Advanced Econometrics

Chapter 17: Time Series Analysis

In Choi

Sogang University

References

- Brockwell, P.J., and R.A. Davis. (1991). Time Series: Theory and Methods. New York: Springer-Verlag.
- Canova, F. (2007). Methods for Applied Macroeconomic Research. New Jersey: Princeton University Press.
- Fuller, W.A. (1976). Introduction to Statistical Time Series. New York: John Wiley & Sons.
- Hamilton, J.D. (1994). Time Series Analysis. New York: Princeton University Press.

Stochastic processes

- A stochastic process is a family of random variables $\{X_t, t \in \mathbb{T}\}$, where $\mathbb{T}$ is an index set.

Example

Random walk

$$X_t = \sum_{i=1}^{t} e_i,$$

where $e_i \sim iid\left(0, \sigma^2\right)$. Note that X_t has a different distribution at each point t.

Weak stationarity and strict stationarity

- If $\{X_t, t \in \mathbb{T}\}$ is a stochastic process such that $Var(X_t) < \infty$ for each $t \in \mathbb{T}$, the autocovariance function $\gamma_X(\cdot, \cdot)$ of $\{X_t\}$ is defined by
$$\gamma_X(r, s) = Cov(X_r, X_s) = E(X_r - EX_r)(X_s - EX_s).$$
Because $Var(X_t) < \infty$ for each $t \in \mathbb{T}$,
$$\begin{aligned} \gamma_X(r, s) &\leq \left[E(X_r - EX_r)^2\right]^{1/2} \left[E(X_s - EX_s)^2\right]^{1/2} \\ &< \infty \end{aligned}$$
by the Cauchy–Schwarz inequality.

Weak stationarity and strict stationarity

- The autocorrelation function $\rho_X(r,s)$ of $\{X_t, t \in \mathbb{T}\}$ is defined by

$$\rho_X(r,s) = \frac{\gamma_X(r,s)}{\sqrt{\gamma_X(r,r)\,\gamma_X(s,s)}}$$

Example

Let $X_t = e_t + \theta e_{t-1},\ e_t \sim iid\left(0, \sigma^2\right)$.

$$\gamma_X(t+h,t) = Cov(X_{t+h}, X_t) \begin{cases} = \left(1+\theta^2\right)\sigma^2, & h = 0 \\ = \theta\sigma^2, & h = \pm 1 \\ = 0, & |h| > 1 \end{cases}$$

$$\rho_X(t+h,t) \begin{cases} = 1, & h = 0 \\ = \frac{\theta}{\left(1+\theta^2\right)}, & h = \pm 1 \\ = 0, & |h| > 1 \end{cases}$$

Weak stationarity and strict stationarity

- A time series $\{X_t, t \in \mathbb{Z}\}$ with index set $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ is said to be (weakly) weakly stationary, if

1. $E\left|X_t^2\right| < \infty$ for all $t \in \mathbb{Z}$
2. $EX_t = m$ for all $t \in \mathbb{Z}$
3. $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ for all $r, s, t \in \mathbb{Z}$

Weak stationarity and strict stationarity

- We may define the autocovariance function of a weakly stationary process as a function of just one variable, which is the difference of two time index.
- Instead of $\gamma_X(r, s)$, write $\gamma_X(r - s)$.

Weak stationarity and strict stationarity

- Let $\gamma_X(h)$ be an autocovariance function of a weakly stationary process $\{X_t, t \in \mathbb{Z}\}$.

1. $\gamma_X(0) \geq 0$
2. $|\gamma_X(h)| \leq \gamma_X(0)$ for all $h = 0, \pm 1, \pm 2, \ldots$

$$\begin{aligned} |\gamma_X(h)| = |Cov(X_{t+h}, X_t)| &\leq (VarX_{t+h})^{1/2} (VarX_t)^{1/2} \\ &= \gamma_X(0) \end{aligned}$$

3. $\gamma_X(-h) = \gamma_X(h)$ for all $h = 0, \pm 1, \pm 2, \ldots$

$$\gamma_X(-h) = EX_{t-h}X_t = EX_tX_{t+h} = \gamma_X(h).$$

Example

Let $X_t = e_t + \theta e_{t-1}$, $e_t \sim iid\left(0, \sigma^2\right)$. Then, $\{X_t\}$ is weakly stationary.

Example

Let $X_t = X_{t-1} + e_t$, $e_t \sim iid\left(0, \sigma^2\right)$ and $X_0 = 0$. Then,

$$X_t = \sum_{i=1}^{t} e_i.$$

Since $Var\left(X_t\right) = t\sigma^2$, $\left\{X_t\right\}$ is not weakly stationary, .

Example

If $X_t \sim N\left(0, \sigma_t^2\right)$, $\left\{X_t\right\}$ is not weakly stationary.

Weak stationarity and strict stationarity

- The sample autocovariance function of the observations $\{x_1, \cdots, x_T\}$ for a weakly stationary time series $\{X_t\}$ is defined by

$$\hat{\gamma}_X(h) = T^{-1} \sum_{j=1}^{T-h} (x_{j+h} - \bar{x})(x_j - \bar{x}), \ 0 \leq h \leq T,$$

  and $\hat{\gamma}_X(h) = \hat{\gamma}_X(-h)$, where $\bar{x} = T^{-1} \sum_{j=1}^{T} x_j$.
- The sample autocorrelation function of the observations $\{x_1, \cdots, x_T\}$ for a weakly stationary time series $\{X_t\}$ is defined by

$$\hat{\rho}_X(h) \equiv \hat{\gamma}_X(h) / \hat{\gamma}_X(0).$$

- $\hat{\rho}_X(1) > 0$ reflects a tendency for successive observations to lie on the same side of the mean, while $\hat{\rho}_X(1) < 0$ reflects a tendency for successive observations to lie on the opposite sides of the mean.

Weak stationarity and strict stationarity

- The time series $\{X_t, t \in \mathbb{Z}\}$ is said to be strictly stationary if the joint distribution of $(X_{t_1}, \cdots, X_{t_k})'$ and $(X_{t_1+h}, \cdots, X_{t_k+h})'$ are the same for all positive integers k and for all $t_1, \cdots, t_k, h \in \mathbb{Z}$.
- Strict stationarity with finite second moments $\Rightarrow$ stationarity. The converse not true in general.
- Gaussian (multivariate normal) weak stationarity $\Rightarrow$ strict stationarity.

Autoregressive processes

- The autoregressive process of order p ($AR(p)$ process) is written as

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + e_t,$$

where $e_t \sim iid\ (0, \sigma^2)$. Or we may write using the lag operator $(B^p y_t = y_{t-p})$

$$(1 - \alpha_1 B - \cdots - \alpha_p B^p)\, y_t = e_t.$$

- Let $y_t = \alpha y_{t-1} + e_t$, $|\alpha| < 1$ and $\hat{\alpha} = \left(\sum_{t=2}^{T} y_{t-1} y_t \right) / \sum_{t=2}^{T} y_{t-1}^2$.
- We have

1. $\hat{\alpha} \xrightarrow{p} \alpha$
2. $\sqrt{T} \left(\hat{\alpha} - \alpha \right) \xrightarrow{d} N \left(0, 1 - \alpha^2 \right)$.

Sketch of proof.

1. Write $\hat{\alpha} - \alpha = \sum_{t=2}^{T} y_{t-1} e_t / \sum_{t=2}^{T} y_{t-1}^2$.Since $y_{t-1} = \sum_{i=0}^{\infty} \alpha^i e_{t-1-i},$, we can obtain using Chebychev's inequality

$$\sum_{t=2}^{T} y_{t-1} e_t / T \xrightarrow{p} 0.$$

Furthermore, $\sum_{t=2}^{T} y_{t-1}^2 / T \xrightarrow{p} \frac{\sigma^2}{1-\alpha^2}$ $(= E(y_{t-1}^2))$.

2. By the central limit theorem,

$$\frac{1}{\sqrt{T}}\sum_{t=2}^{T} y_{t-1}e_t \xrightarrow{d} N\left(0, \sigma^2 p\lim \frac{\sum_{t=2}^{T} y_{t-1}^2}{T}\right) = N\left(0, \frac{\sigma^4}{1-\alpha^2}\right)$$

Hence $\sqrt{T}\left(\hat{\alpha}-\alpha\right) \xrightarrow{d} N\left(0, 1-\alpha^2\right)$.

A continuous-time stochastic process, $\{W(r), 0 \leq r \leq 1\}$, is called Brownian motion or a Wiener process if it satisfies the following conditions.

1. $W(0) = 0$ almost surely.
2. For $0 \leq t_0 \leq t_1 \leq \ldots \leq t_k \leq 1$, $W(t_1) - W(t_0), \ldots, W(t_k) - W(t_{k-1})$ are independent.
3. $W(t) - W(s)$ $(t > s)$ follows $\mathbf{N}(0, t - s)$.

- When $\alpha = 1$, we have

1. $\hat{\alpha} \xrightarrow{p} 1$.
2. $T(\hat{\alpha} - 1) \xrightarrow{d} \left(\int_0^1 W^2(r)\, \mathrm{d}r \right)^{-1} \int_0^1 W(r)\, \mathrm{d}W(r)$.

Note that $\int_0^1 W(r)\, \mathrm{d}W(r) = \frac{1}{2}(W^2(1) - 1) = \frac{1}{2}(\chi^2(1) - 1)$.

- Itô's rule,

$$g(W(t)) - g(0) = \int_0^t g'(W(r))\,\mathrm{d}W(r) + \frac{1}{2}\int_0^t g''(W(r))\,\mathrm{d}r,$$

where $g : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable.

- For $g(W(r)) = W^2(r)$ and $t = 1$, this formula gives

$$\int_0^1 W(r)\,\mathrm{d}W(r) = \frac{1}{2}(W^2(1) - 1) = \frac{1}{2}(\chi^2(1) - 1).$$

- The stochastic integral $\int_0^t G(r)\,\mathrm{d}W(r)$ is equal in distribution to $\mathbf{N}\left(0, \int_0^t G^2(r)\,\mathrm{d}r\right)$, where $G : \mathbb{R} \to \mathbb{R}$ has the property $\int_0^t G^2(r)\,\mathrm{d}r < \infty$.

For the asymptotic distribution, use the weak convergence results

1. $\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r)\,\mathrm{d}r$;
2. $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t \Rightarrow \sigma^2 \int_0^1 W(r)\,\mathrm{d}W(r)$.

- See Phillips (1987, Econometrica) and Chan and Wei (1988, Annals of Statistics) for proofs.

- Let $y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + e_t$, $t = p+1, \cdots, T$.
- In matrix notation, $y = X\alpha + e$, where

$$y = \begin{bmatrix} y_{p+1} \\ \vdots \\ y_T \end{bmatrix}, \ X = \begin{bmatrix} y_p & \cdots & y_1 \\ \vdots & & \\ y_{T-1} & \cdots & y_{T-p} \end{bmatrix}, \ \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix},$$

$$\text{and } e = \begin{bmatrix} e_{p+1} \\ \vdots \\ e_T \end{bmatrix}.$$

- Let $\hat{\alpha} = (X'X)^{-1} X'y$ and $\hat{\sigma}^2 = \frac{(y-X\hat{\alpha})'(y-X\hat{\alpha})}{T-p}$.

- If the process $\{y_t\}$ is weakly stationary,

1. $\hat{\alpha} \xrightarrow{p} \alpha$.
2. $\sqrt{T}\left(\hat{\alpha} - \alpha\right) \xrightarrow{d} N\left(0, \Sigma\right)$, where

$$\Sigma = \sigma^2 \begin{bmatrix} \gamma_0 & & & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & & \gamma_{p-2} \\ & & \ddots & \\ \gamma_{p-1} & & & \gamma_0 \end{bmatrix}^{-1} \text{ and } \gamma_h = E\left(y_{t+h} y_t\right).$$

- If all the roots of the characteristic equation

$$1 - \alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_p z^p = 0$$

lie outside the unit circle, $\{y_t\}$ is weakly stationary. See Fuller (1976) for a proof.

Example

Consider the $AR(2)$ process

$$y_t - y_{t-1} + 0.16 y_{t-2} = e_t.$$

The characteristic equation for this is $1 - z + 0.16z^2 = (1 - 0.8z)(1 - 0.2z)$, which gives $z = \frac{1}{0.8}, \frac{1}{0.2}$.

Example

(continued) Hence, y_t is weakly stationary. We may also express

$$(1-0.8B)(1-0.2B)\,y_t = e_t,$$

which gives $(1-0.8B)\,y_t = \sum_{i=0}^{\infty}(0.2)^i e_{t-i} = u_t$
and $y_t = \sum_{i=0}^{\infty}(0.8)^i u_{t-i}$. (The impact of an event that happened long ago is negligible)

Example

Consider the $AR\left(2\right)$ process $y_t - 1.2y_{t-1} + 0.2y_{t-2} = e_t$. The characteristic equation

$$1 - 1.2z + 0.2z^2 = \left(1 - z\right)\left(1 - 0.2z\right) = 0$$

gives $z = 1,\ \frac{1}{0.2}$.

Example

(continued) Hence, y_t is not weakly stationary. As a matter of fact,

$$\begin{aligned} y_t - y_{t-1} &= (1-0.2B)^{-1} e_t \\ &= \left(1 + 0.2B + 0.2B^2 + \cdots\right) e_t \\ &= \sum_{i=0}^{\infty} (0.2)^i e_{t-i} \\ &= u_t, \text{ say.} \end{aligned}$$

and $y_t = \sum_{i=1}^{t} u_i + y_0$.

- Suppose that the AR model has at least one stable root. Choi (1993, Econometric Theory) shows

$$\sqrt{T}\begin{pmatrix} \hat{\alpha}_1 - \alpha_1 \\ \vdots \\ \hat{\alpha}_p - \alpha_p \end{pmatrix} \Rightarrow \mathbf{N}(0, \sigma_u^2 A \Lambda^{-1} A'),\ \Lambda = \begin{bmatrix} \gamma_0 & & & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & & \gamma_{p-2} \\ & & \ddots & \\ \gamma_{p-1} & & & \gamma_0 \end{bmatrix}$$

Here $A = \begin{bmatrix} 1 & & \underset{\sim}{0} \\ -1 & \ddots & \\ \vdots & \ddots & 1 \\ \underset{\sim}{0} & & -1 \end{bmatrix}$. The covariance $A\Lambda^{-1}A'$, however, does not have full rank since $rank(A\Lambda^{-1}A') = p - 1$.

Moving average processes

- The moving average process of order q (MA(q)) is written as

$$y_t = \sum_{j=0}^{q} \theta_j e_{t-j},$$

where $e_t \sim iid\ 0, \sigma^2)$.

Example

$$y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} \sim MA\,(2)$$

- Consider an $MA(q)$ process

$$y_t = e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} = (1 + \theta_1 B + \cdots + \theta_q B^q)\, e_t.$$

If all the roots of the equation $1 + \theta_1 z + \cdots + \theta_q z^q = 0$ lie outside the unit circle, the MA process can be written as an AR(∞) process, such that

$$e_t = \sum_{j=0}^{\infty} \psi_j y_{t-j} \text{ with } \left|\psi_j\right| < \infty.$$

When $\{e_t\}$ can be expressed in such a way, we call the *MA* process invertible.

Example

Let $y_t = e_t - e_{t-1} = (1 - B)\, e_t$. The root of the equation $1 - z = 0$ is $z = 1$. Hence, this MA process is not invertible. In fact,

$$e_t = \frac{1}{1-B} y_t = \left(1 + B + B^2 + \cdots\right) y_t = \sum_{j=0}^{\infty} \psi_j y_{t-j},$$

where

$$\psi_j = 1 \text{ and } \sum_{j=0}^{\infty} \left|\psi_j\right| = \infty.$$

Example

Let $y_t = e_t - 0.9e_{t-1} = (1 - 0.9B)\, e_t$. Then,

$$1 - 0.9z = 0; \;\; z = \frac{1}{0.9} > 1$$

and

$$e_t = \frac{1}{1 - 0.9B} y_t = \left(1 + 0.9B + 0.9^2 B^2 + \cdots\right) y_t = \sum_{j=0}^{\infty} \psi_j y_{t-j},$$

where

$$\psi_j = 0.9^j \text{ and } \sum_{j=0}^{\infty} \left|\psi_j\right| = \frac{1}{1 - 0.9} < \infty.$$

- Let $y_t = e_t + \theta e_{t-1}$, $|\theta| < 1$.
- The nonlinear least squares estimator of θ has the following property (See Fuller (1976) for proofs).

1. $\hat{\theta} \xrightarrow{p} \theta$
2. $\sqrt{T}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(0, 1 - \theta^2\right)$.

Autoregressive moving average processes

- ARMA(p, q) model

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + e_t + \theta e_{t-1} + \cdots + \theta_q e_{t-q}, \; e_t \sim iid\left(0, \sigma^2\right)$$

- If all roots of the equations

$$a\left(z\right) = 1 - \alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_p z^p = 0$$

and

$$b\left(z\right) = 1 + \theta_1 z + \cdots \theta_q z^q = 0$$

lie outside the unit circle, $\left\{y_t\right\}$ is weakly stationary and invertible.

- If $\{y_t\}$ is weakly stationary and invertible, we have

$$\sqrt{T}\left(\eta_{NLLS} - \eta\right) \xrightarrow{d} N\left(0, V\right).$$

$$\eta' = \left[\begin{array}{cccccc} \alpha_1 & \cdots & \alpha_p & \theta_1 & \cdots & \theta_q \end{array}\right], \; V = \sigma^2 \left[\begin{array}{cc} EU_1 U_1' & EU_1 V_1' \\ EV_1 U_1' & EV_1 V_1' \end{array}\right]^{-1}$$

$$\begin{aligned} a(B)\, u_t &= e_t \,, \; b(B)\, v_t = e_t, \\ U_t &= \left[u_t, ..., u_{t-p+1}\right]' \text{ and } V_t = \left[v_t, ..., v_{t-q+1}\right]'. \end{aligned}$$

Example

$$y_t = \alpha_1 y_{t-1} + e_t + \theta e_{t-1}, e_t \sim iid\left(0, \sigma^2\right)$$

$$\sqrt{T}\begin{pmatrix} \hat{\alpha}_1 - \alpha_1 \\ \hat{\theta}_1 - \theta_1 \end{pmatrix} \xrightarrow{d} N\left(0, V\right),$$

where

$$V = \begin{bmatrix} \left(1-\alpha_1^2\right)^{-1} & \left(1+\alpha_1\theta_1\right)^{-1} \\ \left(1+\alpha_1\theta_1\right)^{-1} & \left(1-\theta_1^2\right)^{-1} \end{bmatrix}^{-1}$$

Example

(continued)

$$U_t - \alpha_1 U_{t-1} = e_t, \;\; V_t - \theta_1 V_{t-1} = e_t$$

$$\Rightarrow \quad EU_t^2 = \frac{\sigma^2}{1-\alpha_1^2}$$

$$EU_t V_t = \frac{\sigma^2}{1+\alpha_1\theta_1}$$

$$EV_t^2 = \frac{\sigma^2}{1-\theta_1^2}$$

$$\Rightarrow \quad V = \sigma^2 \begin{bmatrix} \frac{\sigma^2}{1-\alpha_1^2} & \frac{\sigma^2}{1+\alpha_1\theta_1} \\ \frac{\sigma^2}{1+\alpha_1\theta_1} & \frac{\sigma^2}{1-\theta_1^2} \end{bmatrix} = \begin{bmatrix} \left(1-\alpha_1^2\right)^{-1} & \left(1+\alpha_1\theta_1\right)^{-1} \\ \left(1+\alpha_1\theta_1\right)^{-1} & \left(1-\theta_1^2\right)^{-1} \end{bmatrix}$$

- The process $\{y_t\}$ is said to be an $ARIMA\,(p, d, q)$ process if $\{(1-B)^d\, y_t\}$ is a weakly stationary $ARMA\,(p, q)$ process.

Example

Let

$$(1-\alpha B)\,(1-B)\,y_t = e_t,\ \ |\alpha| < 1.$$

Then, $\{y_t\}$ is an $ARIMA\,(1, 1, 0)$ process, i.e., differencing $\{y_t\}$ once yields $ARMA\,(1, 0)$ process.

- The slowly decaying positive sample autocorrelation suggests the appropriateness of an ARIMA model.

- We need to difference an ARIMA(p, d, q) process d times in order to obtain a weakly stationary process.
- How do we know that $d = 1$? Perform unit root test.
- See Choi (2015) "Almost all about unit roots: foundations, developments and applications," Cambridge University Press for the literature on unit roots.

- Consider the AR(1) model $y_t = \alpha y_{t-1} + e_t$, $e_t \sim i.i.d.(0, \sigma^2)$.
- Let $\hat{\alpha} = \sum_{t=2}^{T} y_t y_{t-1} / \sum_{t=2}^{T} y_{t-1}^2$.
- When $|\alpha| < 1$, as $T \to \infty$,

$$\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{d} N(0, 1 - \alpha^2).$$

and

$$t(\alpha) = \frac{\hat{\alpha} - \alpha}{\sqrt{\hat{\sigma}^2 (\sum y_{t-1}^2)^{-1}}} \xrightarrow{d} N(0, 1),$$

where $\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=2}^{T} (y_t - \hat{\alpha} y_{t-1})^2$.

- However, when $\alpha = 1$,

$$T(\hat{\alpha} - 1) \xrightarrow{d} \left(\int_0^1 W^2(r)\,\mathrm{d}r \right)^{-1} \int_0^1 W(r)\,\mathrm{d}W(r)$$

and

$$t(1) = \frac{\hat{\alpha} - 1}{\sqrt{\hat{\sigma}^2 \left(\sum y_{t-1}^2\right)^{-1}}} \xrightarrow{d} \left(\int_0^1 W^2(r)\,\mathrm{d}r \right)^{-1/2} \int_0^1 W(r)\,\mathrm{d}W(r).$$

- These are known as the Dickey-Fuller unit root test statistics. (See Fuller (1976) for their critical values).

Unit root tests

- The test statistics test the null hypothesis $H_0 : \alpha = 1$ against the alternative $H_1 : |\alpha| < 1$.
- When $T(\hat{\alpha} - 1)$ and $t(1)$ are less than their corresponding critical values, the null is rejected.
- Under the alternative,

$$T(\hat{\alpha} - 1) = T(\hat{\alpha} - \alpha) + T(\alpha - 1) = O_p(\sqrt{T}) + O(T) \xrightarrow{p} -\infty.$$

- Likewise, $t(1) \xrightarrow{p} -\infty$ under the alternative.

Unit root tests

- Alternatively, we may write the model as

$$\Delta y_t = \lambda y_{t-1} + e_t,\ \ e_t \sim i.i.d.(0, \sigma^2)$$

  and test the null hypothesis $H_0 : \lambda = 0$.
- The test statistics are

$$T\hat{\lambda} \text{ and } \frac{\widehat{\lambda}}{\sqrt{\hat{\sigma}^2 \left(\sum y_{t-1}^2\right)^{-1}}}.$$

Unit root tests

- When

$$y_t - \mu = \alpha(y_t - \mu) + e_t,$$

  or

$$y_t = \mu(1-\alpha) + \alpha y_{t-1} + e_t,$$

  $\hat{\alpha}$ also has a nonnormal distribution in the limit if $\alpha = 1$.

- The unit root test statistics for this model are ($\hat{\alpha}$ is the OLS estimator of α):

$$T(\hat{\alpha} - 1)$$

  and

$$\frac{\hat{\alpha} - 1}{\sqrt{\hat{\sigma}^2 \left(\sum_{t=2}^{T} (y_{t-1} - \bar{y}_-)^2\right)^{-1}}}.$$

- Their critical values can be found in Fuller (1976).

Unit root tests

- The AR(p) model

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha y_{t-p} + e_t, \;\; e_t \sim i.i.d.(0, \sigma^2)$$

can be written as

$$\Delta y_t = \lambda y_{t-1} + \sum_{j=1}^{p-1} w_j \Delta y_{t-j} + e_t, \;\; e_t \sim i.i.d.(0, \sigma^2)$$

where $\lambda = \alpha_1 + \dots + \alpha_p$ and $w_j = -\sum_{k=j}^{p} \alpha_k$.

- When there is a unit root, $\alpha_1 + \dots + \alpha_p = 1$. Thus, the null of a unit root can be tested by testing $\lambda = 0$. The t-test for this null hypothesis is called the augmented Dickey-Fuller test. It has the same asymptotic distribution as the t-test for the AR(1) model.

Unit root tests

- If y_0 is a stochastically bounded random variable.

$$T\hat{\lambda} \Rightarrow \frac{\sigma_u}{\sigma_{l\Delta y}} \left(\int_0^1 W(r)^2 \, \mathrm{d}r \right)^{-1} \int_0^1 W(r) \, \mathrm{d}W(r), \qquad (1)$$

where $\sigma^2_{l\Delta y} = \sum_{h=-\infty}^{\infty} Cov(\Delta y_t, \Delta y_{t+h})$.

- The t-ratio using $\hat{\lambda}$ is the augmented Dickey-Fuller test statistic. Its limiting distribution is

$$\left(\int_0^1 W^2(r) \, \mathrm{d}r \right)^{-1/2} \int_0^1 W(r) \, \mathrm{d}W(r).$$

Detrending time series

- We decompose the process $\{y_t\}$ as

$$y_t = m_t + X_t,$$

where m_t is a "trend" component and X_t is a "random noise" component. We need to estimate m_t in order to eliminate a trend.

1. Least squares estimation of m_t.
   Let
   $$m_t = a_0 + a_1 t + \cdots + a_p t^p, \ (t = 1, 2, \cdots, T).$$
   The estimates of $a_0, \cdots, a_p$ are obtained by OLS.
   $$\begin{aligned} y_t &= \hat{a}_0 + \cdots + \hat{a}_p t^p + \hat{X}_t \\ &= \hat{m}_t + \hat{X}_t. \end{aligned}$$
   When $\{X_t\}$ is weakly stationary,
   $$\hat{a}_i \xrightarrow{p} a_i \text{ and } t_{\hat{a}_i} \xrightarrow{d} N(0,1).$$

Detrending time series

2. Differencing

$$\Delta^k y_t = (1 - B)^k y_t.$$

Example

Let

$$y_t = at + X_t,$$

where $\{X_t\}$ is weakly stationary. Then,

$$\Delta y_t = a + \Delta X_t.$$

Hence, $\{\Delta y_t\}$ is weakly stationary.

Detrending time series

3. Double differencing

$$y_t = at + bt^2 + X_t$$
$$\Delta y_t = a + b\left(2t - 1\right) + \Delta X_t$$
$$\Delta^2 y_t = 2b + \Delta^2 X_t$$

Deseasonalizing by differencing

- Suppose

$$Y_t = S_t + X_t,$$

  where S_t denotes a "seasonal" component.
- For example, for quarterly data,

$$S_t = S_{t+4}.$$

Deseasonalizing by differencing

- We can deseasonalize quarterly data by applying $\left(1 - B^4\right)$. That is

$$\begin{aligned} \left(1 - B^4\right) y_t &= y_t - y_{t-4} \\ &= S_t + X_t - S_{t-4} - X_{t-4} \\ &= X_t - X_{t-4}. \end{aligned}$$

- Monthly data is deseasonalized by applying $\left(1 - B^{12}\right)$.