

# Econometrics

## Chapter 10: Basic Concepts and Models in Time Series Analysis

In Choi

Sogang University

## Definition

A stochastic process is a family of random variables  $\{X_t(\omega), t = 1, 2, \dots\}$ .

- For each fixed  $t$ ,  $X_t(\omega)$  is a random variable.
- For fixed  $\omega$ ,  $X_t(\omega)$  is a real valued function of  $t$  (realizations or sample-paths of the process  $\{X_t, t = 1, 2, \dots\}$ )
- In time series analyses, we model data as realizations of stochastic processes.
- In basic statistics, we model data as realizations of a random variable.

## Example

$$X_i \sim iid N(\mu, \sigma^2).$$

## Example

Random walk

$$S_t = \sum_{i=1}^t X_i, \quad t \geq 1$$

$X_i$ 's are iid random variables.

## Example

Branching processes

$$X_{t+1} = \sum_{j=1}^{X_t} Z_{t,j}$$

$Z_{t,j}$  represents the number of offspring of the  $j^{\text{th}}$  individual born in generation  $t$ .

## Definition

If  $\{X_t, t = 1, 2, \dots\}$  is a stochastic process such that  $\text{Var}(X_t) < \infty$  for each  $t$ , the autocovariance function  $\gamma_X(\cdot, \cdot)$  of  $\{X_t\}$  is defined by

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)], \quad r, s = 1, 2, \dots$$

- Autocovariance function shows how  $X_t$  is related to its future and past.
- Most economic time series are correlated over time.

## Example

Let  $X_t = Z_t + Z_{t-1}$ ,  $Z_t \sim iid (0, \sigma^2)$ .

$$\begin{aligned}\gamma(r, s) &= E[X_r - EX_r][X_s - EX_s] \\ &= E(Z_r + Z_{r-1})(Z_s + Z_{s-1}) \\ &= \begin{cases} 2\sigma^2, & r = s \\ \sigma^2, & |r - s| = 1 \\ 0, & \text{o.w.} \end{cases}\end{aligned}$$

## Definition

The time series  $\{X_t, t = 1, 2, \dots\}$  is said to be weakly stationary if

- 1  $E |X_t^2| < \infty$  for all  $t$ .
  - 2  $EX_t = m$  for all  $t$ .
  - 3  $\gamma_X(r, s) = \gamma_X(r + t, s + t)$  for all  $r, s, t$ .
- Weak stationarity = covariance stationarity, stationarity in the wide sense, second-order stationarity.
  - Weak stationarity is the most basic assumption in time series analysis. Without it, it is hard to proceed further.

# Weak stationarity and strict stationarity

- For weakly stationary processes,

$$\gamma_X(r, s) = \gamma_X(r - r, s - r) = \gamma_X(0, s - r) = \gamma_X(h), \quad h = s - r.$$

That is, the covariance function depends only on the time time difference.

- Autocovariance function

$$\rho_X(h) = \gamma_X(h) / \gamma_X(0) \quad \text{for all } t, h = 1, 2, \dots$$

## Example

$$X_t = Z_t + Z_{t-1}, \quad Z_t \sim iid(0, \sigma^2)$$

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

(does not depend on  $t$ )

$\sim$  weakly stationary

$$\rho_X(h) = \begin{cases} 1, & h = 0 \\ 1/2, & h = \pm 1 \\ 0, & |h| > 1 \end{cases}$$



# Weak stationarity and strict stationarity

## Example

$$S_t = \sum_{i=1}^t X_i, \quad X_i \sim iid(0, \sigma^2)$$

$$\begin{aligned} \text{Cov}(S_{t+h}, S_t) &= E \left( \sum_{i=1}^{t+h} X_i \right) \left( \sum_{i=1}^t X_i \right) \\ &= \sigma^2 t. \\ &\sim \text{nonstationary} \end{aligned}$$

## Definition

The time series  $\{X_t, t = 1, 2, \dots\}$  is said to be strictly stationary if the joint distribution of  $(X_{t_1}, \dots, X_{t_k})'$  and  $(X_{t_1+h}, \dots, X_{t_k+h})'$  are the same for all positive integers  $k$  and for all  $t_1, \dots, t_k, h$ .

- strict stationarity with finite second moments  $\Rightarrow$  stationarity.
- Gaussian (multivariate normal) weak stationarity  $\Rightarrow$  strict stationarity.

# Properties of autocovariance function

- Let  $\gamma(h)$  be an autocovariance function of a weakly stationary process  $X_t$ .

①  $\gamma(0) \geq 0$

②  $|\gamma(h)| \leq \gamma(0)$  for all  $h = 0, \pm 1, \pm 2, \dots$

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

③  $\gamma(-h) = \gamma(h)$  for all  $h = 0, \pm 1, \pm 2, \dots$

$$\gamma(-h) = EX_t X_{t-h} = EX_{t-h} X_t = EX_t X_{t+h} = \gamma(h).$$

## Definition

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

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(X_t - \bar{X}), \quad 0 \leq h \leq n,$$

and  $\hat{\gamma}(h) = \hat{\gamma}(-h)$ , where  $\bar{X} = n^{-1} \sum_{t=1}^n X_t$ .

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

## Definition

The process  $\{Z_t\}$  is said to be white noise with mean 0 and variance  $\sigma^2$  ( $\{Z_t\} \sim WN(0, \sigma^2)$ ) iff  $EZ_t = 0$  for all  $t$  and

$$EZ_t Z_s = \begin{cases} 0, & t \neq s \\ \sigma^2, & t = s \end{cases} .$$

## Example

An iid process with zero mean is a white noise process.

## Definition

The process  $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$  is said to be an  $ARMA(p, q)$  process (autoregressive moving average process of order  $p$  and  $q$ ) if

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}$$

where  $Z_t \sim WN(0, \sigma^2)$ .

- In applications,  $X_t, \dots, X_{t-p}$  are observed time series. But  $Z_t, \dots, Z_{t-q}$  are not observed.

- If  $X_t \sim ARMA(p_1, q_1)$  and  $Y_t \sim ARMA(p_2, q_2)$ ,  
 $X_t + Y_t \sim ARMA(P, Q)$  where  $P = p_1 + p_2$  and  
 $Q = \max(p_1 + q_2, p_2 + q_1)$ .
- For example, if  $X_t \sim AR(1)$  and  $Y_t \sim AR(1)$ ,  
 $X_t + Y_t \sim ARMA(2, 1)$ .
- Aggregation of AR processes results in an ARMA process.
- ARMA model is parsimonious compared to the AR model with many lagged variables. But it requires nonlinear optimization for estimation unlike the AR model.

- We may write the *ARMA* process as

$$\phi(B)X_t = \theta(B)Z_t$$

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

$$B^r y_t = y_{t-r}.$$

## Example

*MA*(1) process

$$X_t = Z_t + \theta Z_{t-1}, \quad Z_t \sim WN(0, \sigma^2).$$

## Example

*AR*(1) process

$$X_t = \phi X_{t-1} + Z_t.$$

## Definition

An ARMA( $p, q$ ) process defined by the equations  $\phi(B)X_t = \theta(B)Z_t$  is said to be invertible if there exists a sequence of constants  $\{\pi_j\}$  such that  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  and

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad t = 0, \pm 1, \dots$$

- In other words, an ARMA process is invertible if it can be written as an AR( $\infty$ ) process.



## Theorem

*(Invertibility) Let  $\{X_t\}$  be an ARMA( $p, q$ ) process for which the polynomials  $\phi(\cdot)$  and  $\theta(\cdot)$  have no common zeros. Then  $\{X_t\}$  is invertible iff  $\theta(z) \neq 0$  for all  $z$  such that  $|z| \leq 1$ . The coefficient  $\pi_j$  are determined by the relation*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \phi(z) / \theta(z), \quad |z| \leq 1.$$

## Example

$$\begin{cases} X_t = Z_t - \theta Z_{t-1}, Z_t \sim WN(0, \sigma^2) \\ |\theta| < 1. \end{cases}$$

$$\begin{aligned} X_t &= Z_t - \theta Z_{t-1} \\ (+\theta) X_t &= (+\theta) Z_{t-1} - \theta^2 Z_{t-2} \\ \theta^2 X_{t-2} &= \theta^2 Z_{t-2} - \theta^3 Z_{t-3} \end{aligned}$$

$$\vdots$$

$$\sum_{i=0}^{\infty} X_{t-i} \theta^i = Z_t.$$

Because  $\sum_{i=0}^{\infty} |\theta^i| < \infty$ ,  $Z_t$  is invertible.

## Theorem

*(Weak stationarity) If  $\phi(z) \neq 0$  for all  $z \in \mathbb{C}$  such that  $|z| \leq 1$ , the ARMA process  $X_t$  is weakly stationary.*

## Example

If  $X_t = \phi X_{t-1} + Z_t$  and  $|\phi| < 1$ ,

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j},$$

and  $X_t$  is stationary.

## Example

If  $X_t = \phi X_{t-1} + Z_t$ ,  $\phi = 1$ ,  $\{X_t\}$  is not stationary.

# Detrending and deseasonalizing time series

- Time series  $X_t$  is usually represented as

$$X_t = m_t + S_t + Y_t,$$

where  $m_t$  is a trend component,  $S_t$  a seasonal component and  $Y_t$  a random component.

## Example

GNP

$m_t$  : growth component,  $S_t$  : seasonal component,  $Y_t$  : economic fluctuation

# Detrending and deseasonalizing time series

## Detrending

### 1. Linear regression method

Let  $m_t = a_0 + a_1 t + \cdots + a_p t^p$ . Run OLS regression

$$X_t = \hat{m}_t + \hat{Y}_t$$

to estimate  $a_0, a_1, \cdots, a_p$ . The detrended series is

$$X_t - \hat{m}_t = \hat{Y}_t.$$

In practice, we usually use  $p = 1$ .

# Detrending and deseasonalizing time series

## Detrending

### 2. Differencing

$$\Delta X_t = X_t - X_{t-1}$$

#### Example

$$\begin{aligned} X_t &= at + Y_t \\ \Rightarrow \Delta X_t &= a + \Delta Y_t \end{aligned}$$

# Detrending and deseasonalizing time series

## Deseasonalizing

- The Bureau of Census X-12 ARIMA methods.



# Detrending and deseasonalizing time series

## Differencing to deseasonalize and detrend

- Define

$$\Delta_d X_t = X_t - X_{t-d}.$$

Applying  $\Delta_d$ , we may detrend and deseasonalize a data

$d = 4$ ; quarterly data

$d = 12$ ; monthly data.

# Detrending and deseasonalizing time series

Differencing to deseasonalize and detrend

## Example

Suppose that

$$X_t = \mu t + S_t + Y_t,$$

where  $S_t$  is the seasonal component with the property  $S_t = S_{t-d}$  and  $Y_t$  is a weakly stationary process. Then

$$\begin{aligned}\Delta_d X_t &= (\mu t - \mu(t-d)) + S_t - S_{t-d} + Y_t - Y_{t-d} \\ &= \mu d + Y_t - Y_{t-d}.\end{aligned}$$

This is a weakly stationary process.