Predictive Systems

Pr. El AFIA

Outline: (PART 2)

Chapter 3 : Linear Stochatsic Models

Linear Stationary Models

- AR
- MA
- ARMA

Linear NonStationary Models

- ARIMA
- SARIMA

Autoregressive Moving Average Process of order $(1, 1)$

We may also entertain combinations of AR and MA models:

$ARMA(1,1)$: $$z_t - \varphi_1 z_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

By backshift formulation:

$$(1 - \varphi_1 B) z_t = (1 - \theta_1 B) \varepsilon_t$$

So

$$z_t = \frac{(1 - \theta_1 B)}{(1 - \varphi_1 B)} \varepsilon_t$$

We know that any model can be converted into $MA(\infty)$ or $AR(\infty)$.

Autoregressive Moving Average Process of order $(1,1)$

Thus, the ψ-weights in $MA(\infty)$ will be:

$$\psi(B) = \frac{(1-\theta_1 B)}{(1-\varphi_1 B)}$$

So:

$$z_t = \varepsilon_t + (\varphi_1 - \theta_1)\sum_{i=1}^{\infty} \varphi^{i-1}\varepsilon_{t-i}$$

And, the π-weights in $AR(\infty)$ will be:

$$\pi(B) = \frac{(1-\varphi_1 B)}{(1-\theta_1 B)}$$

So:

$$z_t = (\varphi_1 - \theta_1)\sum_{i=1}^{\infty} \theta^{i-1} z_{t-i} + \varepsilon_t$$

Autoregressive Moving Average Process of order $(1, 1)$

Stationarity and Invertibility conditions of ARMA(1,1)

So that the ψ-weights converge, the stationarity condition is:

$$|\varphi_1| < 1$$

So that the π-weights converge, the invertibility condition is:

$$|\theta_1| < 1$$

ACF of ARMA(1,1)

Let's compute the ACF of the ARMA(1,1) process:

$$z_t - \varphi_1 z_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

We multiply each term by z_{t-k} and taking the expectation:

$$E(z_{t-k} z_t) - \varphi_1 E(z_{t-k} z_{t-1}) = 0$$

Autoregressive Moving Average Process of order $(1,1)$

For $k > 1$:

$$\gamma_k = \varphi_1 \gamma_{k+1}$$

For $k = 0$:

$$\gamma_1 - \varphi_1 \gamma_0 = -\theta_1 \sigma^2$$

For $k = 1$:

$$\gamma_0 - \varphi_1 \gamma_1 = \sigma^2 - \theta_1(\varphi_1 - \theta_1)\sigma^2$$

Eliminating σ^2 from these two equations allows the ACF of the ARMA(1,1) process to be:

$$\begin{cases} \rho_1 = \dfrac{(1-\varphi_1\theta_1)(\varphi_1 - \theta_1)}{1 + \theta_1^2 - 2\theta_1\varphi_1} \quad for\ k = 1 \\ \rho_k = \varphi_1 \rho_{k-1} \quad for\ k > 1 \end{cases}$$

Autoregressive Moving Average Process of order (p, q)

This process is obtained by combining $AR(p)$ and $MA(q)$:

$$z_t - \varphi_1 z_{t-1} - \varphi_2 z_{t-2} - \cdots - \varphi_p z_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

Or using the backshift notation:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) \varepsilon_t$$

$$\varphi(B) z_t = \theta(B) \varepsilon_t$$

Autoregressive Moving Average Process of order (p, q)

Stationarity conditions of $\boldsymbol{ARMA(p,q)}$

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

Therefore, the coefficients in $\psi(B)$ can be obtained by:

$$\psi_0 = 1$$

and

$$\psi_i - \varphi_1\psi_{i-1} - \varphi_2\psi_{i-2} - \cdots - \varphi_p\psi_{i-p} = \begin{cases} -\theta_i & if \quad i = 1, \dots, q \\ 0 & if \quad i > q \end{cases}$$

The polynomial characteristic equation:

$$m^p - \varphi_1 m^{p-1} - \varphi_2 m^{p-2} - \cdots - \varphi_p = 0$$

Autoregressive Moving Average Process of order (p, q)

Stationarity conditions of $\boldsymbol{ARMA(p,q)}$

The stationarity of the ARMA process is related to the AR component in the model and can be checked through the roots of the polynomial characteristic equation:

$$m^p - \varphi_1 m^{p-1} - \varphi_2 m^{p-2} - \cdots - \varphi_p = 0$$

The ARMA model is stationary only if the roots of the polynomial characteristic equation are such that:

$$|\boldsymbol{m_i}| < \mathbf{1} \text{ for } i = 1, \dots, p$$

Autoregressive Moving Average Process of order (p, q)

Invertibility conditions of $\boldsymbol{ARMA(p,q)}$

$$\varphi(B) = \theta(B)\pi(B)$$

Therefore, the coefficients in $\pi(B)$ can be obtained by:

$$\pi_0 = -1$$

and

$$\pi_i - \theta_1\pi_{i-1} - \theta_2\pi_{i-2} - \cdots - \theta_p\pi_{i-p} = \begin{cases} \varphi_i & if \quad i = 1, \dots, p \\ 0 & if \quad i > p \end{cases}$$

The associated polynomial characteristic equation:

$$m^q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \cdots - \theta_q = 0$$

Autoregressive Moving Average Process of order (p, q)

Invertibility conditions of $\boldsymbol{ARMA(p, q)}$

Similar to the stationarity condition, the invertibility of an ARMA process is related to the MA component and can be checked through the roots of the associated polynomial characteristic equation:

$$m^q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \cdots - \theta_q = 0$$

The ARMA model is invertible only if the roots of the polynomial characteristic equation are such that:

$$|\boldsymbol{m_i}| < \mathbf{1} \text{ for } i = 1, \dots, q$$

Autoregressive Moving Average Process of order (p, q)

The ACF and PACF for ARMA(p,q)

As in the stationarity and invertibility conditions, the ACF and PACF of an ARMA process are determined by the AR and MA components, respectively.

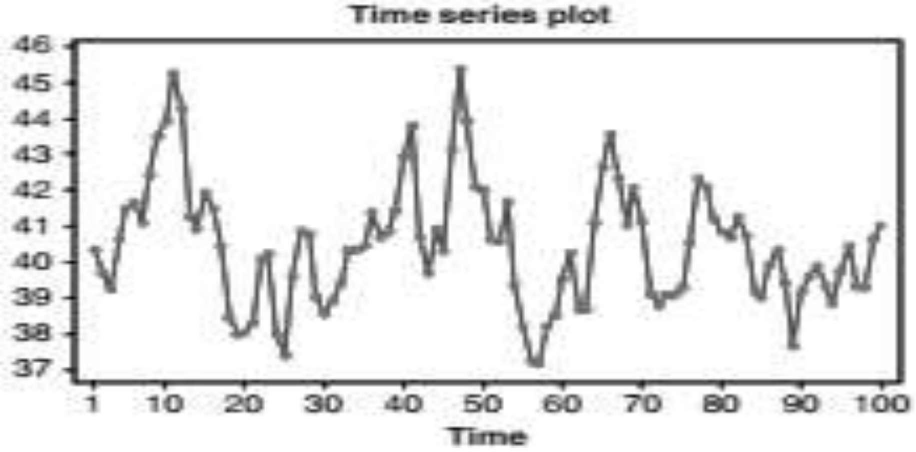
- as $MA(q)$.
- For $k > q - p \quad \Longrightarrow \quad$ ACF will follow the same patterns as $AR(p)$.
- For $k > q - p \quad \Longrightarrow \quad$ PACF will follow the same patterns

It can, therefore, be shown that the ACF and PACF of ARMA exhibit a mixture of exponential decay and/or sinusoidal decay.

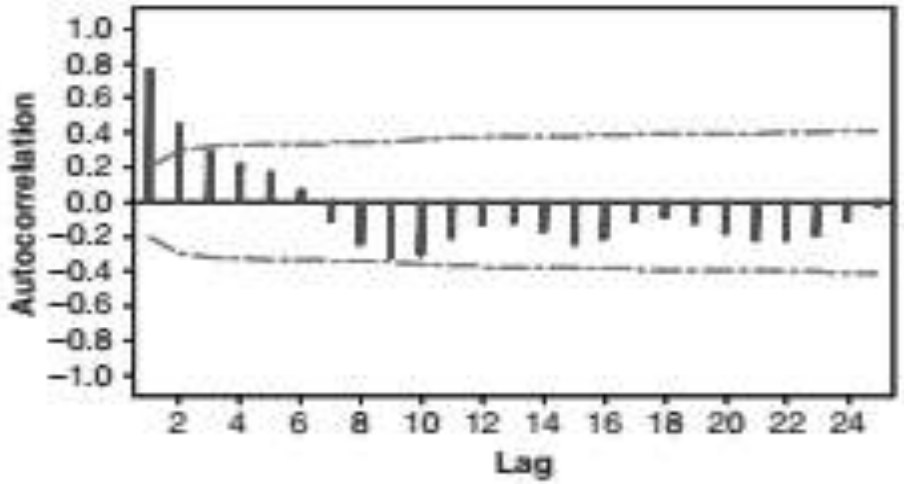
Autoregressive Moving Average Process of order (p, q)

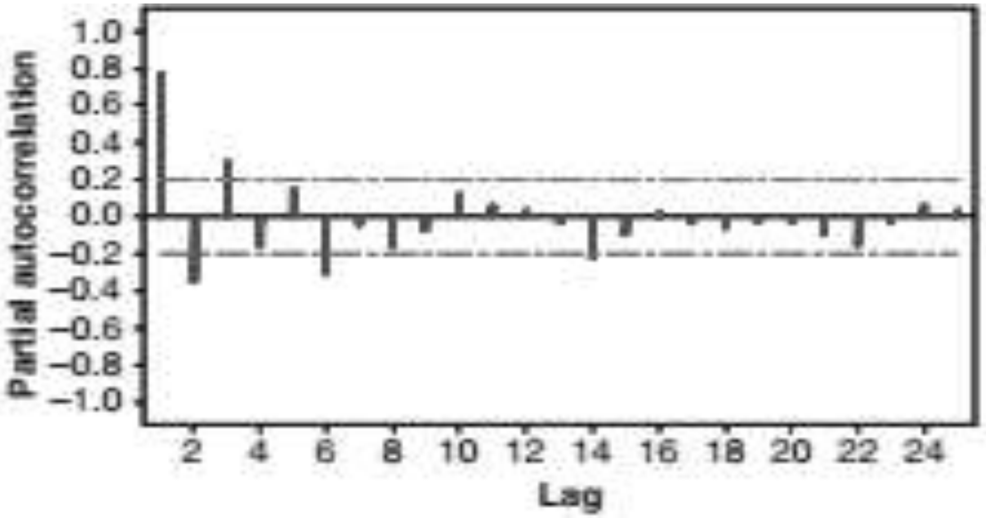
Model	ACF	PACF
MA(q)	Cuts off after lag q	Exponential decay and/or damped sinusoid
AR(p)	Exponential decay and/or damped sinusoid	Cuts off after lag p
ARMA(p,q)	Exponential decay and/or damped sinusoid	Exponential decay and/or damped sinusoid

Autoregressive Moving Average Process of order (p, q)

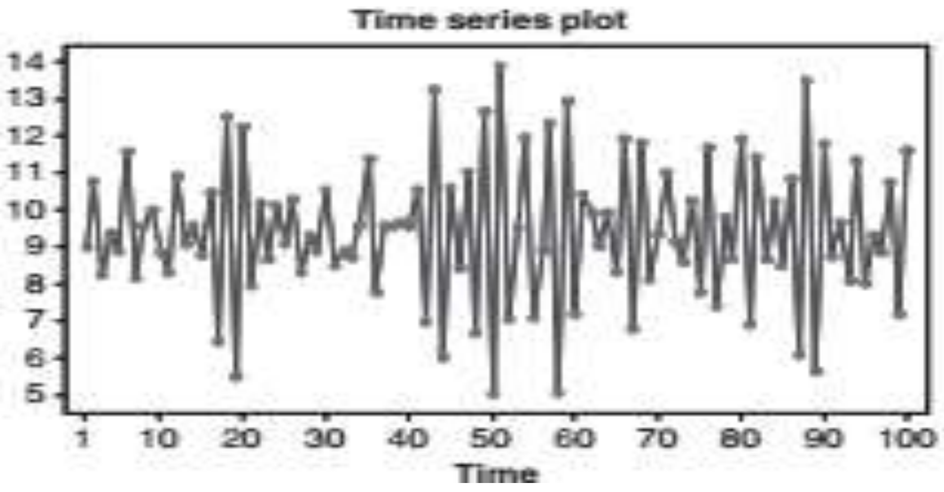


$$z_t = 16 + 0.6z_{t-1} + \varepsilon_t + 0.8\varepsilon_{t-1}$$

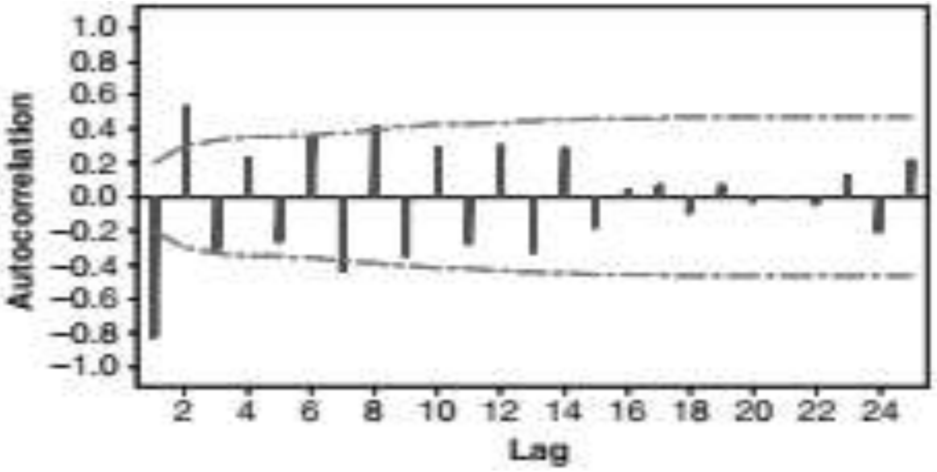


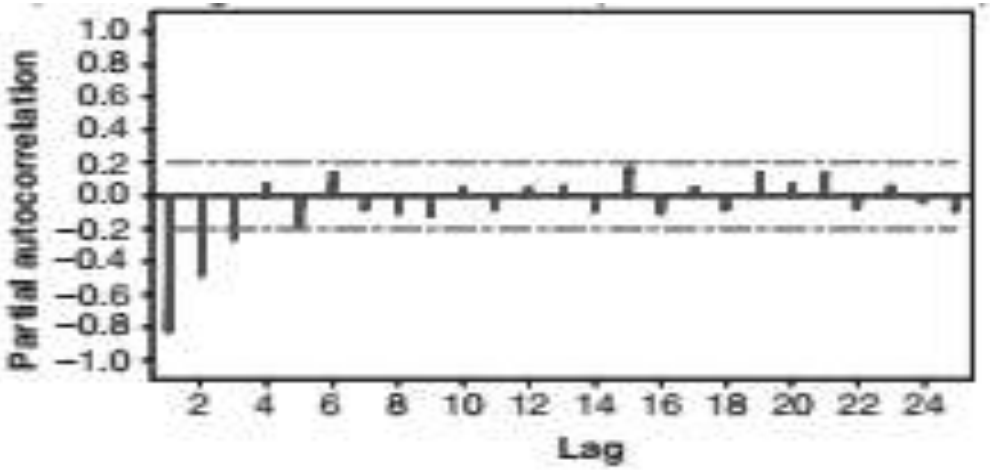


Autoregressive Moving Average Process of order (p, q)



$$z_t = 16 - 0.7z_{t-1} + \varepsilon_t - 0.6\varepsilon_{t-1}$$





Autoregressive Moving Average Process of order (p, q)

All the development had assumed that $\mu = 0$. Nonzero means are easily accommodated by replacing z_t by $z_t - \mu$:

$ARMA(p, q)$: $$\varphi(B)(z_t - \mu) = \theta(B)\varepsilon_t$$

Note that:

$$\varphi(B)\mu = \left(1 - \varphi_1 - \cdots - \varphi_p\right)\mu = \varphi(1)\mu$$

So:

$ARMA(p, q)$: $$\varphi(B)z_t = \varphi(1)\mu + \theta(B)\varepsilon_t = \theta_0 + \theta(B)\varepsilon_t$$

Such that $\theta_0 = \varphi(1)\mu$ is a constant.

Autoregressive Integrated Moving Average Process of order $(\mathrm{p}, \mathrm{d}, \mathrm{q})$

Definition: homogenous nonstationary time series

We call a time series $(z_t)_t$ homogenous nonstationary if it is not stationary but its trend difference of order d, that is, $z_t^{(d)} = (1 - B)^d z_t$ produce a stationary time series.

Notice:

We say that $(z_t)_t$ is generated by an ARIMA(p,d,q) process. Or, $\left(z_t^{(d)}\right)_t$ is generated by an ARMA(p,q) process.

Autoregressive Integrated Moving Average Process of order $(\mathrm{p}, \mathrm{d}, \mathrm{q})$

$ARIMA(p, d, q)$ can be represented as: $\forall t \in \{\mathrm{d}+1, \dots, n\}$

$$z_t^{(d)} = (1-B)^d z_t$$

Such that $\forall t \in \{\max(p, q) + 1, \dots, n\}$:

$$z_t^{(d)} = c + \varphi_1 z_{t-1}^{(d)} + \cdots + \varphi_p z_{t-p}^{(d)} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where ε_t is a white noise.

$\left(z_t^{(d)}\right)_{t\in\mathbb{N}}$ is the trend differenced time series of order d of $(z_t)_{t\in\mathbb{N}}$.

$(\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q)$ and $var(\varepsilon_t)$ are models' parameters.

$(z_{t-1}^{(d)}, z_{t-2}^{(d)}, \dots, z_{t-p}^{(d)})$ are the lagged values of $z_t^{(d)}$.

$(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q})$ are the lagged values of ε_t.

Autoregressive Integrated Moving Average Process of order $(\mathrm{p}, \mathrm{d}, \mathrm{q})$

Using the backshift notation, we obtain:

$$z_t^{(d)} = c + \varphi_1 z_{t-1}^{(d)} + \cdots + \varphi_p z_{t-p}^{(d)} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$\left(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p\right) z_t^{(d)} = c + \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) \varepsilon_t$$

$$\left(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p\right)(1 - B)^d z_t = c + \left(1 + \theta_1 B + \cdots + \theta_q B^q\right) \varepsilon_t$$

$$\varphi(B)(1 - B)^d z_t = c + \theta(B) \varepsilon_t$$

Autoregressive Integrated Moving Average Process of order (p, d, q)

Sometimes, we can take:

$$c = \mu(1 - \varphi_1 - \varphi_2 - \cdots - \varphi_p)$$

So:

$$\left(1 - \varphi_1 B - \cdots - \varphi_p B^p\right) z_t^{(d)} = \mu\left(1 - \varphi_1 - \cdots - \varphi_p\right) + \left(1 + \theta_1 B + \cdots + \theta_q B^q\right)\varepsilon_t$$

$$\left(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p\right)(z_t^{(d)} - \mu) = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right)\varepsilon_t$$

Autoregressive Integrated Moving Average Process of order $(\mathrm{p}, \mathrm{d}, \mathrm{q})$

Random Walk Process $\boldsymbol{ARIMA(0, 1, 0)}$

It is the simplest nonstationary model.

$$\varphi(B)(1-B)^d z_t = c + \theta(B)\varepsilon_t$$

Becomes:

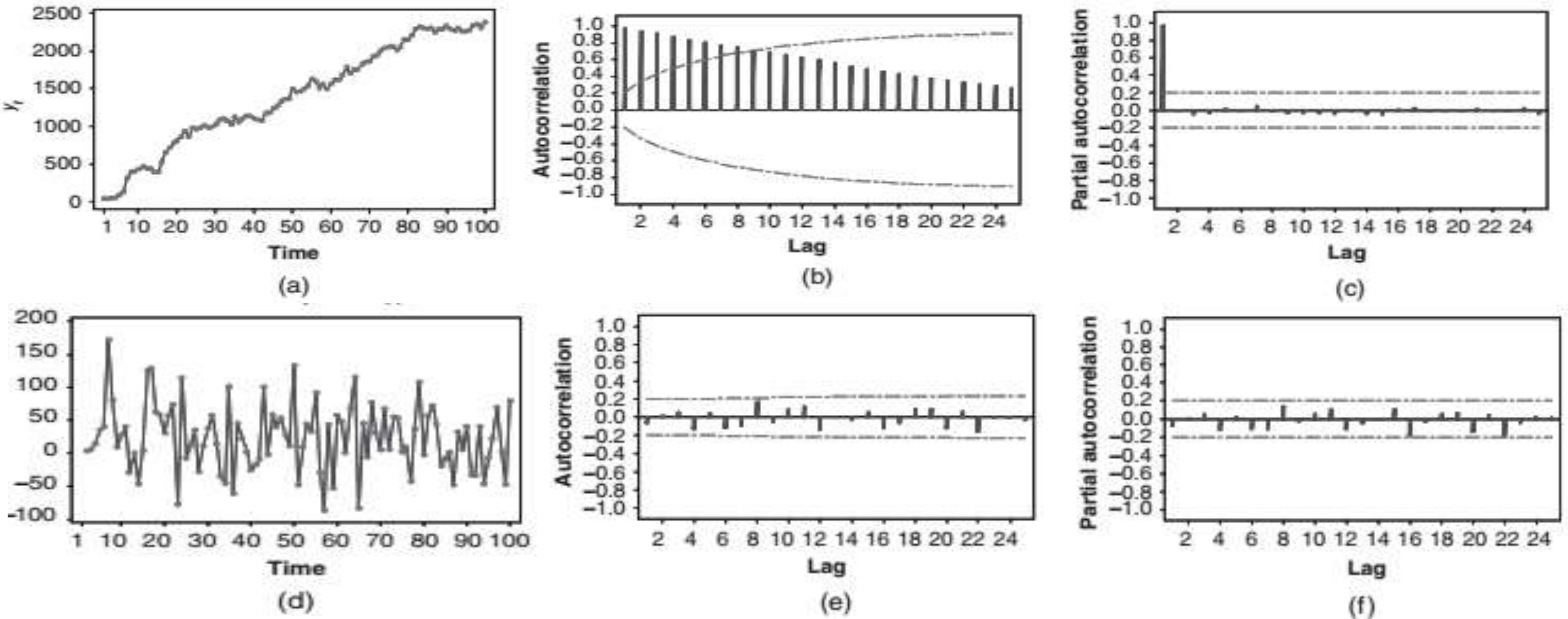
$$(1-B)^1 z_t = c + \varepsilon_t$$

$$z_t - z_{t-1} = c + \varepsilon_t$$

After applying the trend differentiation of order 1, the time series become a white noise

Autoregressive Integrated Moving Average Process of order (p, d, q)

Random Walk Process $ARIMA(0, 1, 0)$



Autoregressive Integrated Moving Average Process of order (p, d, q)

Special cases:

All the stochastic processes that we have already discussed are special cases of the ARIMA model.

Stochastic Process	ARIMA types Process
White noise process	ARIMA(0,0,0)
Random walk process	ARIMA(0,1,0) with no constant
Random walk with drift process	ARIMA(0,1,0) with constant
Autoregressive process	ARIMA(p,0,0)
Moving average process	ARIMA(0,0,q)

Seasonal Processes

Time series data may sometimes exhibit strong periodic patterns.

One way to represent such data is through an additive model where the process is assumed to be composed of two parts:

$$z_t = S_t + N_t$$

Where S_t is the deterministic component of periodicity s and N_t is the stochastic component that can be modeled by an ARMA process.

$$S_t = S_{t+s} = S_{t-s}$$

$$S_t - S_{t-s} = (1 - B^s)S_t = 0$$

Let's apply the operator $(1 - B^s)$ to the $\boldsymbol{z_t}$equation:

$$(1 - B^s)\, z_t = (1 - B^s)\, S_t + (1 - B^s)\, N_t$$

Seasonal Processes

So:

$$z_t^{(s)} = (1 - B^s)\, N_t$$

The process $\left(z_t^{(s)}\right)_t$ can be seen as **seasonally stationary**.

Since an ARMA process can be used to model the stochastic component N_t, in general, we have:

$$\varphi(B)N_t = \theta(B)\varepsilon_t$$

$$\varphi(B)z_t^{(s)} = (1 - B^s)\theta(B)\varepsilon_t$$

where ε_t is a white noise.

Seasonal Processes

Notice that the operator $(1 - B^s)$ makes the data stationary. However, the seasonally differenced data may still show strong autocorrelation at lags: $s, 2s, \ldots, ps$.

So, the seasonal ARMA model is:

$$(1 - \varphi_1^* B^s - -\varphi_2^* B^{2s} - \cdots - \varphi_P^* B^{Ps}) z_t^{(s)} = \left(1 - \theta_1^* B^s - \theta_2^* B^{2s} - \cdots - \theta_Q^* B^{Qs}\right) \varepsilon_t$$

Or :

$$\varphi^*(B^s)(1 - B^s) z_t = \theta^*(B^s) \varepsilon_t$$

Hence, a more general seasonal ARIMA model of orders (p,d,q)(P,D,Q) with period s is:

$$\boldsymbol{\varphi^*(B^s)\varphi(B)(1 - B)^d(1 - B^s)^D z_t = c + \theta^*(B^s)\theta(B)\varepsilon_t}$$

Time Series Model Building

This process includes four steps:

- Model Identification.
- Parameter estimation.
- Model selection.
- Diagnostic checking.
- Forecasting.

Time Series Model Building

Model Identification

This step is based on:

- Plotting the time series.
- Apply differentiation filters.
- Apply the distributional transformations.
- Check for stationarity using statistical tests (Dickey-Fuller, KPSS,…).
- Plotting its SACF and SPACF of the original time series and its differenced/transformed one.
- Identify the order of the model based on the SACF and SPACF.

Time Series Model Building

Model Identification

ρ_k can be estimated using the SACF$(k) = r_k$. And φ_{kk} can be estimated using SPACF$(k) = \hat{\varphi}_{kk}$.

To identify the order, we should define the confidence interval for SACF and SPACF values.

For the SPACF the confidence interval is approximately: (for any k)

$$\pm\frac{2}{\sqrt{n}}$$

Such that n is the size of the time series.

Time Series Model Building

Model Identification

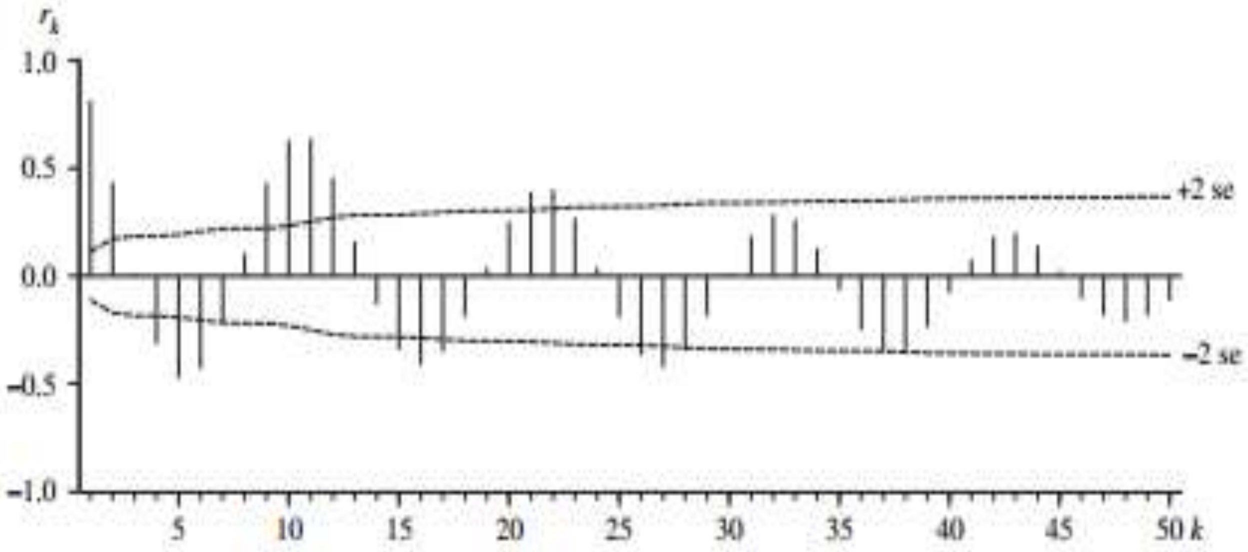
For the SACF the confidence interval is approximately:

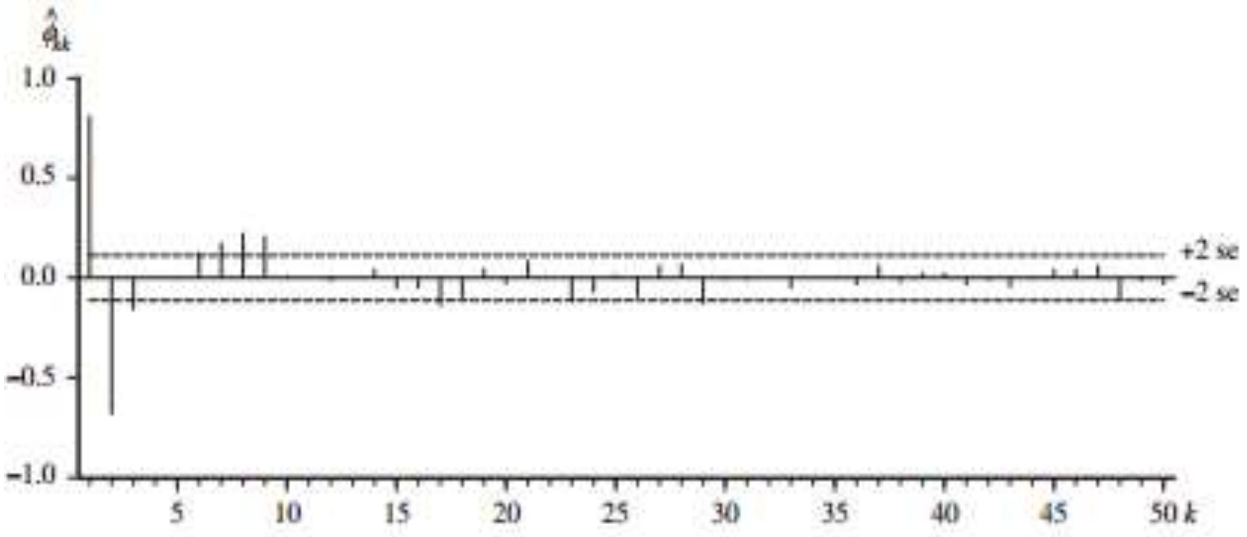
$$\begin{cases} \pm\dfrac{2}{\sqrt{n}} \quad for \quad r_1 \\ \pm 2\sqrt{\dfrac{1+2r_1^2}{n}} \quad for \quad r_2 \\ \vdots \\ \pm 2\sqrt{\dfrac{1+2r_1^2+\cdots+2r_{k-1}^2}{n}} \quad for \quad r_k \end{cases}$$

Time Series Model Building

Model Identification

If the SPACF and SACF are inside the confidence interval, it is said that they are not significant.





Time Series Model Building

Parameter estimation

This step consists on estimating the model unknown parameters:

$$\boldsymbol{\alpha} = \{\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_p, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_q\}$$

There are several methods such as the methods of:

- Moments.
- Maximum likelihood:

$$\widehat{\boldsymbol{\alpha}} \in \underset{\boldsymbol{\alpha}}{\mathbf{argmax}}\{\mathbf{ln}(\boldsymbol{\mathcal{L}}(\boldsymbol{\alpha}, (\boldsymbol{z}_t)_t))\}$$

- Least squares:

$$\widehat{\boldsymbol{\alpha}} \in \underset{\boldsymbol{\alpha}}{\mathbf{argmin}}\left\{\sum_{t=1}^{n} \hat{\varepsilon}_t^2\right\} = \underset{\boldsymbol{\alpha}}{\mathbf{argmin}}\left\{\sum_{t=1}^{n} (\boldsymbol{z}_t - \hat{\boldsymbol{z}}_t)^2\right\}$$

that can be employed to estimate the parameters in the tentatively identified model.

Time Series Model Building

Model selection

In fact, several models can be used to fit the data. Which model to select? The selection of the best model is done based on some criteria, which are called "information criteria".

Definition: information criteria

It is a quantitative value that measures the amount of information lost by a model with respect to other models.

Time Series Model Building

Model selection

There exist different types of these criteria:

(AIC) Akaike Information Criterion: $AIC = 2\mathrm{k} - 2\ln(\mathcal{L}_{max})$

(AICc) the corrected version of AIC: $AICc = 2\mathrm{k} - 2\ln(\mathcal{L}_{max}) + \frac{2k^2+2k}{n-k-1}$

(BIC) Schwarz Bayesian Information Criterion: $BIC = \ln(\mathrm{n})\mathrm{k} - 2\ln(\mathcal{L}_{max})$

Where:

$\mathrm{k} = |\alpha|$: the number of parameters.

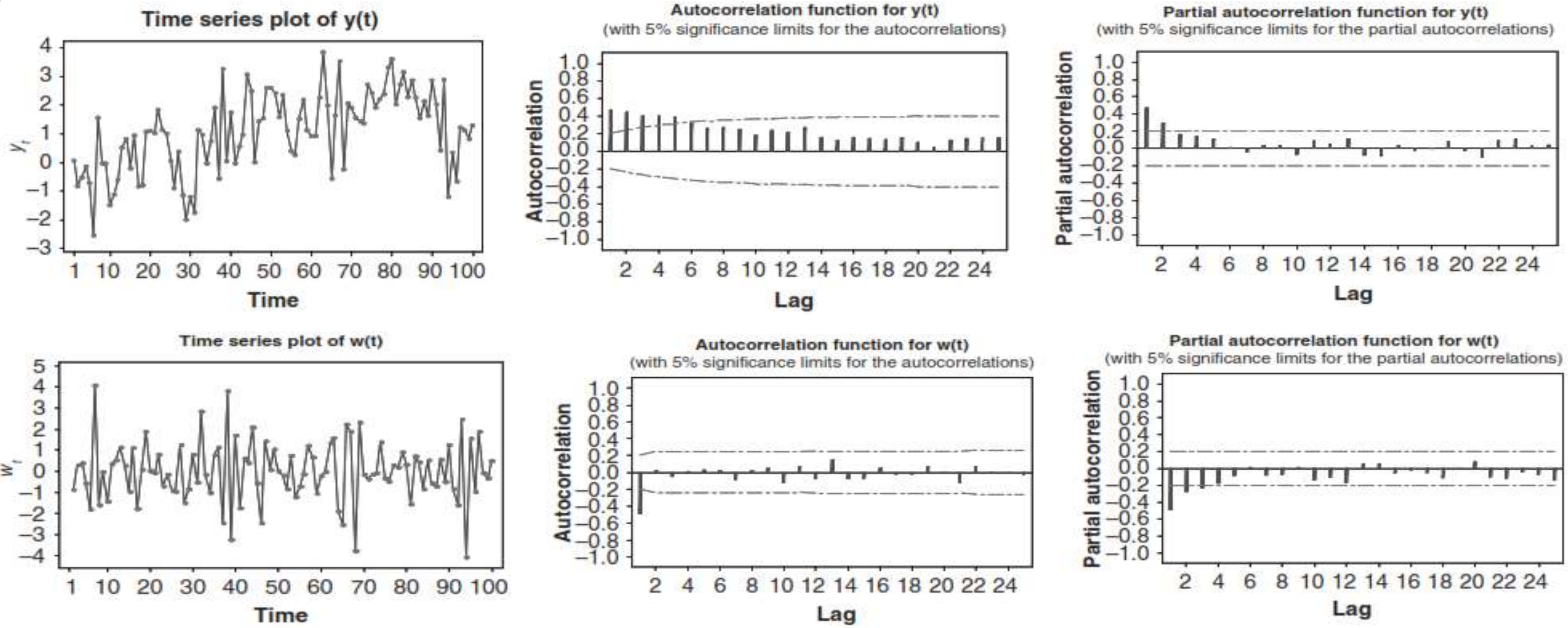
$\mathcal{L}_{\max}$: the maximum of likelihood function.

n: the size of the time series.

Time Series Model Building

Model selection

The objective is to select the model that has smaller values of AIC, AICc and BIC



Time Series Model Building

Diagnostic checking

The third stage, diagnostic checking, is to examine the residuals.

$$\hat{\varepsilon}_t = z_t - \hat{\varphi}_1 z_{t-1} - \cdots - \hat{\varphi}_p z_{t-p} - \hat{\theta}_1 \hat{\varepsilon}_{t-1} - \hat{\theta}_2 \hat{\varepsilon}_{t-2} - \cdots - \hat{\theta}_q \hat{\varepsilon}_{t-q}$$

Before, making any forecast, we should check if the residuals are white noise based on:

- Statistical tests such as Box-Pierce test.
- Using SACF.

Time Series Model Building

Diagnostic checking

Box-Pierce test

It is based on computing the Ljung-Box statistic:

$$Q_{LB} = n(n+2)\sum_{i=1}^{k}\frac{1}{n-i}{r_i}^2$$

Such that :

k= lag.

n is the size of the residuals time series.

r_i is the estimation of the ACF of residuals.

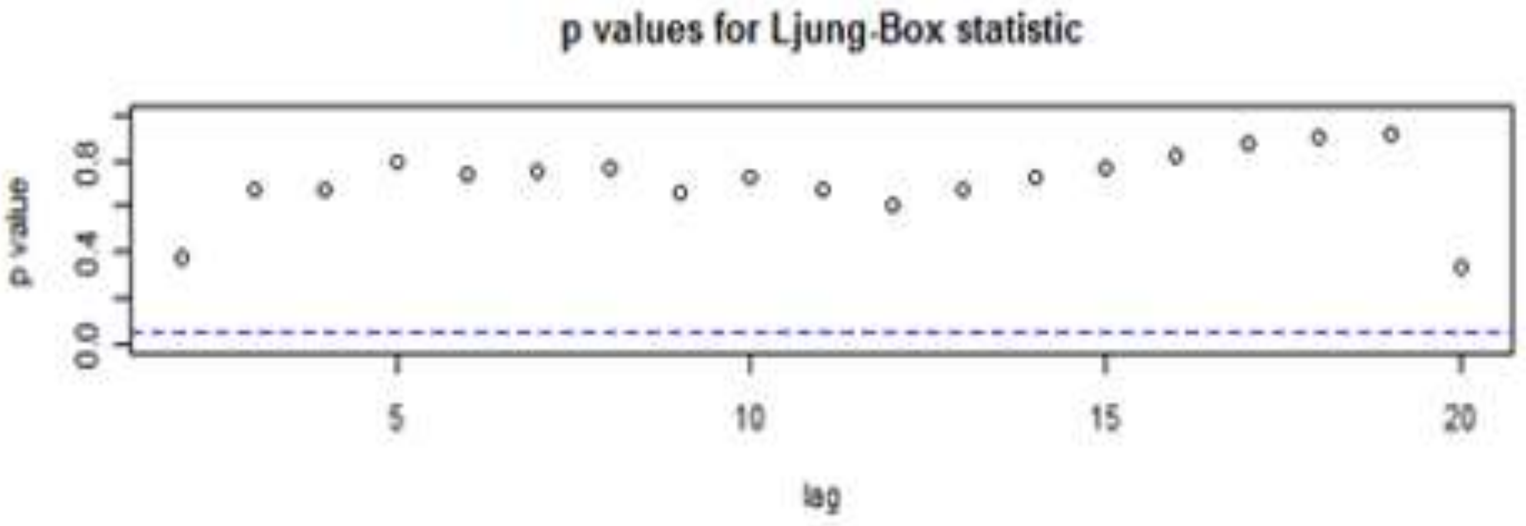
Time Series Model Building

Diagnostic checking

- The null hypothesis of this test is: the residuals are white noise.
- The alternative hypothesis of this test: the residuals are not white noise.

To reject the null hypothesis we should have:

$$p\text{-value} \leq 0.05$$



Time Series Model Building

Forecasting

Once an appropriate time series model has been fit, it may be used to generate forecasts of future observations.

If we denote the current time by n, the forecasts for z_{n+h} is called the h-period ahead forecasts and denoted by $\hat{z}_{n+h}$.

$$\hat{z}_{n+h} = E(z_{n+h}|z_n, z_{n-1}, \dots, z_1) = \mu + \sum_{i=h}^{\infty} \psi_i \varepsilon_{n+h-i}$$

$$\hat{z}_{n+h} = \mu + \psi_h \varepsilon_n + \psi_{h+1} \varepsilon_{n-1} + \psi_{h+2} \varepsilon_{n-2} + \cdots$$

Time Series Model Building

Forecasting

Such that:

$$\varepsilon_n = z_n - \hat{z}_n$$
$$\varepsilon_{n-1} = z_{n-1} - \hat{z}_{n-1}$$
$$\varepsilon_{n-2} = z_{n-2} - \hat{z}_{n-2}$$
$$\vdots$$

Another approach that is often used in practice, is to use the model equation and to substitute t by $t + h$.

For example:

$$z_t = \mu + \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \varepsilon_t$$

We want to compute one-period ahead forecasts:

$$\hat{z}_{n+1} = \mu + \varphi_1 z_n + \varphi_2 z_{n-1}$$