# Introduction to Time Series Analysis

Juan de Dios Tena Horrillo Office. 10.1.02B jtena@est-econ.uc3m.es

1

### OUTLINE

Chapter 1. Univariate ARIMA models.

- Chapter 2. Model fitting and checking.
- Chapter 3. Prediction and model selection.
- Chapter 4. Outliers and influential observations.
- Chapter 5. Vector Autoregressive and Vector Error Correction Models.
- Chapter 6. Nonlinear Time Series Modelling.

### **TEXTBOOKS**

- Peña, D., Tiao, G.C. and Tsay, R.S. (2000). A Course in Time Series Analysis. Wiley & sons.
- ✓ Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Box, G.E.P., Jenkins, G.M. and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control,* 3<sup>rd</sup>. Ed. Prentice-Hall, Englewood Cliffs, NJ.
- Enders, W. (2004). Applied Econometric Time Series . Wiley & sons.
- Lütkepohl, H. and Krätzig, M. (2004). Applied Time Series Econometrics. Cambridge University Press.
- Franses, P.H. and van Dick, D. (2000). Non-linear time series models in empirical finance. Cambridge University Press.

### GRADING

• Final exam (70%).

• Class participation and empirical project (30%).

### WEB RESOURCES

o Global Insight.

o Time series data library.

http://datamarket.com/data/list/?q=provider:tsdl

• Macroeconomic time series

http://www.fgn.unisg.ch/eurmacro/macrodata/index.html

o Instituto Nacional de Estadística

http://www.ine.es

### o <u>Eurostats</u>

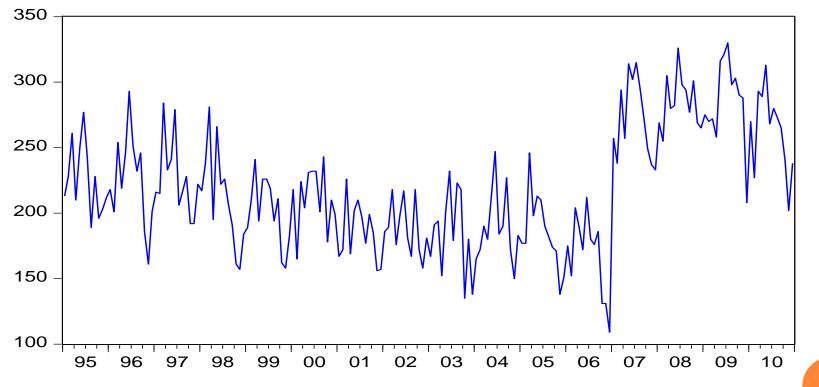
http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

 Time series: sequence of observations taken at regular intervals of time

# • Data in bussines, engineering, enviroment, medicine, etc.

#### SUICIDES IN SPAIN, TOTAL NUMBER

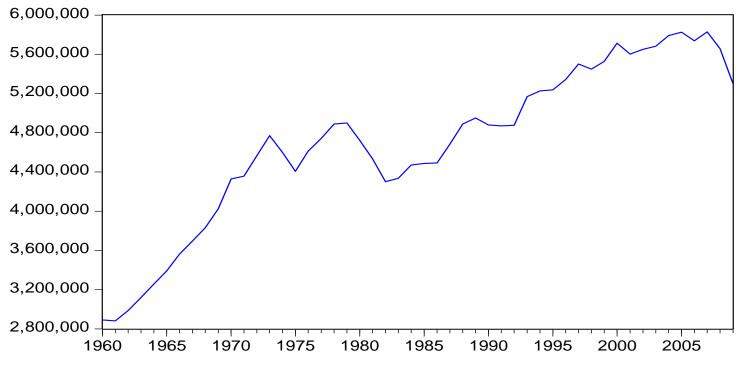
SUICIDES



Introduction to time series (2013)

#### CO2 EMISSIONS IN THE USA

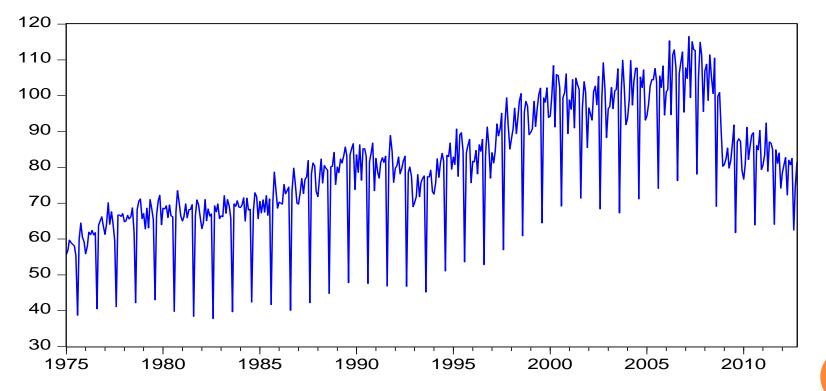
CO2



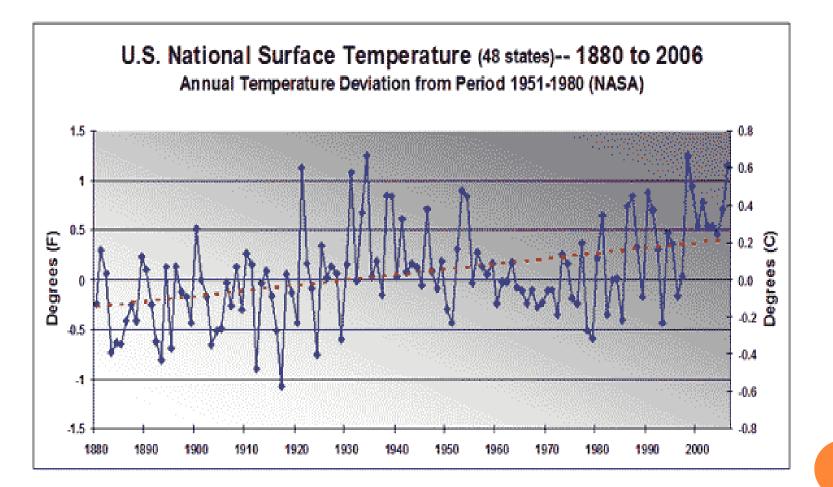
Source: World Databank

### **INDUSTRIAL PRODUCTION INDEX (SPAIN)**

IPI

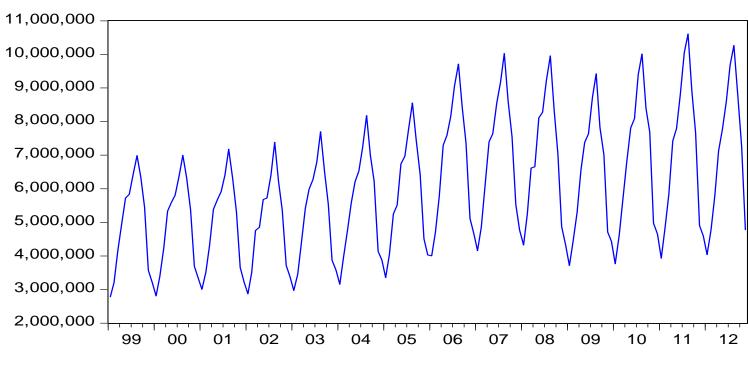


Introduction to time series (2013)



10

#### **OVERNIGHT STAYS IN SPANISH HOTELS**



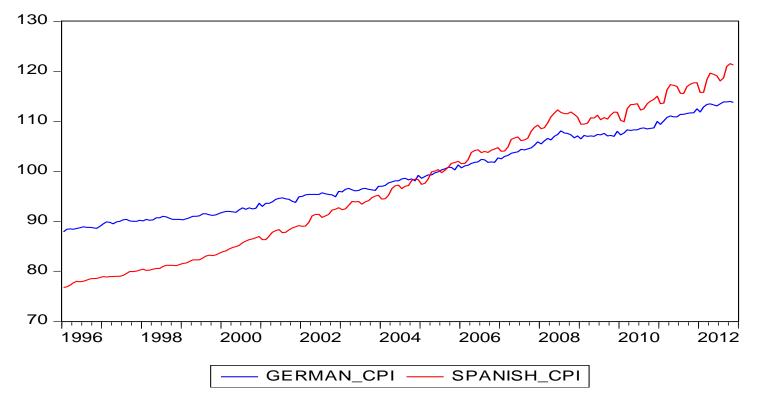
TURISTS

Source: ine

Introduction to time series (2013)

11

#### **CONSUMER PRICE INDEX IN GERMANY AND SPAIN**

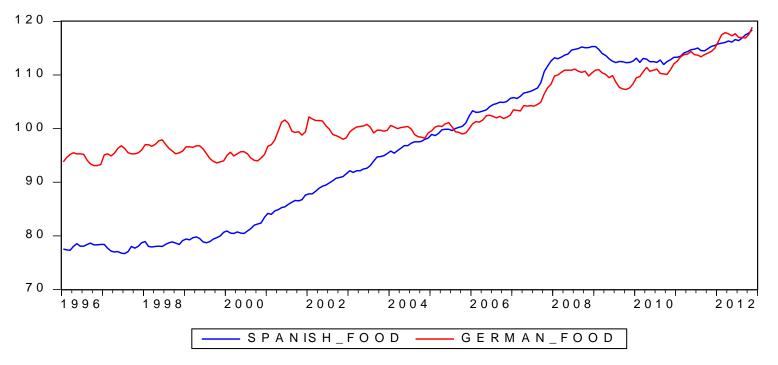


Introduction to time series (2013)

12

Source: Eurostats

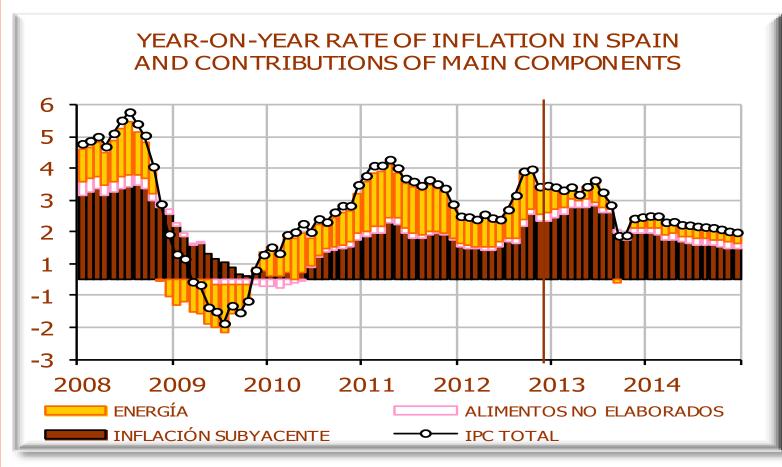
### CONSUMER PRICE INDEX FOR FOOD AND NON-ALCOHOLIC BEVERAGES IN GERMANY AND SPAIN



Introduction to time series (2013)

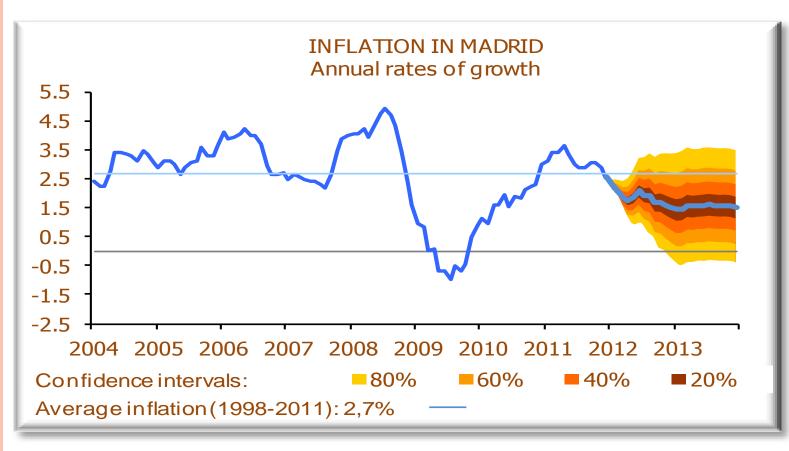
13

Source: Eurostats



14

Source: Instituto Flores de Lemus



Source: Instituto Flores de Lemus



Introduction to time series (2013)

16

#### Decomposition of additive time series

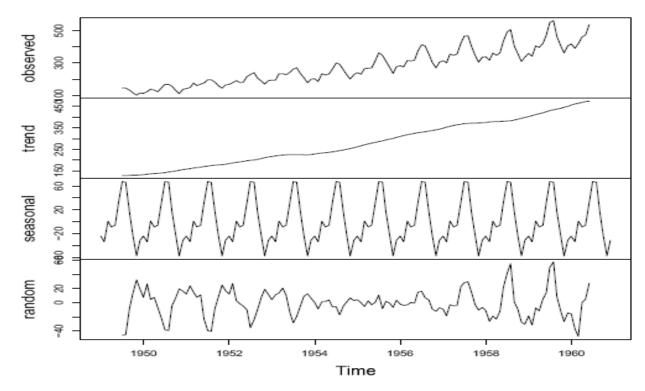


Figure: Additive seasonal decomposition of the airline data

 Knowledge of the dynamic structure will help to produce accurate forecasts of future observations and design optimal control schemes.

# Main objectives

- Understanding the dynamic structure of the observations in a single series.
- Ascertain the leading, lagging and feedback relationships among several variables.
- Analyzing the impact of other variables that can be used for policy decisions.
- What to do when there are breaks in the evolution of the series.

# Chapter 1

# **Univariate ARIMA Models**

Recommended readings: chapters 1 to 3 of Peña et al. (2000)

# CHAPTER 1. CONTENTS.

### **1.1. Introduction.**

Definitions, examples,

### **1.2.** Properties of univariate ARMA.

ARMA weights. Stationarity and covariance structure. Autocorrelation function.

Partial autocorrelación function.

- **1.3. Model especification.**
- 1.4. Examples.

• Stochastic Process. Real-valued random variable  $Z_t$  that follows a distribution  $f_t(Z_t)$ .

• The T-dimensional variable  $Z_{t_1}, Z_{t_2}, \dots, Z_{t_T}$ will have a joint distribution that depends on  $t_1, \dots, t_T$ .

• Time series  $z_{t_1}, z_{t_2}, \dots, z_{t_T}$  will denote a particular realization of the stochastic process.

**O** To simplify notation  $z_1, z_2, ..., z_T$ .

### Assumptions

1. The process is stationary.

2. The joint distribution of  $z_1, z_2, ..., z_T$  is a multivariate normal distribution.

Implications of stationarity

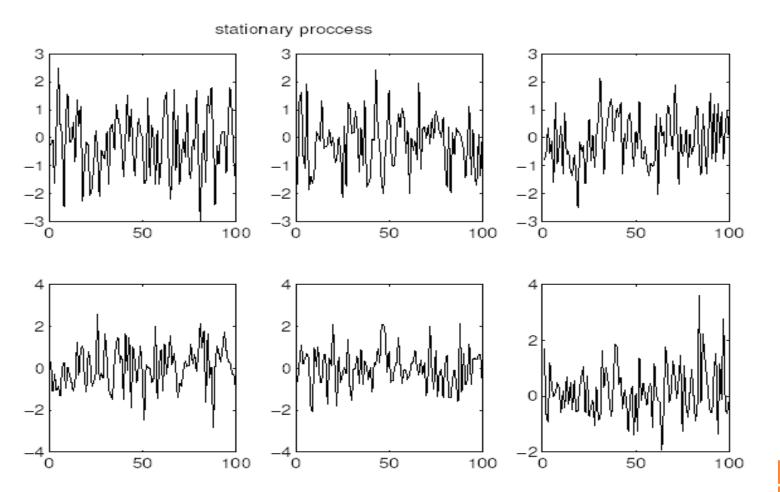
- The joint distribution remains constant over time
- It is then true that

$$f(z_1, z_2, \dots, z_T) = f(z_{1+k}, z_{2+k}, \dots, z_{T+k})$$

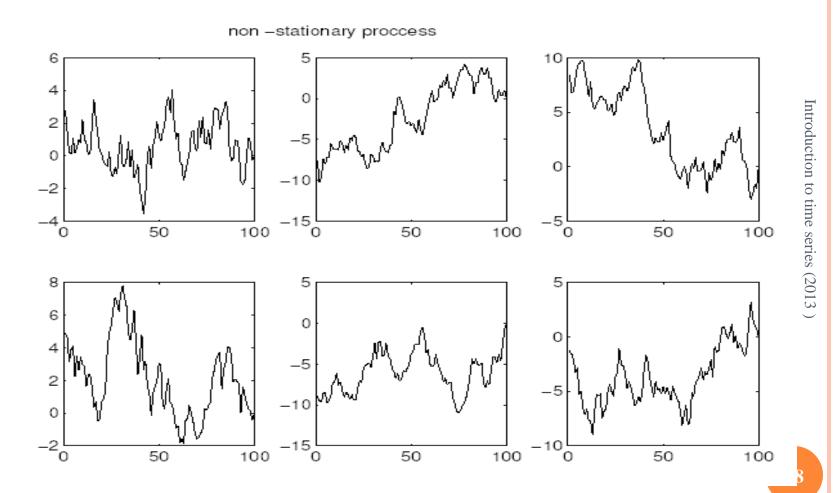
 $Ez_t = \mu_z; \quad Vz_t = V_z$ 

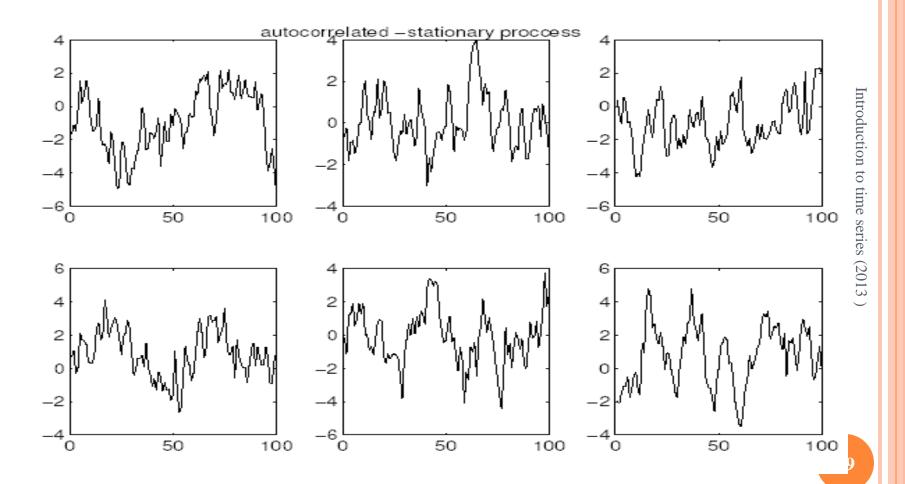
25

- Stationarity implies a constant mean and bounded deviations from it.
- Strong requirement, few actual economic series will satisfy it.



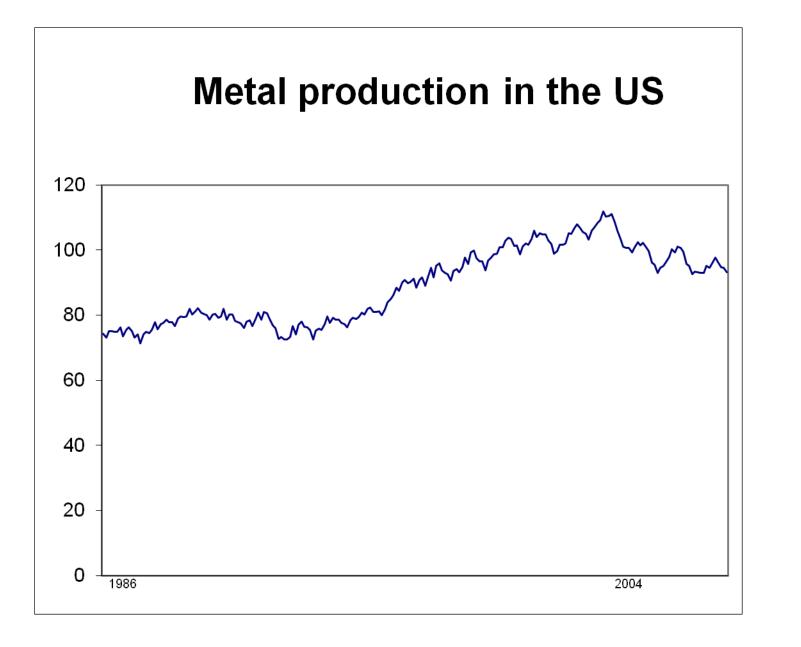
Introduction to time series (2013)





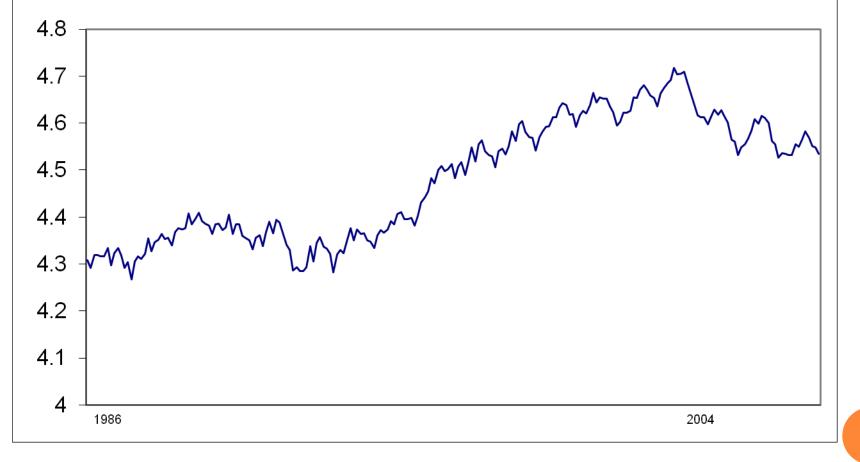
### Transformations to achieve stationarity

- Constant variance: log/level plus outlier correction.
- Stationary in mean: differencing.



31





Introduction to time series (2013)

32

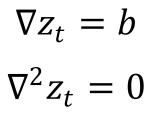
**Differencing** • Let B be the backward operator  $B^{j}z_{t} = z_{t-j}$ 

• We shall use the operators:

- Regular difference
- Seasonal difference

$$\nabla = 1 - B$$
$$\nabla_s = 1 - B^s$$

### If $z_t$ is a deterministic trend ( $z_t = a + bt$ )



### where

$$\nabla^2 z_t = \nabla(\nabla z_t)$$

# O In general, ∇<sup>d</sup> will reduce a polynomial of degree d to a constant.

• Example: for a quarterly series  $z_t$ ,  $\nabla_4 z_t = z_t - z_{t-4}$ will cancel a constant, but it will also cancel other deterministic periodic functions.

Homogeneous difference equation

$$\nabla_4 z_t = (1 - B^4) z_t = z_t - z_{t-4}$$

Characteristic equation:  $r^4 - 1 = 0$ Solution is given by  $r = \sqrt[4]{1}$  (four roots on the unit series (2013 circle)

$$r_1 = 1, r_2 = -1, r_3 = i, r_4 = -i$$

Two real roots and two complex conjugates with modulus 1 and frequency  $\omega = \pi/2$ .

#### Therefore,

$$\nabla_4 = (1 - B)(1 + B)(1 + B^2)$$

- 1. One in the zero frequency (trend)
- 2. One in the twice-a-year seasonality  $\omega = \pi$
- 3. Associated with once-a-year seasonality  $\omega = \pi/2$

To see this just recall that the complementary function to the difference equation  $z_t = bz_{t-1} + cz_{t-2}$  will eventually be of the form:

 $z_t = (r^t)[A\cos\theta t + B\sin\theta t]$ 

where r is a positive constant and  $\theta$  is an angle measured in radians. A and B are arbitrary constants to enable the solution to satisfy any starting point of  $z_t$ .

Hence, when the two solutions of the characteristic equations are *i* and -i,  $\sin \theta = 1$  and  $\theta = \frac{\pi}{2}$ .

# UNIVARIATE ARMA MODELS

"Building block" is the white noise process

$$a_t \succ Niid(0, \sigma^2)$$

Implications:

$$E(a_{t}) = E(a_{t} / a_{t-1}, a_{t-2}, a_{t-3}, ...) = 0$$
  

$$E(a_{t}, a_{t-j}) = Cov(a_{t}, a_{t-j}) = 0$$
  

$$Var(a_{t}, a_{t-j}) = Var(a_{t} / a_{t-1}, a_{t-2}, ...) = \sigma_{a}^{2}$$

39

# General ARMA(p,q) processes:

$$\phi(B)z_t = c + \theta(B)a_t$$

- $\bigcirc$   $z_t$  is the observable time series.
- $\circ a_t$  is sequence of white noise.
- *c* is the constant term.

#### Autoregressive polynomial

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

Moving-average polynomial

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

 The AR and MA polynomials are assumed to have no common factor

• Stationarity implies that all zeros of  $\phi(B)$  are restricted to lie outside the unit circle and in this case:

$$c = (1 - \phi_1 - \dots - \phi_p)\mu$$

where  $\mu$  is the mean of the series.

 Overall behavior of the series remains the same over time.

- Overall behavior of the series remains the same over time.
- In real world, however, time series data often exhibits a drifting behavior. Nonstationarity can be modeled allowing some of the zeroes in  $\phi(B)$  to be equal to one. Thus,

$$\phi(B)(1-B^d)z_t = c + \theta(B)a_t$$

• This is known as the ARIMA (p,d,q) model.

• Some special cases of ARIMA(p,d,q):

- AR(1)
- MA(1)
- ARMA(1,1)
- IMA(1,1)
- ARIMA(0,1,1)(0,1,1)

• AR(1) to MA( $\infty$ ) by recursive substitution.

$$z_{t} = \phi z_{t-1} + a_{t}$$

$$z_{t} = \phi (\phi z_{t-2} + a_{t-1}) + a_{t} = \phi^{2} z_{t-2} + \phi a_{t-1} + a_{t}$$

$$z_{t} = \phi^{k} z_{t-k} + \phi^{k-1} a_{t-k+1} + \dots + \phi^{2} a_{t-2} + \phi a_{t-1} + a_{t}$$

# • If $|\phi| < 1$ , then

$$z_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}$$

#### which is a MA( $\infty$ )

- o Some considerations:
- ARMA models are not unique (variety of possible representations.
- Parsimony is always a desirable property.
- AR representations are easiest to estimate (OLS) and to be interpretated.

 For simplicity, we will assume zero mean and a starting point m. The ARIMA (p,d,q)

$$\phi(B)z_t = \theta(B)a_t$$

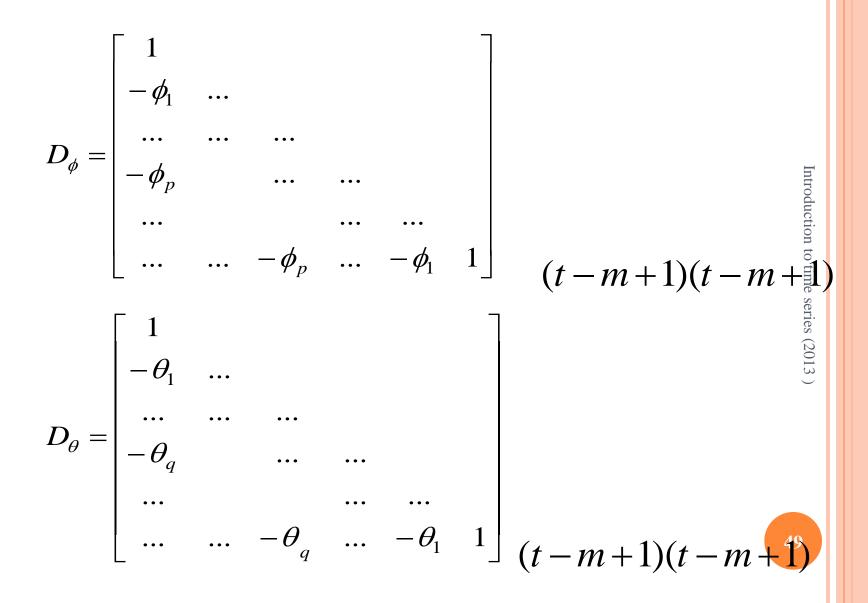
• Can be rewritten as:

$$D_{\phi}z = D_{\theta}a + \omega$$

where, 
$$z = (z_m, ..., z_t)'$$
  $a = (a_m, ..., a_t)'$ 

$$\omega = (w_m, \dots, w_{m-r-1}, 0, \dots, 0)'$$

 $W_m, ..., W_{m-r-1}$  are r=max(p,q) initial values.



#### • ARMA weights

$$z = D_{\phi}^{-1} D_{\theta} a + D_{\phi}^{-1} \omega$$

#### Where,

$$D_{\phi}^{-1}D_{\theta} = \begin{bmatrix} 1 & & & \\ \psi_{1} & & & \\ & \ddots & & & \\ \psi_{t-m} & \cdots & \psi_{1} & 1 \end{bmatrix}$$

• The  $\Psi$ -weights can be obtained by equating coefficients of powers of B from the relations:

$$\phi(B)\psi(B) = \theta(B)$$

where

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

• The ARIMA (p,d,q) model can then be rewritten in the MA form as:

$$z_t = a_t + \sum_{h=1}^{t-m} \psi_h a_{t-h}$$

#### • In the same way,

$$D_{\theta}^{-1}D_{\phi}z - D_{\theta}^{-1}\omega = a$$

• Where,

$$D_{\theta}^{-1} D_{\phi} = \begin{bmatrix} 1 & & & \\ -\pi_{1} & & & \\ & \ddots & & & \\ -\pi_{t-m} & \cdots & -\pi_{1} & 1 \end{bmatrix}$$

53

# • With $\theta(B)\pi(B) = \phi(B)$ .

• The AR form of the model is then,

$$z_t = \sum_{h=1}^{t-m} \pi_h z_{t-h} + a_t$$

- These two expressions are of fundamental importance in understanding the nature of the model.
- The <u>MA form</u> with the  $\Psi$  weights, shows how the observation  $z_t$  is affected by current and past shocks or innovations.
- The <u>AR form</u> with the  $\pi$  weigths, indicates how the observation is related to its own past values.

#### Examples : AR(1) model

• MA representation:

$$z_t = a_t + \phi_1 a_{t-1} + \phi^2 a_{t-2} + \dots$$

• AR representation:

$$z_t = \phi z_{t-1} + a_t$$

#### Examples : MA(1) model

• MA representation:

$$z_t = a_t - \theta a_{t-1}$$

• AR representation:

$$z_t = -\theta z_{t-1} - \dots - \theta^j z_{t-j} - \dots - a_t$$

Introduction to time series (201

**58** 

#### Examples : IMA(1,1) model

• MA representation:

$$z_t = a_t + (1 - \theta)(a_{t-1} + a_{t-2} + \dots)$$

• AR representation:

$$z_t = (1-\theta)z_{t-1} + \theta(1-\theta)z_{t-2} + \theta^2(1-\theta)z_{t-3} + \dots + \ddot{a}_t$$

# STATIONARITY CONDITION

o Consider,

$$z_t = a_t + \sum_{h=1}^{t-m} \psi_h a_{t-h}$$

- Since the innovations are assumed normally distributed, it follows that the observations are also normally distributed.
- It is seen that, if in the characteristic equation:

$$\psi_{k} = A_{1}r_{1}^{k} + \dots + A_{p0}r_{p0}^{k}$$

### STATIONARITY CONDITION

• (Where the A's denote polinomials in k and the r's are the distinct zeroes of  $\phi(B)$ )

o Then as 
$$\left| r_{j} 
ight| < 1,$$
 we have that

$$t - m \rightarrow \infty$$
,

 $E(z_t) \rightarrow 0, \quad \operatorname{cov}(z_t, z_{t+k}) \rightarrow \sigma_a^2 \left(\sum_{h=0}^{\infty} \psi_h \psi_{h+k}\right)$ 

# STATIONARITY CONDITION

• So that  $z_t$  will be stationary in this asymptotic sense.

• This is the stationarity condition of an ARMA model, and is equivalent to require that the zeroes of  $\phi(B)$  are lying outside the unit circle.

# LAG K AUTOCOVARIANCE

#### • Let us denote,

$$\gamma(k) = \operatorname{cov}(z_t, z_{t+k}) = \operatorname{cov}(z_t, z_{t-k}) = \gamma(-k)$$

• For an alternative expressions of the autocovariance function in terms of the ARMA parameters,

$$z_{t-k}(z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p}) = z_{t-k}(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})$$

### LAG K AUTOCOVARIANCE

• By taking expectations on both sides and using the MA form, we obtain, for  $k \ge 0$ ,

$$\gamma(k) = \sum_{h=1}^{p} \phi_h \gamma(k-h) + g_k$$

• Where,

$$g_{k} = \begin{cases} -\sigma_{a}^{2} \sum_{h=0}^{q-k} \psi_{h} \theta_{h+k} & k = 0, \pm 1, \pm 2, \dots \\ 0 & k > q \end{cases}$$

### AUTOCORRELATION FUNCTION

• The autocorrelation function is defined as

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}, \quad k = 0, \pm 1, \pm 2, \dots$$

o By substitution of  $\gamma(k)$ , we obtain that

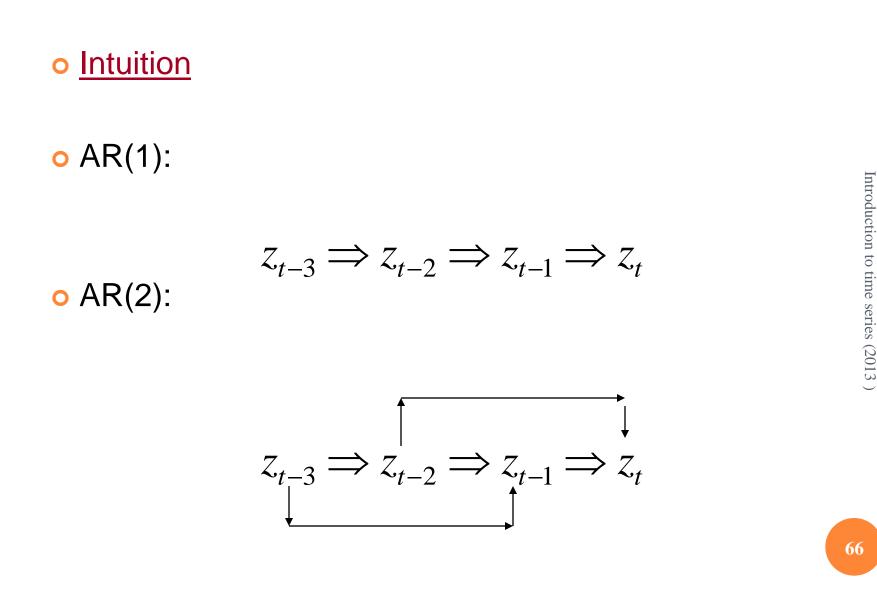
$$\rho(k) = \sum_{h=1}^{p} \phi_h \rho(k-h) + \frac{g_k}{\gamma(0)} \quad k \ge 0$$

# AUTOCORRELATION FUNCTION

• When  $\phi(B) = 0$ , in a MA(q) model, then,

$$\rho(k) = \begin{cases} -\theta_q (1 + \theta_1^2 + \dots \theta_q^2)^{-1}, & k = q \\ 0, & k > q \end{cases}$$

 That is, the autocorrelation function cuts off after lag q.



- The autocorrelation function only takes into account that  $z_t$  and  $z_{t-2}$  are related in both cases.
- But if we want to measure the direct relationship (without the intermediate  $Z_{t-1}$ ), we found that it is zero for the AR(1) and different from zero for the AR(2) model

 <u>The partial autocorrelation function</u> is, therefore, a measure of the linear relation among observations kperiods apart, independently of the intermediate values.

 Consider first an stationary AR(p) model. The autoregressive coefficients are related to the autocorrelations by the Yule-Walker equations,

 $G_p \phi_p = \rho_p$ 

• where, $G_p$  is the (pxp) matrix,

$$G_{p} = \begin{bmatrix} 1 & \rho(-1) & \dots & \rho(-p+2) & \rho(-p+1) \\ \rho(1) & 1 & \vdots & \rho(-p+2) \\ \vdots & & \ddots & & \vdots \\ \rho(p-2) & & 1 & \rho(-1) \\ \rho(p-1) & \rho(p-2) & \dots & \rho(1) & 1 \end{bmatrix}$$

70

• Regarding this as system of p equations and p unknowns (the  $\phi$  coefficients), the solution is, for p>1, the ratio of two determinants

$$\phi_p = |H_p| / |G_p|$$

### where

$$H_{p} = \begin{bmatrix} 1 & \rho(-1) & \dots & \rho(-p+2) & \rho(1) \\ \rho(1) & 1 & \vdots & \rho(2) \\ \vdots & \ddots & & \vdots \\ \rho(p-2) & & 1 & \rho(p-1) \\ \rho(p-1) & \rho(p-2) & \dots & \rho(1) & \rho(p) \end{bmatrix}$$

## PARTIAL AUTOCORRELATION FUNCTION

•A pxp matrix, and  $\phi_p = \rho(1)$  for p=1. This leads to define, for any stationary model

$$\varphi_k \begin{cases} \rho(k) & k = 1 \\ |H_k| / |G_k| & k > 1 \end{cases}$$

•Which is known as the Partial Autocorrelation Function.

## PARTIAL AUTOCORRELATION FUNCTION

It has the property that, for a stationary AR(p) model,

$$\varphi_k = \begin{cases} \phi_k & k = p \\ 0 & k > p \end{cases}$$

In other words,  $\varphi_k$  vanishes for k>p when the model is AR(p).

## PROPERTIES

	SAF	PAF
AR(P)	slow decay towards zero	p different from zero
MA(q)	q different from zero	slow decay towards zero
ARMA(p,q)	Slow decay towards zero	Slow decay towards zero

## EXAMPLES

## o <u>AR(1) model</u>

• Variance:

• Model:

$$\sigma_z^2 = \phi^2 \sigma_z^2 + \sigma_a^2$$

 $z_t = \phi z_{t-1} + a_t$ 

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \phi^2}$$

Introduction to time series (2013)

### • Autocovariance function:

$$\gamma(k) = \begin{cases} \sigma_z^2 & k = 0 \\ \frac{\phi \sigma_a^2}{1 - \phi^2} & k = 1 \\ \phi \gamma(k - 1) & k > 1 \end{cases}$$

### • Autocorrelation function:

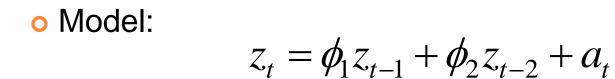
$$\rho(k) = \begin{cases} 1 & k = 0 \\ \phi & k = 1 \\ \phi^k & k > 1 \end{cases}$$

• Partial autocorrelation:

$$\varphi(1) = \rho(1)$$

## EXAMPLES

## o AR(2) model



### • Variance:

$$\sigma_{z}^{2} = \left(\frac{1-\phi_{2}}{1+\phi_{2}}\right) \frac{\sigma_{a}^{2}}{\left\{\left(1-\phi_{2}\right)^{2}-\phi_{1}^{2}\right\}}$$

### • Autocorrelation function:

$$\rho(k) = \begin{cases} \frac{\phi_1}{1 - \phi_2} & k = 1 \\ \phi_2 + \frac{\phi_1^2}{1 - \phi_2} & k = 2 \end{cases}$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad k > 2$$

## EXAMPLES

## o MA(1) model

• Model:

$$z_t = a_t - \theta a_{t-1}$$

• Variance:

$$\sigma_z^2 = \sigma_a^2 + \theta^2 \sigma_a^2$$
$$\sigma_z^2 = \sigma_a^2 (1 + \theta^2)$$

### • Autocovariance function:

$$\gamma(k) = \begin{cases} \sigma_z^2 & k = 0 \\ -\theta \sigma_a^2 & k = 1 \\ 0 & k > 1 \end{cases}$$



### • Autocorrelation function:

$$\rho(k) = \begin{cases} 1 & k = 0 \\ -\theta & k = 1 \\ 1 + \theta^2 & k = 1 \\ 0 & k > 1 \end{cases}$$

- The class of ARMA(p,q) models is extensive.
- Guidelines are needed in selecting a member of the class to represent the time series data at hand

- Box and Jenkins (1976) proposed an iterative model building strategy.
  - Tentative specification or identification of a model.
  - Efficient estimation of model parameters.
  - Diagnostic checking of fitted model for further improvement.

- <u>Tentative specification</u>. The aim is to employ statistics that:
  - 1. Can be readily calculated from the data.
  - 2. Allow the user to tentatively select a model.

- Three methods:
  - The Sample Autocorrelation Function (SAF)
  - The Sample Partial Autocorrelation Function (SPAF)
  - Use of diagnostic tools (AIC,BIC) to find the best model (TRAMOautomatic).

# • Sample Autocorrelation Function. The SACF of $z_t$ are defined as

$$\hat{\rho}_k = C_k / C_0 \quad k = 1, 2, \dots$$

with 
$$C_{j} = \sum_{t=1}^{n-j} (z_{t} - \overline{z})(z_{t+j} - \overline{z})$$

and  $\overline{z}$  the sample mean of the n available observations

#### • Properties:

o 1.For stationary models,

$$\hat{\rho}_k \to \rho_k \quad n \to \infty$$

### 2. When there exists a unit root in the AR polynomial

$$\hat{\rho}_k \xrightarrow{p} 1$$

• If the SACF of the original series is persistently close to 1 as k increases, one forms the first difference,  $\nabla z_t$ , and studies its SACF to determine whether further differencing is called for.

 Once stationary is achieved, a cutting off pattern after, say a lag "q", in the SACF will then lead to tentative specification of a MA(q) model.

• The Sample Partial Autocorrelation Function:

$$\hat{\varphi}_k \quad k=1,2,\ldots$$

• Are obtained by replacing the  $\mathcal{P}_k$  in

$$\varphi_k \begin{cases} \rho(k) & k = 1 \\ |H_k| / |G_k| & k > 1 \end{cases}$$

o by their sample estimates  $\hat{
ho}_k$ 

### Properties:

o 1.For stationary models,

$$\hat{\varphi}_k \to \varphi_k \quad n \to \infty$$

• 2. The  $\hat{\varphi}_k$  are asymptotically normally distributed • 3. For a stationary AR(p) model

$$Var(\hat{\varphi}_k) \cong n^{-1} \quad k > p$$

- Properties 1, 2 and 3 make SPACF a convenient tool for specifying the order or a stationary AR model (cutting off pattern after lag p)
- Not valid for non-stationary models.

### Weakness of the SACF and SPACF.

- 1. Subjective judgment is often required to decide on the order of differencing.
- 2. For stationary ARMA models, both SACF and SPACF tend to exhibit a gradual "tapering off" behavior, making the specification very difficult.

### **Diagnostic Tools**

- The program (TRAMO), in an automatic way, specifies a set of possible models, estimate them and select the best one based on AIC and BIC criteria.
- However an accurate jugment is always necessary to interprect the output of TRAMO.