_{t=1}^{T} X_{t-1}^2} = \rho + \frac{\boxed{T^{-1}} \sum_{t=1}^{T} X_{t-1} u_t}{\boxed{T^{-2}} \sum_{t=1}^{T} X_{t-1}^2}$$

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

We know

$$T^{-1} \sum X_{t-1} u_t \Longrightarrow \frac{1}{2} \sigma^2 \left\{ (W(1))^2 - 1 \right\}$$

$$T^{-2} \sum X_{t-1}^2 \Longrightarrow \sigma^2 \int_0^1 (W(r))^2 dr$$

Therefore

$$\boxed{T\left(\hat{\rho}_T - 1\right) \Longrightarrow \frac{\frac{1}{2}\left\{ W(1)^2 - 1 \right\}}{\int_0^1 (W(r))^2 dr}}$$

$$\boxed{\sqrt{T}\left(\hat{\rho} - \rho\right) \Longrightarrow N(0, 1-\rho^2)} \quad \text{when } |\rho| < 1$$

2013

The Dickey-Fuller statistic is

$$\boxed{t\,\hat{\rho}}$$

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}} = \boxed{\frac{(\hat{\rho}_T - 1)}{\left\{ S_T^2 \Big/ \sum_{t=1}^{T} X_{t-1}^2 \right\}^{\frac{1}{2}}}}$$ D-F test

where $S_T^2 = \frac{1}{T-1} \sum_{t=1}^{T} \left(X_t - \hat{\rho}_T X_{t-1} \right)^2$

To find the A.D of t_T we need to rewrite t_T as:

$$t_T = T(\hat{\rho}_T - 1) \left\{ T^{-2} \sum_{t=1}^{T} X_{t-1}^2 \right\}^{1/2} \frac{1}{\{S_T^2\}^{\frac{1}{2}}}$$

$$= \frac{T^{-1} \sum_{1}^{T} X_{t-1} u_t}{\left\{ T^{-2} \sum_{1}^{T} X_{t-1}^2 \right\}^{1/2} \{ S_T^2 \}^{\frac{1}{2}}}$$

$$\Rightarrow \frac{\frac{1}{2} \sigma^2 \left\{ [W(1)]^2 - 1 \right\}}{\left\{ \sigma^2 \int_0^1 (W(r))^2 dr \right\}^{\frac{1}{2}} \{ \sigma^2 \}^{\frac{1}{2}}} =$$

$\boxed{S_T^2 \xrightarrow{P} \sigma^2 \text{ by Consistency of } \hat{\rho}_T}$

$$\boxed{= \frac{1/2 \left\{ [W(1)]^2 - 1 \right\}}{\left\{ \int_0^1 W(r)^2 dr \right\}^{\frac{1}{2}}}}$$ D-F distribution

OTHER THINGS TO KNOW

Case (b') DGP: $X_t = \alpha + X_{t-1} + u_t$

Regression $X_t = a + \rho X_{t-1} + u_t$

$$X_t = X_0 + \alpha t + (u_1 + u_2 + \dots + u_t)$$
$$= X_0 + \alpha t + \xi_t$$

$$\sum_1^T X_{t-1} = \sum_1^T [X_0 + \alpha(t-1) + \xi_{t-1}]$$

$$\boxed{T^{-1}} \to \underline{\underline{X_0}} \qquad \boxed{T^{-2}} \to \underline{\underline{\alpha/2}} \qquad \boxed{T^{-3/2}} \to \underline{\underline{\sigma \int_0^1 W(r)\,dr}}$$

So we need to divide $\sum_1 X_{t-1}$ by $\boxed{T^2}$

to get something that "makes sense"

$$\boxed{T^{-2} \sum_1^T X_{t-1}} = T^{-1} X_0 + T^{-2} \sum_1^T \alpha(t-1) + T^{-1/2} \left\{ T^{-3/2} \sum_1^T \xi_{t-1} \right\}$$

$$\xrightarrow{p} 0 + \alpha/2 + 0 = \boxed{\alpha/2}$$

Similarly

$$\sum_1^T X_{t-1}^2 = \underbrace{\sum_1^T X_0^2}_{O_p(T)} + \underbrace{\sum_1^T \alpha^2 (t-1)^2}_{O_p(T^3)} + \underbrace{\sum_1^T \xi_{t-1}^2}_{O_p(T^2)} + \underbrace{2 \sum 2 y_0 \alpha (t-1)}_{O_p(T^2)} + \underbrace{\sum 2 y_0 \xi_{t-1}}_{O_p(T^{3/2})} + \underbrace{\sum 2\alpha(t-1)\xi_{t-1}}_{O_p(T^{5/2})}$$

so

$$\boxed{T^{-3} \sum_1^T X_{t-1}^2 \xrightarrow{p} \alpha^2/3}$$

Finally

$$\sum_1^T X_{t-1} u_t = \underbrace{X_0 \sum_1^T u_t}_{O_p(T^{1/2})} + \underbrace{\sum_1^T \alpha (t-1) u_t}_{O_p(T^{3/2})} + \underbrace{\sum_1^T \xi_{t-1} u_t}_{O_p(T)}$$

so

$$\boxed{T^{-3/2} \sum_1^T X_{t-1} u_t \Rightarrow T^{-3/2} \sum_1^T \alpha (t-1) u_t}$$

Putting every piece together

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

where

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

$$\boxed{\text{so } t_{\hat{\rho}} \ ???}$$

Case (c) and (c'): DGP $X_t = \cancel{\alpha} + X_{t-1} + u_t$

Regression $X_t = a + bt + \rho X_{t-1} + u_t$

$t_{\hat{\rho}} \Longrightarrow$ D.F distribution that DOES NOT DEPEND ON α or σ

Think on a strategy to test for unit roots.

What if u_t is correlated

$$X_t = X_{t-1} + u_t \quad \text{and} \quad u_t = \Psi(L)\, e_t$$

with $e_t \sim \underline{\underline{iid}}$

NOTE:

(1) $X_t = \rho X_{t-1} + u_t \qquad H_0: \rho = 1 \qquad H_A: |\rho| < 1$

$$\Delta X_t = (\rho - 1) X_{t-1} + u_t \qquad H_0: (\rho-1) = 0 \qquad H_A: (\rho-1) < 0$$

(2)

$$\Psi(L)^{-1} = A(L) = A(1) + (1-L)\, A^*(L)$$

$$A(0) = 1$$

(3)

$$\Delta X_t = \rho X_{t-1} + \Psi(L)\, e_t$$

(4)

$$\Psi^{-1}(L)\, \Delta X_t = (\rho-1)\Psi^{-1}(L) X_{t-1} + e_t$$

or

$$\Delta X_t = (\rho-1) A(1) X_{t-1} + (\rho - 1)\, \Delta A^{+}(L) X_{t-1}$$

$$+ (-A(L) + A(0))\, \Delta X_t + e_t$$

(5)

$$\Delta X_t = a^* X_{t-1} + \text{LAGS of } \Delta X_t + e_t$$

(6) $\boxed{t_{\hat{a}^*}}$ has the same AD as

$t_{\hat{\rho}}$ in case (a)

DF and deterministic terms

(1) $$X_t = \rho X_{t-1} + u_t$$

$$X_t = a + \rho X_{t-1} + u_t$$

$$X_t = a + bt + \rho X_{t-1} + u_t$$

(2) $$X_t = \begin{pmatrix} \text{polynomial} \\ \text{in } t \end{pmatrix} + \rho X_{t-1} + u_t$$

Under the H_0 of $X_t = X_{t-1} + u_t$

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}} \Rightarrow \frac{\int_0^1 W(r)\, dW(r)}{\left(\int_0^1 W^2(r)\, dr\right)^{\frac{1}{2}}} \quad \text{in (1)}$$

In general (2)

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}} \Rightarrow \frac{\int_0^1 W_D(r)\, dW(r)}{\left(\int_0^1 W_D^2(r) \right)^{\frac{1}{2}}}$$

where $W_D(r) = W(r) - \left[\int_0^1 W D'\right]\left[\int_0^1 DD'\right]^{-1} D(r)$ is the

Hilbert projection in $L_2[0,1]$ of W onto the space orthogonal to D. $D(r) = (1, r, \dots r^p)'$

Phillips - Perron tests (Z_α and Z_t)

$$X_t = \rho X_{t-1} + u_t \qquad \text{with} \qquad u_t = C(L)\varepsilon_t$$

$$T(\hat{\rho} - 1) = \frac{\frac{1}{T}\sum_1^T X_{t-1} u_t}{\frac{1}{T^2}\sum_1^T X_{t-1}^2} \Rightarrow \frac{\left[\int B(r)\, dB(r) + \lambda\right]}{\left[\int B^2(r)\, dr\right]}$$

$$t_{\hat{\rho}} \Rightarrow \frac{(\hat{\rho}_T - 1)}{\left(\hat{\sigma}_u^2 / \sum_1^T X_{t-1}^2\right)^{\frac{1}{2}}} \Rightarrow \frac{\int_0^1 B(r)\, dB(r) + \lambda}{\sigma_u \left[\int_0^1 B^2(r)\, dr\right]^{\frac{1}{2}}}$$

where

$\sigma_u^2 = V(u_t)$, $B(r)$ is a BM with variance $w^2 = \sigma_\varepsilon^2 C(1)^2$

and $\lambda = \sum_{j=1}^{\infty} E(u_0 u_j)$. $\boxed{w^2 = \sigma_u^2 + 2\lambda}$

《THINK ON WHAT DOES IT HAPPEN IF $u_t = \varepsilon_t$?》

$$Z_\alpha = T(\hat{\alpha} - 1) - \frac{\hat{\lambda}}{\left(\frac{1}{T^2}\sum_1^T X_{t-1}^2\right)} \Rightarrow \frac{\int_0^1 W\, dW}{\int W^2\, dr}$$

$$Z_t = \hat{\sigma}_u \hat{w}^{-1} t_\alpha - \hat{\lambda}\left\{\hat{w}\left(T^{-2}\sum_1^T X_{t-1}^2\right)^{\frac{1}{2}}\right\}^{-1} \Rightarrow \frac{\int W\, dW}{\left(\int W^2\, dr\right)^{\frac{1}{2}}}$$

where $W(r) = \frac{1}{w} B(r)$.

Efficient Unit Root Tests

- When there is no trend/intercept in the model, the DF test is very close to being optimal - i.e for local alternatives it comes close to having optimal power in relation to the best test computed for the known local alternative using the Neyman-Pearson Lemma (this is the so-called power envelope) and is based on the likelihood ratio

$$\frac{L_{H_0}(\rho=1)}{L_{H_1}(\rho=1+\frac{c}{T}, \text{given } c)}$$

- However, when there is a trend in the model the DF rely on trend removal by regression and it turns out that efficiency can be gained by "improving" the trend removal process

$$y_t = \beta' D_t + X_t$$

with $\underbrace{D_t = (1, t, ...)}_{\text{Deterministic components}}$ and $X_t = \rho X_{t-1} + \varepsilon_t$.

- Under local alternative $\rho = 1 + \frac{c}{T}$
- You could use OLS to estimate the deterministic components. This is what DF does
- It makes sense to think that if we find a more efficient method to estimate β, we would get better power
- Elliot, Rothember, Stock (1996, Econometrica) proposes to use GLS to estimate β under H_1, so for different values of "c".

$$\tilde{\beta} = (\tilde{D}'\tilde{D})^{-1} \tilde{D}' \tilde{y}$$

where
$$\tilde{D}_t = D_t - \rho D_{t-1} = (1 - \rho L) D_t$$
$$\tilde{y}_t = y_t - \rho y_{t-1} = (1 - \rho L) y_t$$

So $\tilde{y}_t = \tilde{D}_t \beta + X_t$

$$\tilde{X}_t = \tilde{y}_t - \tilde{D}_t \tilde{\beta} = \text{detrended data}$$

Now unit root test on $\underline{\underline{\tilde{X}_t}}$.

The AD of the DF on the GLS detrending data depends on $\underline{\underline{c}}$.

For $D = (1, t)$, ERS advises to use $c = 13.5$ and they provide the CVs for that.

Additional Aspects of Unit Root tests

* It is comonly claimed that UR tests have serious
  - size problems (PP tests)
  - power problems (D-F tests)

My opinion is in

"No Lack of Relative Power of the DF Tests" JTSA (1996), 17, 37-47 by J. Gonzalo and T. Lee

* Why the null is unit root (non-stationarity) instead of |root| < 1 (stationarity) ?

  Think on terms of

  type I - II errors.

To solve the size problems of the PP tests, Perron and Ng (1996)

"Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties" RES 63, 435-63.

propose some modifications of the PP tests.

$$\boxed{MZ_\alpha = Z_\alpha + \frac{1}{2} T(\hat{\rho}_T - 1)^2}$$

Under H_0 $MZ_\alpha \overset{d}{=} Z_\alpha$

Define $MSB = \left(T^{-2} \sum_1^T X_{t-1}^2 / \hat{\omega}^2\right)^{\frac{1}{2}}$.

It can be checked that

$$Z_t = MSB . Z_\alpha$$

This suggests a new modified PP

$$MZ_t = MSB . MZ_\alpha$$

$$\boxed{MZ_t = Z_t + \frac{1}{2}\left(\frac{\sum_1^T X_{t-1}^2}{\hat{\omega}^2}\right)^{\frac{1}{2}}(\hat{\rho}_T - 1)^2}$$

Under H_0 $MZ_t \overset{d}{=} Z_t$

Think on how to implement it.

Testing Stationarity

H_0: no unit root

H_a: unit root

$$y_t = D_t + X_t + v_t$$

$$X_t = X_{t-1} + u_t$$

$$H_0: \sigma_u^2 = 0$$

$$H_a: \sigma_u^2 > 0$$

Let $\hat{e}_t$ be the residuals from the regression of y_t on D_t and $\hat{\sigma}_v^2 = \frac{1}{T}\sum_1^T \hat{e}_t^2$, then the LM statistic can be constructed as follows

$$LM = \frac{\frac{1}{T^2}\sum_1^T S_t^2}{\hat{\sigma}_v^2}$$

where S_t is the partial sum process of the residuals $\sum_{j=1}^{t} \hat{e}_j^2$.

Under H_0, this $LM \Rightarrow \int_0^1 V_{(r)}^2 dr$, where $V(r)$ is a Brownian Bridge $(B(r) - rB(1))$