CH3 Nonstationary AR Process

Some nonstationary processes
Winner process
Limiting theorem on random walk
Application to regression models
Unit root test
Steps of unit root test

Trend and Random Walk

"**Trend**": permanent or nondecaying component of a time series. The trend has a permanent effect on a series.

"Trend stationary": $y_t = y_0 + a_0 t + A(L)\varepsilon_t$, where $A(L)\varepsilon_t$ is a stationary component of y_t.

"Not trend stationary": $y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i + a_0 t$. $\sum_{i=1}^{t} \varepsilon_i$ is a stochastic trend component of y_t. Each ε_i shock has a permanent change in the conditional mean of $\{y_t\}$.

Random walk model: $y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i \qquad y_t = y_{t-1} + \varepsilon_t$, where ε_t is a white-noise.

Nonstationary and Difference stationary

An ε_i shock $(i < t)$ has a permanent effect on y_t and hence the forecasts for y_{t+s}.

$$\boxed{\begin{aligned} Ey_t &= Ey_{t-s} = y_0, \\ Var(y_t) &= t\sigma^2,\ Var(y_{t-s}) = (t-s)\sigma^2, \\ E_t y_{t+1} &= E_t y_{t+s} = y_t,\ s > 0. \end{aligned}}$$

$$Cov(y_t, y_{t-s}) = Cov\left(\sum_{i=1}^{t} \varepsilon_i, \sum_{i=1}^{t-s} \varepsilon_i\right) = (t-s)\sigma^2,$$

$$\rho_s = \frac{(t-s)\sigma^2}{\sqrt{t\sigma^2}\sqrt{(t-s)\sigma^2}} = (1 - s/t)^{1/2},$$

as s increases, ρ_s will decline, and hence the ACF for a random walk process will show a slight tendency to decay. It is impossible to use the ACF to distinguish between a unit root process and a stationary process with a near-unit root.

Other random walk models

Random walk plus drift model: $y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i + a_0 t \qquad y_t = a_0 + y_{t-1} + \varepsilon_t,$

It is difference-stationary. $\quad E_t y_{t+s} = y_t + a_0 s.$

Random walk plus noise model:

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i + \eta_t \qquad y_t = y_{t-1} + \varepsilon_t + \Delta\eta_t,$$

where η_t is a white-noise with variance σ_η^2 and independent of ε_{t-s} for all t and s; $\eta_0 = 0$.

It is Nonstationary but Difference-stationary.

$$\begin{aligned} Ey_t &= Ey_{t-s} = y_0, \\ Var(y_t) &= t\sigma^2 + \sigma_\eta^2,\ Var(y_{t-s}) = (t-s)\sigma^2 + \sigma_\eta^2, \\ E_t y_{t+1} &= E_t y_{t+s} = y_t,\ s > 0. \end{aligned}$$

$$Cov(y_t, y_{t-s}) = Cov\left(\sum_{i=1}^{t} \varepsilon_i + \eta_t, \sum_{i=1}^{t-s} \varepsilon_i + \eta_{t-s}\right) = (t-s)\sigma^2,$$

$$\rho_s = \frac{(t-s)\sigma^2}{\sqrt{t\sigma^2 + \sigma_\eta^2}\sqrt{(t-s)\sigma^2 + \sigma_\eta^2}},$$

Other random walk models

Trend plus noise model:

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i + \eta_t \qquad\qquad y_t = a_0 + y_{t-1} + \varepsilon_t + \Delta\eta_t,$$

where η_t is a white-noise process with variance σ_η^2 and $E\varepsilon_t\eta_{t-s} = \mathbf{0}$ for all t and s. Nonstationary: a linear deterministic trend $a_0 t$, a stochastic trend $\sum_{i=1}^{t} \varepsilon_i$, and a pure white-noise η_t. It is difference-stationary.

General trend plus irregular model:

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i + A(L)\eta_t \qquad y_t = a_0 + y_{t-1} + \varepsilon_t + A(L)\Delta\eta_t,$$

where $A(L)\eta_t$ is a stationary process. Nonstationary: a linear deterministic trend $a_0 t$, a stochastic trend $\sum_{i=1}^{t} \varepsilon_i$, and a stationary component $A(L)\eta_t$. Shocks to a stationary series are necessarily temporary; the effects of the irregular component will dissipate and do not affect its long-run mean level. But the trend components will determine the trend of the y_t process.

Remove the trend

1) Differencing:

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i + a_0 t \qquad \Delta y_t = a_0 + \varepsilon_t,$$

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i + \eta_t \qquad \Delta y_t = \varepsilon_t + \Delta\eta_t,$$

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i + \eta_t \qquad \Delta y_t = a_0 + \varepsilon_t + \Delta\eta_t,$$

$$y_t = y_0 + a_1 t + \varepsilon_t \qquad \Delta y_t = a_1 + \varepsilon_t - \varepsilon_{t-1}.$$

2) Detrending:

$$y_t = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n + \varepsilon_t,$$

regressing $\{y_t\}$ on the deterministic polynomial time trend

Difference or Detrend?

- **Difference or detrend?** Inappropriate method to eliminate trend will lead to a serious problem. (1) First-differencing the TS (trend-stationary) process will introduce a noninvertible unit root process into the MA component of the model. Examine the example
$$A(L)y_t = a_0 + a_1 t + B(L)\varepsilon_t.$$
Detrending yields a satationary and invertible ARMA process, but first-difference deduces that
$$A(L)\Delta y_t = a_1 + (1-L)B(L)\varepsilon_t$$
of which the second term $(1-L)B(L)\varepsilon_t$ makes $A(L)\Delta y_t$ noninvertible. (2) Detrending a DS process may not eliminate all the trend components (the stochastic component of the trend). Study the general trend plus irregular model
$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i + A(L)\eta_t.$$

Spurious Regression

High R-squared, significant t-test,
but no economic meaning.

Examine the model:

$$y_t = a_0 + a_1 z_t + e_t,$$

where $\{y_t\}$ and $\{z_t\}$ are two independent random walk processes, i.e.

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_{yt} \\ z_t &= z_{t-1} + \varepsilon_{zt} \end{aligned}$$

with two independent white-noise processes ε_{yt} and ε_{zt}.

Central Limit Theorem

CLT for a Martingale Difference Sequence: Let $\{Y_t\}$ be a scalar martingale difference sequence with $\bar{Y}_T = (1/T)\sum_{t=1}^{T} Y_t$. Suppose that (a) $E(Y_t^2) = \sigma_t^2 > 0$ with $(1/T)\sum_{t=1}^{T}\sigma_t^2 \to \sigma^2 > 0$; (b) $E|Y_t|^r < \infty$ for some r>2 and all t; (c) $(1/T)\sum_{t=1}^{T} Y_t^2 \to \sigma^2$ in probability. Then

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Y_t \to N(0,\sigma^2). \qquad \blacksquare$$

Theorem (consistent estimator of second moment) Suppose

$$Y_t = \sum_{j=1}^{\infty} \psi_j u_{t-j} \text{ with } \sum_{j=1}^{\infty} |\psi_j| < \infty$$

and $\{u_t\}$ is i.i.d. and for some $r > 2$, $E|u_t|^r < \infty$. Then as $T \to \infty$,

$$\frac{1}{T}\sum_{t=1}^{T} y_t y_{t-k} \to E\left(y_t y_{t-k}\right) \quad \text{in probability.} \qquad \blacksquare$$

An Application: Limit Theorem

AR(1) model $y_t = a_1 y_{t-1} + u_t$ with $|a_1| < 1$, where $\{u_t\}$ is i.i.d $(0, \sigma^2)$.

$$\hat{a}_1 = a_1 + \left(\sum_{t-1}^{T} y_{t-1}^2\right)^{-1} \sum_{t-1}^{T} u_t y_{t-1}.$$

$$\begin{aligned} E\left(u_t y_{t-1}\right)^2 &= EE[(u_t y_{t-1})^2 | y_{t-1}] = EE[u_t^2 | y_{t-1}]E\left(y_{t-1}\right)^2 \\ &= \sigma^2 \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_1^{i+j} E\left(u_{t-i}u_{t-j}\right) = \sigma^4 \sum_{i=1}^{\infty} a_1^{2i} = \frac{\sigma^4}{1-a_1^2} \end{aligned}$$

$$\frac{1}{T}\sum_{t=1}^{T} E\left(u_t y_{t-1}\right)^2 = \frac{\sigma^4}{1-a_1^2}.$$

$$\boxed{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t y_{t-1} \to N\left(0, \frac{\sigma^4}{1-a_1^2}\right).}$$

$$\boxed{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \to E y_{t-1}^2 = \sigma^2 \sum_{i=1}^{\infty} a_1^{2i} = \frac{\sigma^2}{1-a_1^2}.}$$

$$\begin{aligned} \sqrt{T}(\hat{a}_1 - a_1) &= \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t y_{t-1} \\ &\to \left(\frac{\sigma^2}{1-a_1^2}\right)^{-1} N\left(0, \frac{\sigma^4}{1-a_1^2}\right) = N(0, 1-a_1^2). \quad \text{as } T \to \infty, \end{aligned}$$

What is the limiting: AR(1) with a_1=1?

Now study

$$y_t = a_1 y_{t-1} + \varepsilon_t,\ a_1 = 1,$$

where $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$.The OLS estimator for a_1 satisfies

$$T(\hat{a}_1 - 1) = \frac{(1/T)\sum_{t=1}^{T} \varepsilon_t y_{t-1}}{(1/T^2)\sum_{t=1}^{T} y_{t-1}^2},$$

which is not asymptotically normally distributed. We will show that the asymptotic distribution of $T(\hat{a}_1 - 1)$ is related with Winner process.

Need Winner Process, etc

Winner Process 1/4

Winner Process (Standard Brownian Motion) $W(\cdot)$ is defined as a continuous-time stochastic process, associating each date $t \in [0,1]$ with the scalar $W(t)$ such that

(i) $W(0) = 0$;

(ii) For any dates $0 \leq t_1 < t_2 < \cdots < t_k = 1$, the changes

$$[W(t_2) - W(t_1)], \quad [W(t_3) - W(t_2)], \quad \cdots, \quad [W(t_k) - W(t_{k-1})]$$

are independent random variables;

(iii) For any $0 \leq t < s \leq 1$, $W(s) - W(t) \sim N(0, s-t)$.

Winner Process 2/4

Notes:

1) For any $t \in [0,1]$, select $dt > 0$ such that $0 < t - dt < 1$. $W(t) - W(t-dt) = \eta_t \sim N(0, dt)$ or

$$W(t) = W(t-dt) + \eta_t \quad \text{with } \eta_t \sim N(0, dt)$$

is a random walk with the time-step dt.

2) Other continuous-time processes can be generated from standard Brownian motion. For example, $B(t) \equiv \sigma W(t) \sim N(0, \sigma^2 t)$, which is called as Brownian motion with variance σ^2; $Z(t) = W^2(t) \sim t\chi^2(1)$.

3) $W(t)$ is continuous but not differentiable in $t \in [0,1]$, where the distance of $W(t_1)$ and $W(t_2)$ is defined as

$$d(t_1, t_2) = \sqrt{E\left(W(t_1) - W(t_2)\right)^2}$$

for any $t_1, t_2 \in [0,1]$ and $t_2 > t_1$. The reasons are that, for any $t_0 \in [0,1], dt > 0, t_0 + dt \in [0,1]$, $W(t_0 + dt) - W(t_0) \sim N(0, dt)$, and hence

$$\begin{aligned} d(t_0, t_0 + dt) &= \sqrt{E\left(W(t_0 + dt) - W(t_0)\right)^2} \\ &= \sqrt{Var\left(W(t_0 + dt) - W(t_0)\right)} \\ &= \sqrt{dt} \to 0, \text{ as } dt \to 0 \end{aligned}$$

$$\lim_{dt \to 0^+} \frac{d(t_0, t_0 + dt)}{dt} = \lim_{dt \to 0^+} \frac{\sqrt{dt}}{dt} = \infty.$$

Winner Process: Functional CLT 3/4

Define $X_T(r) \equiv \frac{1}{T}\sum_{t=1}^{[Tr]} \varepsilon_t.$

$$X_T(r) = \begin{cases} 0, & 0 \le r < 1/T \\ \varepsilon_1/T, & 1/T \le r < 2/T \\ (\varepsilon_1+\varepsilon_2)/T, & 2/T \le r < 3/T \\ \vdots & \\ (\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_{T-1})/T, & (T-1)/T \le r < 1 \\ (\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_T)/T, & r = 1 \end{cases}$$

$$\sqrt{T}X_T(r) \equiv \frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} \varepsilon_t = \frac{\sqrt{[Tr]}}{\sqrt{T}}\frac{1}{\sqrt{[Tr]}}\sum_{t=1}^{[Tr]} \varepsilon_t \longrightarrow N(0, r\sigma^2) \quad \text{as } T \longrightarrow \infty,$$

$$\sqrt{T}X_T(r)/\sigma \equiv \frac{1}{\sqrt{T}\sigma}\sum_{t=1}^{[Tr]} \varepsilon_t \longrightarrow N(0, r).$$

For any $1 > r_2 > r_1 > 0$, $\sqrt{T}\left(X_T(r_2) - X_T(r_1)\right)/\sigma \to N(0, r_2 - r_1)$,

independent of $\sqrt{T}X_T(r)/\sigma$, provided that $r < r_1$.

The limiting distribution of $\sqrt{T}X_T(r)$ is the same as the distribution of the Wiener process $B(r) = \sigma W(r)$

Winner Process 4/4

- **Continuous Mapping Theorem:** If $S_T(\cdot) \to S(\cdot)$ in distribution and $g(\cdot)$ is a continuous function, then $g(S_T(\cdot)) \to g(S(\cdot))$ in distribution.

For example, since $\sqrt{T}X_T(\cdot)/\sigma \to W(\cdot)$, we have $\sqrt{T}X_T(\cdot) \to \sigma W(\cdot)$ and $\sqrt{T}X_T(r) \sim N(0, \sigma^2 r)$; $S_T(\cdot) \equiv \left(\sqrt{T}X_T(\cdot)\right)^2 \to \sigma^2 W^2(\cdot)$; $\int_0^1 \sqrt{T}X_T(r)dr \to \sigma \int_0^1 W(r)dr$.

- **Stochastic Integral**: $\int_0^1 f(r)dW(r)$

$$\int_0^1 dW = W(1) \sim N(0,1)$$

$$\int_0^1 W(t)dW(t) = \frac{1}{2}\left[W^2(1) - 1\right] \sim \frac{1}{2}\left[\chi^2(1) - 1\right]$$

$$\int_0^1 rdW(r) \sim N(0, 1/3)$$

Theorem: Application to Random Walk 1/4

Theorem 1 Suppose that $y_t = y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. Then as $T \to \infty$,

$$\boxed{T^{-1/2} \sum_{t=1}^{T} \varepsilon_t \to \sigma W(1)}$$

$$T^{-1} \sum_{t=1}^{T} y_{t-1}\varepsilon_t \to \frac{1}{2}\sigma^2 (W^2(1) - 1)$$

$$T^{-3/2} \sum_{t=1}^{T} t\varepsilon_t \to \sigma W(1) - \sigma \int_0^1 W(r)dr$$

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} \to \sigma \int_0^1 W(r)dr$$

$$T^{-5/2} \sum_{t=1}^{T} t y_{t-1} \to \sigma \int_0^1 rW(r)dr$$

$$T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \to \sigma^2 \int_0^1 W^2(r)dr.$$

Theorem: Application to Random Walk 2/4

Theorem 1 Suppose that $y_t = y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. Then as $T \to \infty$,

$$\boxed{T^{-1}\sum_{t=1}^{T} y_{t-1}\varepsilon_t \to \frac{1}{2}\sigma^2(W^2(1) - 1)}$$

note that $y_t^2 = (y_{t-1} + \varepsilon_t)^2 = y_{t-1}^2 + \varepsilon_t^2 + 2\varepsilon_t y_{t-1}$. Therefore,

$$
\begin{aligned}
T^{-1}\sum_{t=1}^{T} \varepsilon_t y_{t-1} &= \frac{1}{2T}\sum_{t=1}^{T}\left(y_t^2 - y_{t-1}^2 - \varepsilon_t^2\right) \\
&= \frac{1}{2T}y_T^2 - \frac{1}{2T}\sum_{t=1}^{T}\varepsilon_t^2 = \frac{1}{2}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t\right)^2 - \frac{1}{2T}\sum_{t=1}^{T}\varepsilon_t^2 \\
&= \frac{1}{2}\left(\sqrt{T}X_T(1)\right)^2 - \frac{1}{2T}\sum_{t=1}^{T}\varepsilon_t^2 \\
&\to \frac{1}{2}(\sigma W(1))^2 - \frac{1}{2}\sigma^2 = \frac{1}{2}\sigma^2\left(W^2(1) - 1\right).
\end{aligned}
$$

Theorem: Application to Random Walk 3/4

Theorem 1 Suppose that $y_t = y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$. Then as $T \to \infty$,

$$\boxed{T^{-3/2} \sum_{t=1}^{T} t\varepsilon_t \to \sigma W(1) - \sigma \int_0^1 W(r)dr} \qquad \boxed{T^{-3/2} \sum_{t=1}^{T} y_{t-1} \to \sigma \int_0^1 W(r)dr}$$

since

$$X_T(r) = \begin{cases} 0, & 0 \le r < 1/T \\ y_1/T, & 1/T \le r < 2/T \\ y_2/T, & 2/T \le r < 3/T \\ \vdots & \\ y_{t-1}/T, & (T-1)/T \le r < 1 \\ y_T/T, & r = 1 \end{cases} \qquad T^{-3/2} \sum_{t=1}^{T} y_{t-1} = \int_0^1 \sqrt{T} X_T(r)dr \to \int_0^1 \sigma W(r)dr.$$

On the other hand, it follows from $T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-3/2}[\varepsilon_1 + (\varepsilon_1 + \varepsilon_2) + (\varepsilon_1 + \varepsilon_2 + \varepsilon_3) + \cdots + (\varepsilon_1 + \cdots + \varepsilon_{T-1})]$

$$= T^{-3/2} \sum_{t=1}^{T} (T-t)\varepsilon_t = T^{-1/2} \sum_{t=1}^{T} \varepsilon_t - T^{-3/2} \sum_{t=1}^{T} t\varepsilon_t \qquad \boxed{T^{-3/2} \textstyle\sum_{t=1}^{T} t\varepsilon_t = T^{-1/2} \sum_{t=1}^{T} \varepsilon_t \quad T^{-3/2} \sum_{t=1}^{T} y_{t-1}}$$

Theorem: Application to Random Walk 4/4

Theorem 1 Suppose that $y_t = y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$. Then as $T \to \infty$,

$$\boxed{T^{-5/2} \sum_{t=1}^{T} t y_{t-1} \to \sigma \int_0^1 rW(r)dr} \qquad \boxed{T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \to \sigma^2 \int_0^1 W^2(r)dr.}$$

$$\begin{aligned} t y_{t-1} &= t(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_{t-1}) \\ &= T \int_{(t-1)/T}^{t/T} ([Tr]+1)(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_{[Tr]})dr \\ &= T^2 \int_{(t-1)/T}^{t/T} ([Tr]+1) X_T(r)dr, \end{aligned}$$

$$T^{-5/2} \sum_{t=1}^{T} t y_{t-1} = T^{-1/2} \sum_{t=1}^{T} \int_{(t-1)/T}^{t/T} ([Tr]+1) X_T(r)dr$$

$$= \sum_{t=1}^{T} \int_{(t-1)/T}^{t/T} \frac{[Tr]+1}{T} \sqrt{T} X_T(r)dr = \int_0^1 \frac{[Tr]+1}{T} \sqrt{T} X_T(r)dr$$

$$\to \sigma \int_0^1 rW(r)dr \text{ as } T \to \infty.$$

$$T^{-2} \sum_{t=1}^{T} y_{t-1}^2 = T^{-2} \sum_{t=1}^{T} (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_{t-1})^2$$

$$= \sum_{t=1}^{T} \int_{(t-1)/T}^{t/T} \left(\sqrt{T} X_T(r)\right)^2 dr = \int_0^1 \left(\sqrt{T} X_T(r)\right)^2 dr$$

Application to Regression Models

Theorem 2 Consider the following three models:

$$\begin{aligned} \text{Model 1} &: \quad y_t = a_1 y_{t-1} + \varepsilon_t, \;\; a_1 = 1; \\ \text{Model 2} &: \quad y_t = a_0 + a_1 y_{t-1} + \varepsilon_t, \;\; a_0 = 0, \;\; a_1 = 1; \\ \text{Model 3} &: \quad y_t = a_0 + a_1 y_{t-1} + \beta t + \varepsilon_t, \;\; a_0 = 0, \;\; a_1 = 1, \;\; \beta = 0. \end{aligned}$$

(i) In Model 1, the data are generated by a random walk: $y_t = y_{t-1} + \varepsilon_t$, but the model is estimated by OLS: $y_t = a_1 y_{t-1} + \varepsilon_t$. The OLS estimator for a_1 satisfies that

$$T(\hat{a}_1 - 1) \to \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W^2(r)dr} = \frac{(W^2(1) - 1)/2}{\int_0^1 W^2(r)dr}$$

(ii) In Model 2, the data are generated by a random walk: $y_t = y_{t-1} + \varepsilon_t$, but the model is estimated by OLS: $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$. The OLS estimator for a_1 satisfies that

$$T(\hat{a}_1 - 1) \to \frac{\int_0^1 \bar{W}(r)dW(r)}{\int_0^1 \bar{W}^2(r)dr},$$

where $\bar{W}(r) = W(r) - \int_0^1 W(r)dr$ is the demeaned Brownian motion.

Application to Regression Models

Theorem 2 Consider the following three models:

$$\begin{aligned} \text{Model 1} &: \quad y_t = a_1 y_{t-1} + \varepsilon_t, \ \ a_1 = 1; \\ \text{Model 2} &: \quad y_t = a_0 + a_1 y_{t-1} + \varepsilon_t, \ \ a_0 = 0, \ \ a_1 = 1; \\ \text{Model 3} &: \quad y_t = a_0 + a_1 y_{t-1} + \beta t + \varepsilon_t, \ \ a_0 = 0, \ \ a_1 = 1, \ \ \beta = 0. \end{aligned}$$

(iii) In Model 3, the data are generated by a random walk: $y_t = y_{t-1} + \varepsilon_t$, but the model is estimated by OLS: $y_t = a_0 + a_1 y_{t-1} + \beta t + \varepsilon_t$. The OLS estimator for a_1 satisfies that

$$T(\hat{a}_1 - 1) \rightarrow \frac{\int_0^1 W^*(r)dW(r)}{\int_0^1 W^{*2}(r)dr},$$

where $W^*(r) = W(r) - 4\left(\int_0^1 W(r)dr - \frac{3}{2}\int_0^1 rW(r)dr\right) + 6r\left(\int_0^1 W(r)dr - 2\int_0^1 rW(r)dr\right)$ is the demeaned and detrended Brownian motion.

Limiting Distribution of t-Statistic

Model 1: Under the null: $a_1 = 1$, the OLS estimator $\hat{a}_1$ is a consistent estimator, and hence

$$\hat{\sigma}_T^2 \equiv \frac{1}{T-1} \sum_{t=1}^{T} \left(y_t - \hat{a}_1 y_{t-1}\right)^2 \to \sigma^2 \quad \text{in probability.}$$

Then

$$\begin{aligned} t_T \quad &\equiv \quad \frac{\hat{a}_1 - 1}{s.e.(\hat{a}_1)} = \frac{T^{-1} \sum_{t=1}^{T} y_{t-1} \varepsilon_t}{\hat{\sigma} \sqrt{T^{-2} \sum_{t=1}^{T} y_{t-1}^2}} \\ &\to \quad \frac{\frac{1}{2}\sigma^2 (W^2(1) - 1)}{\sigma \sqrt{\sigma^2 \int_0^1 W^2(r) dr}} = \frac{\int_0^1 W(r) dW(r)}{\sqrt{\int_0^1 W^2(r) dr}}. \end{aligned}$$

The t-statistic converge weakly to a functional of the Brownian motion with asymmetric limit distribution.

Limiting Distribution of t-Statistic

Model 2: Under the null: $a_1 = 1$,

$$t_T \equiv \frac{\hat{a}_1 - 1}{s.e.(\hat{a}_1)} \rightarrow \frac{\int_0^1 \bar{W}(r)dW(r)}{\sqrt{\int_0^1 \bar{W}^2(r)dr}}.$$

Model 3: Under the null: $a_1 - 1$,

$$t_T \equiv \frac{\hat{a}_1 - 1}{s.e.(\hat{a}_1)} \rightarrow \frac{\int_0^1 W^*(r)dW(r)}{\sqrt{\int_0^1 W^{*2}(r)dr}}.$$

OLS estimator for other random walks

- **Random Walk with a Drift: What is the Limit Distribution of the OLS Estimator for the autoregression?**

Consider the following random walk plus drift process:

$$y_t = a + \rho y_{t-1} + \varepsilon_t,$$

where $a \neq 0$, $\rho = 1$, and ε_t is i.i.d.

OLS estimator for other random walks

- **Random Walk with a Drift: What is the Limit Distribution of the OLS Estimator for the autoregression?**

  Consider the following random walk plus drift process:

$$y_t = a + \rho y_{t-1} + \varepsilon_t,$$

  where $a \neq 0$, $\rho = 1$, and ε_t is i.i.d.

- **Example:** Suppose that the true model is a unit root process with a constant

$$y_t = \alpha + y_{t-1} + u_t, \tag{10}$$

  where u_t are i.i.d.$(0, \sigma^2)$. However, we use a **Trend-Stationary Model**

$$y_t = c + \beta t + v_t \tag{11}$$

  for the estimation of β. What is the limit distribution of the OLS estimator $\hat{\beta}$? What happens to the conventional t-test for $\beta = 0$ in Model (11)?

OLS estimator for other random walks

- **Random Walk with a Drift: What is the Limit Distribution of the OLS Estimator for the autoregression?**

Consider the following random walk plus drift process:

$$y_t = a + \rho y_{t-1} + \varepsilon_t,$$

where $a \neq 0$, $\rho = 1$, and ε_t is i.i.d.

Unit Root Tests

1) **Dickey-Fuller (DF) Test:** The three basic models used for regression

$$\text{Model 1:} \quad y_t = \rho y_{t-1} + \varepsilon_t,$$
$$\text{Model 2:} \quad y_t = \alpha + \rho y_{t-1} + \varepsilon_t,$$
$$\text{Model 3:} \quad y_t = \alpha + \delta t + \rho y_{t-1} + \varepsilon_t.$$

can be equivalently written, respectively, as

$$\text{Model 1:} \quad \Delta y_t = \gamma y_{t-1} + \varepsilon_t, \tag{12}$$
$$\text{Model 2:} \quad \Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t, \tag{13}$$
$$\text{Model 3:} \quad \Delta y_t = \alpha + \delta t + \gamma y_{t-1} + \varepsilon_t \tag{14}$$

Model used for regression	Null Hypothesis H_0	Test Statistic
Model 3: $\Delta y_t = \alpha + \delta t + \gamma y_{t-1} + \varepsilon_t$	$\gamma = 0$	τ_τ
	$\gamma = \delta = 0$	ϕ_3
	$\alpha = \gamma = \delta = 0$	ϕ_2
Model 2: $\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$	$\gamma = 0$	τ_μ
	$\alpha = \gamma = 0$	ϕ_1
Model 1: $y_t = \rho y_{t-1} + \varepsilon_t,$	$\gamma = 0$	τ

Unit Root Tests

Note: Problems:

(i) For y_t, AR(p)?

(ii) For the error term, MA(q)?

(iii) When to include the constant or time trend in the regression model?

(iv) More than one unit roots?

(v) Structural change?

2) **Augmented Dickey-Fuller (ADF) Test (Said and Dickey test):** Extend the DF test to ARMA(p,q) models

$$y_t = a_1 y_{t-1} + \cdots + a_p y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

or

$$(1 - A(L))y_t \equiv (1 - a_1 L - \cdots - a_p L^p) y_t = (1 + b_1 L + \cdots + b_q L^q)\varepsilon_t \equiv B(L)\varepsilon_t,$$

- **A Four-step Test Approach: when the form of the DGP is completely unknown.** See Enders's book: P213 **Figure 4.13**. Note that 1) Plotting the data is usually an important indicator of presence of deterministic regressors; 2) Theoretical consideration might suggest the appropriate deterministic regressors. See GDP and Unit Roots example on P214.