UNIT 16 TIME SERIES MODELS

Structure

16.1 INTRODUCTION

In Unit 15, you have learnt that there are two types of stationary processes: strict stationary and weak stationary processes. You have also learnt how to determine the values of autocovariance and autocorrelation coefficients, and to plot a correlogram for a stationary process. In this unit, we discuss various time series models.

In Sec. 16.2 of this unit, we introduce an important class of linear stationary processes, known as Moving Average (MA) and Autoregressive (AR) processes and describe their key properties. We discuss Autoregressive Moving Average (ARMA) models in Sec. 16.3. We also discuss their properties in the form of autocorrelations and the fitting of suitable models to the given data. We discuss how to deal with models with trend by considering integrated models, called the Autoregressive Integrated Moving Average (ARIMA) models in Sec. 16.4.

Objectives

After studying this unit, you should be able to:

- describe a linear stationary process;
- explain autoregressive and moving average processes;
- fit autoregressive moving average models;
- describe and use the ARIMA models; and
- explore the properties of AR, MA, ARMA and ARIMA models.

16.2 LINEAR STATIONARY PROCESSES

In Unit 15, we have considered discrete time stationary processes and their properties. Note that the sequences of random variables $\{Y_i\}$ are mutually independent and identically distributed. If a discrete stationary process consists of such sequences of i.i.d. variables, it is called a **purely random process**. Sometimes it is called white noise.

Recall that the random variables are normally distributed with mean zero and variance σ^2. Similarly, a purely random process has constant mean and variance, i.e.,

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) = \begin{cases} \sigma_Y^2 & k = 0 \\ 0 & k = 1, 2, 3, ... \end{cases} \quad ...(1)$$

In this section, we consider some particular cases of a linear process. Let Y_t be a stochastic process with mean μ. We can express it as a weighted sum of previous random noises (shocks). Thus, we have

$$Y_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + ... \quad \ldots (2)$$

Here a_t, (t = 0, 1, 2, …) represent white noises with mean zero, variance σ_a^2 and ψ_i (i = 1, 2, …) represent weights. For the linear process to be stationary, the following conditions on weights are required, i.e.,

$$\sum \psi_i^2 < \infty, \quad \sum \left|\psi_i^2\right| < \infty \quad \ldots (3)$$

Then, the autocovariance is given by

$$\text{Cov}(a_t, a_{t+k}) = 0 \qquad \text{for } k \neq 0 \quad \ldots (4)$$

For simplicity, we denote the process by X_t:

$$X_t = Y_t - \mu \quad \ldots (5)$$

Therefore, the process X_t has mean zero and we can write it as:

$$X_t = Y_t - \mu = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + ... \quad \ldots (6)$$

Under the above mentioned conditions on weights ψ_i, the model can also be expressed as

$$X_t = a_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... \quad \ldots (7)$$

Let us now consider two particular cases of the linear stationary processes.

16.2.1 Moving Average (MA) Process

The moving average processes have been often used in econometrics. For example, the economic indicators are affected by many random events such as government decisions, strikes and shortages of raw materials, etc. They have immediate effects as well as effects of lower magnitude in past periods. Such processes have been successfully modelled by moving average processes.

Suppose we write the linear process as

$$X_t = \beta_0 a_t + \beta_1 a_{t-1} + ... + \beta_q a_{t-q} \quad \ldots (8)$$

where β_i, (i = 0, 1, 2, …, q) are constants. This process is known as the **moving average process** of order q and is abbreviated as MA(q) process. The white noises (a_t) are scaled so that $\beta_0 = 1$. The mean and variance of X_t are given by

$$E(X_t) = 0 \qquad \text{and} \qquad V(X_t) = \sigma_a^2 \left(1 + \sum_{i=1}^{q} \beta_i^2\right) \quad ...(9)$$

and autocovariance is given as

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) \quad \ldots (10)$$

$$\begin{aligned} \gamma_k = \gamma_{-k} &= \text{Cov}\left(a_t + \beta_1 a_{t-1} + ... + \beta_q a_{t-q}, a_{t+k} + \beta_1 a_{t+k-1} + ... + \beta_q a_{t+k-q}\right) \\ &= 0 \qquad \text{for } k > q \\ &= \sigma_a^2 \left(\beta_k + \beta_1 \beta_{k+1} + ... + \beta_{q-k} \beta_q\right) \text{ for } k = 1, 2, ..., q \end{aligned} \quad \ldots (11)$$

The autocorrelation function (acf) of the MA(q) process is given by

$$\rho_k = \frac{\left(\beta_k + \beta_1 \beta_{k+1} + \ldots + \beta_{q-k} \beta_q\right)}{\left(1 + \sum_{i=1}^{q} \beta_i^2\right)}, \qquad k = 1, 2, \ldots, q \qquad \ldots(12)$$

Note that the autocorrelation function (acf) becomes zero, if lag k is greater than the order of the process, i.e., q. This is a very important feature of moving average (MA) processes.

First and Second Order Moving Average (MA) Processes

For the first order moving average {MA (1)} process, we have

$$X_t = a_t + \beta_1 a_{t-1} \qquad \ldots (13)$$

The mean and variance are obtained for q =1 as

$$E(X_t) = 0, \qquad V(X_t) = \sigma_a^2 \left(1 + \beta_1^2\right) \qquad \ldots (14)$$

and the autocorrelation coefficient is obtained for q=1 as

$$\rho_1 = \frac{\beta_1}{\left(1 + \beta_1^2\right)} \qquad \ldots (15)$$

Similarly, for the second order **Moving Average MA(2) process** we have

$$X_t = a_t + \beta_1 a_{t-1} + \beta_2 a_{t-2} \qquad \ldots (16)$$

For q = 2, the mean and variance are given as

$$E(X_t) = 0, \qquad V(X_t) = \sigma_a^2 \left(1 + \beta_1^2 + \beta_2^2\right) \qquad \ldots (17)$$

The autocorrelation coefficients are given as

$$\rho_1 = \frac{\left(\beta_1 + \beta_1 \beta_2\right)}{\left(1 + \beta_1^2 + \beta_2^2\right)}, \qquad \rho_2 = \frac{\beta_2}{\left(1 + \beta_1^2 + \beta_2^2\right)} \qquad \ldots (18)$$

There is no requirement on the constants β_1 and β_2 for stationarity. However, for unique representation of the model, the autocorrelation coefficients should satisfy the condition of **invertibility,** which is satisfied when the roots of

$$\theta(B) = 1 + \beta_1 B + \beta_2 B^2 + \ldots + \beta_q B^q = 0 \qquad \ldots (19a)$$

lie outside the unit-circle, i.e., roots $|B| > 1$.

For MA (1) process, we have

$$\theta(B) = 1 + \beta_1 B = 0 \quad \Rightarrow B = -1/\beta_1 \qquad \ldots (19b)$$

Therefore, if $|B| > 1$, this implies that $|\beta_1| < 1$. Hence, for invertibility

$$|\beta_1| < 1 \qquad \ldots (20)$$

Let us consider an example of the moving average process.

Example 1: Consider a time series consisting of 60 consecutive daily over shots from an underground gasoline tank at a filling station. The sample mean and estimate of σ_a^2 with some sample autocorrelations are given as:

Sample mean = 4.0; $\sigma_a^2 = 4515.46$

$r_1 = -0.5$, $r_2 = 0.1252$, $r_3 = -0.2251$, $r_4 = 0.012$, $r_5 = 0.0053$

Check whether a moving average MA (1) process can be fitted to the data and obtain preliminary estimates of the parameters.

Solution: We are given the sample mean of 60 observations as 4.0 and estimate of σ_a^2, i.e., $\hat{\sigma}_a^2 = 3415.72$.

The MA (1) model is written as

$$Y_t = \mu + a_t + \beta_1 a_{t-1}$$

$$X_t = Y_t - \mu = a_t + \beta_1 a_{t-1}$$

If the process is purely random, all the autocorrelations (r_k) should be in the range of

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

In this case, $$\pm \frac{2}{\sqrt{N}} = \pm \frac{2}{\sqrt{60}} = \pm 0.258$$

Here we see that of the given autocorrelations, only r_1 lies outside the range, given by ±0.258. This suggests that moving average MA (1) model could be a suitable model since only ρ_1 is significantly different from zero and ρ_k, k >1 lie within the range ±0.258.

Equating r_1 to ρ_1 given by equation (15) and using the method of moments, we get

$$r_1 = -0.5 \quad \Rightarrow \frac{\beta_1}{\left(1+\beta_1^2\right)} = -0.5$$

On simplifying the above equation, we get

$$\hat{\beta}_1 = -0.1$$

Hence, the model MA(1) becomes $X_t = a_t - 0.1\, a_{t-1}$

Thus, $$Y_t = -4.0 + \left(a_t - 0.1\, a_{t-1}\right)$$

where a_t is white noise with estimated variance of 4515.46.

You may now like to solve the following exercise to check your understanding about MA processes.

E1) Show that the autocorrelation function of MA(2)

$$X_t = a_t + 0.74\, a_{t-1} - 0.19\, a_{t-2}$$

is given by

$$\rho_k = \begin{cases} 0.3675 & k = \pm 1 \\ -0.1289 & k = \pm 2 \\ 0 & \text{Otherwise} \end{cases}$$

In Sec.16.2.1, we have considered estimation of parameters $\beta_1, \beta_2 \ldots$ by the method of moments, i.e., by equating autocorrelations to their expected values. This method is not a very efficient method of estimation of parameters. For moving average processes, usually the maximum likelihood method is used which gives more efficient estimates when N is large. We do not discuss it here as it is beyond the scope of this course.

16.2.2 Autoregressive (AR) Process

A stationary process Y_t is said to be an autoregressive process of order p, abbreviated as AR (p), if

$$Y_t - \mu = \alpha_1(Y_{t-1} - \mu) + \alpha_2(Y_{t-2} - \mu) + \ldots + \alpha_p(Y_{t-p} - \mu) + a_t \quad \ldots (21)$$

which is written as

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + a_t \quad \ldots (22)$$

where $X_t = Y_t - \mu$ and a_t is white noise. It is similar to a multiple regression model, where we regress X_t on its past values and that is why it is called an **autoregressive process**.

A linear stationary process can always be expressed as an autoregressive process of suitable order. Unlike the moving average (MA) process, which puts no restrictions on parameters for stationarity, autoregressive (AR) process requires certain restrictions on the parameters α for stationarity. An autoregressive (AR) process can also be written as

$$\left(1 - \alpha_1 B - \alpha_2 B^2 - \ldots - \alpha_p B^p\right) X_t = a_t \quad \ldots (23)$$

$$\text{or} \quad \phi(\beta)\, X_t = a_t$$

where B is the backward shift operator, defined as

$$B X_t = X_{t-1},\ B^2 X_t = X_{t-2}, \ldots\ldots\ B^p X_t = X_{t-p} \quad \ldots (24)$$

For an AR (p) process to be stationary, the roots of

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \ldots - \alpha_p B^p = 0 \quad \ldots (25)$$

must lie outside the unit circle.

First Order Autoregressive {AR(1)} (Markov) Process

Suppose, we write the linear model as

$$X_t = \alpha_1 X_{t-1} + a_t \quad \ldots (26)$$

By repeatedly using this equation you can see that X_t can be expressed as weighted sum of infinite numbers of past noises a_t, i.e.,

$$X_t = a_t + \alpha_1 a_{t-1} + \alpha_2 a_{t-2} + \ldots \quad \ldots (27)$$

Autocorrelations ρ_k are obtained by multiplying equation (27) by X_{t-k} and taking expectations of the results. Then we get

$$\rho_k = \alpha_1 \rho^{k-1} = \ldots\ldots\ldots = \alpha_1^k \quad \ldots (28)$$

Thus, $\rho_0 = 1, \quad \rho_1 = \alpha_1$

From equation (27) $\quad \sigma_x^2 = \alpha_1 \sigma_x^2 + \sigma_a^2 \quad \ldots (29)$

which gives $\quad \sigma_x^2 = \sigma_a^2 / \left(1 - \alpha_1^2\right) \quad \ldots (30)$

and σ_x^2 is positive if $|\alpha_1| < 1$. Thus, for stationarity

$$|\alpha_1| < 1 \quad \ldots (31)$$

When α_1 is positive and large, the time plot becomes smooth and shows a slow changing trend. When α_1 is large and negative, the time plot shows a very rapid zig-zag movement. It is because of negative autocorrelations. If one value of autocorrelation is above mean, the next value of the autocorrelation is very likely to be below mean, and so on.

Second order Autoregressive {AR (2)} process

This process is obtained by taking p = 2 and the model is

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + a_t \qquad \ldots (32)$$

For stationarity, the following restrictions are placed on the coefficients:

$$\alpha_2 + \alpha_1 < 1; \quad \alpha_2 - \alpha_1 < 1 \text{ and } -1 < \alpha_2 < 1 \qquad \ldots (33)$$

For autoregressive AR(2) model, the first two autocorrelations ρ_1 and ρ_2 are obtained as follows:

On multiplying equation (32) by X_{t-1} and X_{t-2} and taking expectations and dividing the results by σ_x^2, we get

$$\rho_1 = \alpha_1 + \alpha_2 \rho_1 \qquad \ldots (34a)$$

$$\rho_2 = \alpha_1 \rho_1 + \alpha_2 \qquad \ldots (34b)$$

On simplifying the above equations, we obtain

$$\alpha_1 = \frac{\rho_1(1-\rho_2)}{(1-\rho_1^2)} \qquad \alpha_2 = \frac{(\rho_2 - \rho_1^2)}{(1-\rho_1^2)} \qquad \ldots (35)$$

Similarly, ρ_1 and ρ_2 can be expressed in terms of α_1 and α_2 as

$$\rho_1 = \frac{\alpha_1}{(1-\alpha_2)}, \qquad \rho_2 = \alpha_2 + \frac{\alpha_1^2}{(1-\alpha_2)} \qquad \ldots (36a)$$

$$\sigma_x^2 = \sigma_a^2 / (1 - \rho_1\alpha_1 - \rho_2\alpha_2) \qquad \ldots (36b)$$

Multiplying equation (32) by X_{t-k}, taking expectations and dividing by σ_x^2 gives the autocorrelation function of AR (2) process as

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}, \qquad X > 0 \qquad \ldots (37)$$

We can obtain ρ_k for different values of k by using equation (37) for k=1, 2, ….

Let us consider an example of AR(2) process.

Example 3: Consider an auto regressive AR(2) model

$$X_t = 0.80 X_{t-1} - 0.60 X_{t-2} + a_t$$

Verify whether the series is stationary.

(i) Obtain ρ_k for k = 1, 2, …, 5, and (ii) plot the correlogram.

Solution: We have an autoregressive AR (2) model

$$X_t = 0.80 X_{t-1} - 0.60 X_{t-2} + a_t \qquad \ldots (i)$$

Now from equation (36a), the autocorrelations ρ_1 and ρ_2 are given as

$$\rho_1 = \frac{\alpha_1}{(1-\alpha_2)}, \qquad \rho_2 = \alpha_2 + \frac{\alpha_1^2}{(1-\alpha_2)}$$

$$\therefore \rho_0 = 1, \ \rho_1 = \frac{0.80}{(1+0.60)} = 0.50 \quad \text{and} \quad \rho_2 = -0.50 + \frac{(0.80)^2}{(1+0.60)} = -0.20$$

We obtain the values of the autocorrelations ρ_3, ρ_4 and ρ_5 using equation (37) and get

$$\rho_3 = \alpha_1\rho_2 + \alpha_2\rho_1,$$

$$\rho_3 = 0.80 \times -0.20 + (-0.60) \times (0.50) = -0.46$$

$$\rho_4 = 0.80 \times -0.46 + (-0.60) \times (-0.20) = -0.25$$

$$\rho_5 = 0.80 \times -0.25 + (-0.60) \times (-0.46) = 0.076$$

The correlogram of the given AR process is shown in Fig. 16.1.

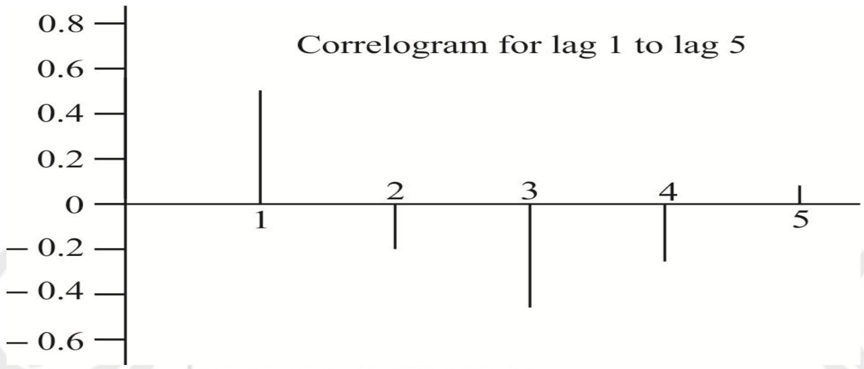

Fig 16.1: Correlogram of the model.

You may now like to solve the following exercises to check your understanding about MA processes.

E2) Consider an AR (2) process given by

$$X_t = X_{t-1} - 0.5\, X_{t-2} + a_t$$

Verify whether the series is stationary or not.

a) Obtain ρ_k for k =1, 2, ..., 5 and b) plot the correlogram

E3) For each of the following processes, write the model using B notations and then determine whether the processes are stationary or not:

a) $X_t = 0.3X_{t-1} + a_t$

b) $X_t = \left(X_{t-1} + X_{t-2}\right)/12 + a_t$

16.2.3 Fitting an Autoregressive Process

Suppose N observations on a time series $y_1, y_2, \ldots, y_N$ are available. We now wish to fit an autoregressive (AR) process of suitable order. Therefore, we need to know the order of autoregressive (AR) process, that is, p. Suppose, we know the order p. Then we have to estimate parameters $\mu, \alpha_1, \alpha_2, \ldots, \alpha_p, \sigma^2_x$, etc. We calculate the autocorrelations from the data. Usually μ is estimated by $\hat{Y}$. Hence, by subtracting $\hat{Y}$ from Y_t, we calculate

$$X_t = Y_t - \hat{Y} \qquad \ldots (38a)$$

For the given r_k, we have to calculate parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$ of the model:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + a_t \qquad \ldots (38b)$$

For an autoregressive (AR) process, the least squares estimates of the parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$ are obtained by minimising S:

$$S = \sum^{N}\left(X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \ldots - \alpha_p X_{t-p}\right)^2 \qquad \ldots (39)$$

with respect to $\alpha_1, ..., \alpha_p$, and equating the result to zero. This method provides good estimates.

If Y_t, t = 1, 2, …, N is the observed series $X_t = Y_t - \hat{Y}$ are used in equation (39). This looks very similar to multiple regression estimates and by differentiating S with respect to $\alpha_1, \alpha_2, \ldots, \alpha_p$ and equating the result to zero, we get a set of k equations

$$R\hat{\alpha} = r \qquad \ldots (40)$$

where R is a matrix of autocorrelations given by

$$R = \begin{bmatrix} 1 & r_1 \ldots\ldots\ldots\ldots r_{p-1} \\ r_1 & 1 \ldots\ldots\ldots\ldots r_{p-2} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ r_{p-1} & r_{p-2} \ldots\ldots\ldots\ldots 1 \end{bmatrix} \qquad \ldots (41)$$

and $r' = (r_1, r_2, ..., r_p)$ is the row matrix corresponding to the column matrix r. Thus, $\hat{\alpha}$ is obtained by solving the simultaneous equations (40) using inverse of R matrix denoted by R^{-1} as

$$\hat{\alpha} = R^{-1} r \qquad \ldots (42)$$

16.2.4 Determining the Order of an Autoregressive Model

For fitting the model, we have to estimate the order of the autoregressive model for the data at hand. For the first order autoregressive model, the autocorrelation function (acf) reduces exponentially as follows:

$$\rho_k = \alpha_1^k \qquad \text{as } |\alpha_1| < 1.$$

Hence, for an autoregressive process AR (1), the exponential reduction of autocorrelation function (acf) gives a good indication that the autoregressive process is of order 1. However, this is not true for correlogram of higher orders. For two and higher order autoregressive models, the autocorrelation function (acf) can be a combination of damped exponential or cyclical functions and may be difficult to identify.

One way is to start fitting the model by taking p = 1 and then p = 2, and so on. As soon as the contribution of the last α_p fitted is not significant, which can be judged from the reduction in the value of residual sum of squares, we should stop and take the order as p −1. An alternative method is to calculate what is called **partial autocorrelation function.**

16.2.5 Partial Autocorrelation Function (pacf)

For an autoregressive AR (p) process, the partial autocorrelation function (pacf) is defined as the value of the last coefficient α_p. We start with p =1 and calculate pacf. Hence, for the AR (1) process, pacf (1) is

$$\alpha_1 = \rho_1 \qquad \ldots (43a)$$

For AR (2), the pacf is given by

$$\alpha_2 = \frac{\left(\rho_2 - \rho_1^2\right)}{\left(1 - \rho_1^2\right)} \qquad \ldots (43b)$$

as described earlier. In this way, we can go on calculating pacf(3) as α_3 and α_p, p = 4, 5, …. We can estimate these partial autocorrelation functions by substituting estimated autocorrelations r_k in place of ρ and then test the significance. When partial autocorrelation function (pacf) is zero, its asymptotic standard error is $1/\sqrt{N}$. Hence, we calculate partial autocorrelation functions (pacf) by increasing the order by one every time. As soon as this lies within range of $\pm 2/\sqrt{N}$, we stop and take the order as the last significant partial autocorrelation function (pacf). This is indicated when pacf lies outside the range of $\pm 2/\sqrt{N}$. In the following steps, we give partial autocorrelation functions (pacf) up to autoregressive AR(3) process:

$$\text{pacf}(1) = \rho_1 = \alpha_1;\ \text{pacf}(2) = \frac{\left(r_2 - r_1^2\right)}{\left(1 - r_1^2\right)} = a_2 \qquad \ldots (44a)$$

and
$$\text{pacf}(3) = \frac{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & r_3 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{vmatrix}} \qquad \ldots (44b)$$

where |…| means the determinant of the matrix.

Let us now calculate partial autocorrelation functions for stationary processes.

Example 4: Find the pacf of the AR(2) process:

$$X_t = 0.333 X_{t-1} + 0.222 X_{t-2} + a_t$$

Solution: For this process, $\alpha_1 = 0.333$ and $\alpha_2 = 0.222$. We use the expressions of ρ_1 and ρ_2 as given in equation (36a) and get

$$\rho_1 = \frac{\alpha_1}{(1-\alpha_2)} = \frac{0.333}{0.778} = 0.428$$

$$\rho_2 = \alpha_2 + \frac{\alpha_1^2}{(1-\alpha_2)} = 0.222 + \frac{0.111}{0.778} = 0.365$$

Now, from equations (43a and b),

$$\text{pacf}(1) = \alpha_1 = \rho_1 = 0.428$$

and
$$\text{pacf}(2) = \frac{r_2 - r_1^2}{1 - r_1^2} = \frac{(0.365 - 0.183)}{(1 - 0.183)} = 0.222$$

Also, $\text{pacf}(k) = 0$, for $k \geq 3$.

Example 5: Suppose for a time series of length N = 100, the three autocorrelation coefficients are $r_1 = 0.806$, $r_2 = 0.428$, $r_3 = 0.070$. Calculate the pacfs and estimate the order of autoregressive model to be fitted.

Solution: Equating r_k to ρ_k (k= 1, 2, 3), from equations (44a and b), we have

$$\text{pacf}(1) = r_1 = 0.806$$

$$\text{pacf}\ (2) = \frac{\left(r_2 - r_1^2\right)}{\left(1 - r_1^2\right)} = \frac{\left(0.428 - 0.650\right)}{0.350} = -\ 0.634$$

$$\text{pacf}\ (3) = \frac{\begin{vmatrix} 1 & 0.806 & 0.428 \\ 0.806 & 1 & 0.806 \\ 0.428 & 0.806 & 0.070 \end{vmatrix}}{\begin{vmatrix} 1 & 0.806 & 0.428 \\ 0.806 & 1 & 0.806 \\ 0.428 & 0.806 & 1 \end{vmatrix}} = 0.077$$

and $\quad$ range = $\pm 2/\sqrt{N} = \pm 2/10 = \pm 0.2$

The partial autocorrelation functions pacf (1) and pacf (2) lie outside this range and pacf (3) lies inside this range. Since the least significant pacf is pacf (2), the order of the model is 2 and the autoregressive model AR(2) is suggested for this process.

You may now like to solve the following exercises to check your understanding about MA processes.

E4) For the AR (2) process

$$X_t = 1.0\ X_{t-1} - 0.5 X_{t-2}\ + a_t$$

calculate ρ_1 and ρ_2. State whether the model is stationary. Also calculate pacf (1) and pacf (2).

E5) For the model

$$X_t = 1.5\ X_{t-1} - \ 0.6\ X_{t-2}\ + a_t$$

obtain ρ_1 and ρ_2. Is the process stationary?

E6) Find the autocorrelation function (acf) of the process

$$X_t = \ X_{t-1} - 0.25 X_{t-2}\ + a_t$$

and obtain ρ_1 and ρ_2.

E7) The following table gives the number of workers trained during 1980-2010.

(t)	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
y_t	4737	5117	5091	3468	4320	3825	3673	3694	3708	3333
(t)	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
y_t	3367	3614	3362	3655	3963	4405	4595	5045	5700	5716
(t)	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
y_t	5138	5010	5353	6074	5031	5648	5506	4230	4827	3885

Some autocorrelations are given below:

$r_1 = 0.732$, $r_2 = 0.661$, $r_3 = 0.557$, $r_4 = 0.385$, $r_5 = 0.272$, $r_6 = 0.119$, $r_7 = 0.019$, $r_8 = -0.139$, $r_9 = -0.268$, $r_{10} = -0.375$, $\bar{y} = 4503.00$ and $\sigma_y = 836.74$

i) Draw the time plot.
ii) Plot the correlogram.
iii) Calculate pacf (1) and pacf (2) and test their significance.
iv) Which one of the models, AR(1) or AR(2), will be more suitable for this data?
v) Fit the suitable model.

16.3 AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

A finite order moving average process can be written as an infinite order autoregressive process. Similarly, a finite order autoregressive process can be written as an infinite order moving average process. We would like to fit a model, which has the least number of parameters. This property is called **parsimony** (most economical). Hence, a combination of autoregressive (AR) and moving average (MA) models may turn out to be the most parsimonious. We represent a combination of AR(p) and MA(q) model as ARMA(p, q) and write

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} + a_t + \beta_1 a_{t-1} + \beta_2 a_{t-2} + ... + \beta_q a_{t-q} \quad \text{… (45)}$$

Using the backward shift operator B, we can write equation (45) as

$$\Phi(B)\, X_t = \theta(B)\, a_t \quad \text{… (46)}$$

where

$$\Phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - ..... - \alpha_p B^p \quad \text{(AR)} \quad \text{… (47a)}$$

$$\theta(B) = 1 + \beta_1 B + \beta_2 B^2 + ... + \beta_q B^q \quad \text{(MA)} \quad \text{… (47b)}$$

The conditions of stationarity and invertibility are the same as for autoregressive (AR) and moving average (MA) processes, respectively, i.e., the roots of

$$\Phi(B) = 0 \text{ and } \theta(B) = 0 \quad \text{… (47c)}$$

must lie outside the unit circle. So the modulus of roots of B must be greater than one.

An ARMA (1, 1) model can be written as

$$X_t = \alpha X_{t-1} + a_t + \beta\, a_{t-1} \quad \text{… (48a)}$$

which can be written using backward operator B as

$$(1 - \alpha B)\, X_t = (1 + \beta B)\, a_t \quad \text{… (48b)}$$

For a stationary and invertible ARMA (1, 1) process,

$$|\alpha| < 1, \qquad |\beta| < 1$$

On multiplying equation (48a) by X_t, X_{t-1} and X_{t-k} and taking expectations, we obtain

$$\gamma_0 = \frac{\sigma_a^2 \left(1 + \beta^2 + 2\alpha\beta\right)}{\left(1 - \alpha^2\right)} \quad \text{… (49a)}$$

$$\gamma_1 = \alpha\gamma_0 + \beta\sigma_a^2 \quad \text{… (49b)}$$

$$\gamma_k = \alpha\gamma_{k-1},\ k \geq 2 \quad \text{… (49c)}$$

We also obtain

$$\rho_1 = \frac{(1 + \alpha\beta)(\alpha + \beta)}{\left(1 + \beta^2 + 2\alpha\beta\right)} \quad \text{… (49d)}$$

$$\rho_k = \alpha\rho_{k-1},\ k \geq 2 \quad \text{… (49e)}$$

Thus, the autocorrelation function decays exponentially from the starting value ρ_1, which depends on α and β.

Let us take up an example of the ARMA model.

Example 6: Write the following ARMA (1,1) model

$$X_t = 0.5\, X_{t-1} + a_t - 0.3\, a_{t-1}$$

using backward operator B. Is the process stationary and invertible? Calculate ρ_1 and ρ_2 for the process.

Solution: Since $\alpha = 0.5$ and $\beta = -0.3$, from equation (48b), the model is written using backward operator B as:

$$(1-0.5B)X_t = (1-0.3B)a_t$$

In this case, from equations (47a and b) we have

$$\Phi(B) = 1 - 0.5\,B \text{ and } \theta(B) = 1 - 0.3B$$

Therefore, for stationarity and invertibility, from equation (47c), the roots of $1 - 0.5\,B = 0$ and $1 - 0.3B = 0$ must lie outside the unit circle. The roots of these equations are:

$$B = 1/0.5 = 2.0 \quad \text{and} \quad B = 1/0.3 = 3.33$$

Since both roots lie outside the unit circle, the process is stationary and invertible. From equations (49 d and e),

$$\rho_1 = (1+\alpha\,B)(\alpha+\beta)/(1+\beta^2+2\alpha\beta) = 0.215, \text{ and}$$

$$\rho_2 = \alpha\rho_1 = 0.107$$

You may now like to try out an exercise.

E8) Show that the ARMA (1, 1) model

$$X_t - 0.5\, X_{t-1} = a_t - 0.5\, a_{t-1}$$

can be equivalently written as $X_t = a_t$, which is a white noise model.

16.4 AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODELS

In Units 13 and 14, we have discussed that the actual time series often contains trend and seasonal components. In that sense, most of the time series we come across are non-stationary as their mean changes with time. In these units, we have tried to take moving average to remove seasonal component and then we have estimated trend. In this section, we incorporate trend and seasonal effects in the model and then by making suitable operations on the series, transform them to stationary series. Then we apply the methods of stationary models discussed so far.

If a time series is non-stationary because of changes in mean, we can take the difference of successive observations. The modified series is more likely to be stationary. Sometimes more than one difference of successive observations is required to get a modified stationary model. Such a model is called an **integrated** model because the stationary model that is fitted to the modified series has to be summed or integrated to provide a model for the original non-stationary series. The first difference of series X_t is defined as W_t:

$$W_t = \nabla X_t = (1-B)\, X_t = X_t - X_{t-1} \qquad \ldots (50)$$

where ∇ is the **difference operator**. This is called the difference of order 1. We may define a modified series of order d as

$$W_t = \nabla^d X_t = (1-B)^d X_t \qquad \dots (51)$$

where d takes values 1, 2, ….

For d =2, this operation takes differences twice:

$$W_t = \nabla^2 X_t = (1-B)(1-B) X_t = (1-B)(X_t - X_{t-1})$$
$$= X_t - X_{t-1} - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} \qquad \dots (52)$$

In general, the ARIMA model can be written as:

$$W_t = \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_p W_{t-p} + a_t + \beta_1 a_{t-1} + \beta_2 a_{t-2} + \dots + \beta_q a_{t-q} \qquad \dots (53a)$$

or using backward operator B, it can be written as:

$$\Phi(B) W_t = \theta(B) a_t \qquad \dots (53b)$$

or $$\Phi(B)(1-B)^d X_t = \theta(B) a_t \qquad \dots (53c)$$

It is denoted by ARIMA (p, d, q). The operator Φ (B) $(1-B)^d$ has d roots of B equal to 1. For d = 0, the series is an ARMA process. In practice, the first or second difference make the process stationary. A random walk model is an example of the ARIMA model.

Consider the time series

$$X_t = X_{t-1} + a_t \qquad \dots (54a)$$

which can be written as

$$(1-B) X_t = a_t \qquad \dots (54b)$$

It is clearly non-stationary as one root of

$$\Phi(B) = 1 - B = 0 \qquad \dots (54c)$$

lies on the unit circle. To make it stationary, we take one difference of X_t, as

$$W_t = X_t - X_{t-1} = a_t$$

So the time series can be written as ARIMA (0,1,0). W_t is a white noise process and stationary.

A plot of the first difference looks like a plot of a stationary process without any trend. The plot of autocorrelations and partial autocorrelations provide the idea of the process.

Example 7: For the model

$$(1-0.2B)(1-B)X_t = (1-0.5B)a_t$$

find p, d, q and express it as ARIMA (p, d, q). Determine whether the process is stationary and invertible.

Solution: We are given the model

$$(1-0.2B)(1-B)X_t = (1-0.5B)a_t$$

a) In this case, from equations (53 b and c), we can write the given model as

$$(1-0.2B)(1-B)^1 X_t = (1-0.5B)a_t$$

which implies that $W_t = (1 - B) X_t$, i.e., d = 1 and from equation (53a)

$$X_t - 0.2X_{t-1} = a_t - 0.5a_{t-1}$$

This implies that p = 1 and q = 1. Hence, the process is ARIMA (1,1,1)

b) $\Phi(B) = (1 - B)(1 - 0.2B) = 0 \Rightarrow B = 1$ and $B = 5$ and

$$\theta(B) = (1 - 0.5B) = 0 \Rightarrow B = 1/0.5 = 2.0$$

One of the roots of $\Phi(B) = (1 - B)(1 - 0.2B) = 0$ is 1. Hence, the process is non-stationary. However, the root of $\theta(B) = 0$ lies outside the unit circle. Hence, it is invertible. For the first difference $W_t = (1 - B) X_t$, the process is stationary and invertible.

You may now lke to try some more exercises for practice.

E9) Consider the time series

$$X_t = \beta_1 + \beta_2 t + a_t$$

where β_1 and β_2 are known constants and a_t is a white noise with variance σ^2.

Determine whether X_t is stationary. If X_t is not stationary, find a transformation that produces a stationary process.

E10) Suppose that the correlogram of a time series consisting of 100 observations has

r_1=0.31, r_2=0.37, r_3= – 0.05, r_4= 0.06, r_5= – 0.21, r_6=0.11, r_7=0.08, r_8 = 0.05, r_9=0.12, r_{10}= – 0.01

Suggest an ARIMA model which may be appropriate for this case.

Let us now summarise the concepts that we have discussed in this unit.

16.5 SUMMARY

1. The sequences of random variables $\{Y_i\}$ are mutually independent and identically distributed. If a discrete stationary process consists of such sequences of i.i.d. variables, it is called a **purely random process**. Sometimes it is called white noise.
2. The moving average processes are used successfully to model stationary time series in econometrics. The MA(q) process of order q is given as

$$X_t = \beta_0 a_t + \beta_1 a_{t-1} + ... + \beta_q a_{t-q}$$

where β_i, (i = 0, 1, 2, …, q) are constant.
3. The autocorrelation function (acf) of the MA (q) process is given by

$$\rho_k = \frac{(\beta_k + \beta_1 \beta_{k+1} + ... + \beta_{q-k} \beta_q)}{\left(1 + \sum_{i=1}^{q} \beta_i^2\right)}, \qquad k = 1, 2, ..., q$$

It becomes zero if lag k is greater than the order of the process, i.e., q. This is a very important feature of moving average (MA) processes.
4. A linear stationary process can always be expressed as an **autoregressive process** of suitable order. Unlike moving average (MA) process, which puts no restrictions on parameters for stationarity, autoregressive (AR) process requires certain restrictions on the parameter α for stationarity.

5. We can estimate the partial autocorrelation functions by substituting estimated autocorrelations r_k in place of ρ. Then we test the significance. When the partial autocorrelation function (pacf) is zero, its asymptotic standard error is $1/\sqrt{N}$. Hence, we calculate the partial autocorrelation functions (pacf) by increasing the order one at a time. As soon as the order lies within range of $\pm 2/\sqrt{N}$, we stop and take the order as the last significant partial autocorrelation function (pacf).

6. A finite order moving average process can be written as an infinite order autoregressive process. Similarly, a finite order autoregressive process can be written as an infinite order moving average process.

7. If a time series is non-stationary because of changes in mean, we can take the difference of successive observations. The modified series is more likely to be stationary. Sometimes more than one difference is required. Such a modified model is called an **integrated** model because the stationary model that is fitted to the modified series has to be summed or integrated to provide a model for the original non-stationary time series.

16.6 SOLUTIONS/ANSWERS

E1) From equation (8), we are given that

$$q=2,\ \beta_0=1,\ \beta_1 = 0.60 \text{ and } \beta_2 = -0.3$$

for the model.

Using equation (12), for q=2, we get

$$\rho_1 = \frac{(\beta_1 + \beta_1\beta_2)}{(1+\beta_1^2+\beta_2^2)} = \frac{0.6+(0.6\times -0.3)}{1+(0.6)^2+(-0.3)^2} = 0.29$$

$$\rho_2 = \frac{2\beta_2}{(1+\beta_1^2+\beta_2^2)} = \frac{2\times -0.3}{1+(0.6)^2+(-0.3)^2} = -0.4138$$

$$\rho_k = 0, \qquad k \geq 3$$

E2) For the series, $\alpha_1 = 1$, $\alpha_2 = -0.5$.

a) From equation (33), the stationarity conditions are

$$\alpha_2 + \alpha_1 < 1, \alpha_2 - \alpha_1 < 1, -1 < \alpha_2 < 1$$

$$\Rightarrow \alpha_2 + \alpha_1 = 0.5, \alpha_2 - \alpha_1 = -1.5, -1 < \alpha_2 < 1$$

All three conditions of stationarity are satisfied in this case. Hence, the process is stationary. From equation (36a), we get

$$\rho_2 = \frac{\alpha_1}{(1-\alpha_2)} = \frac{1}{1.5} = 0.667,\ \rho_2 = \alpha_2 + \frac{\alpha_1^2}{(1-\alpha_2)} = 0.167$$

We calculate the other values of ρ_k from equation (37) with $\alpha_1 = 1$, $\alpha_2 = -0.5$. Thus,

$$\rho_k = \rho_{k-1} - 0.5\,\rho_{k-2}$$

$$\Rightarrow \rho_3 = \rho_2 - 0.5\,\rho_1 = -0.166,$$

$$\rho_4 = \rho_3 - 0.5\,\rho_2 = -0.250, \text{ and}$$

$$\rho_5 = \rho_4 - 0.5\,\rho_3 = -0.166$$

b) The correlogram for the process is shown in Fig. 16.2.

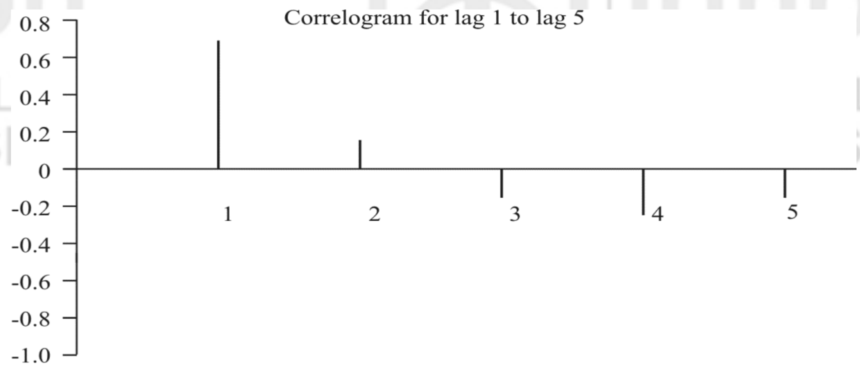

Fig. 16.2: Correlogram for AR(2) Model.

E3) a) Here $\alpha_1 = 0.3$. Therefore, using equation (23), we can write the model $X_t = 0.3X_{t-1} + a_t$ in B notation as

$$(1-0.3B)X_t = a_t \Rightarrow \phi(B) = (1-0.3B)$$

From $\phi(B)=0$, we get $1-0.3B=0$ or $B=1/0.3=3.333$

which lies outside the unit circle. Hence, the process is stationary.

c) Here $\alpha_1 = \alpha_2 = 1/12$. Using equation (23), we can write the model $X_t = (X_{t-1}+X_{t-2})/12 + a_t$ in B notation as

$$\{1-(B+B^2)/12\}X_t = a_t \Rightarrow \phi(B) = \frac{1-B-B^2}{12}$$

The two roots of $\phi(B) = 1-(B+B^2)/12 = 0$ are given by $B = 3, -4$,

and for both, $|B| > 1$. Hence, the process is stationary.

E4) For the AR (2) process

$$X_t = 1.0X_{t-1} + 0.5X_{t-2} + a_t,$$

we have

$$\alpha_1 = 1.0,\ \alpha_2 = -0.5$$

Therefore, from equation (36a),

$$\rho_1 = \alpha_1/(1-\alpha_2) = 1/1.5 = 0.667$$

and $\rho_2 = \alpha_2 + \alpha_1^2/(1-\alpha_2) = -0.5 + 1/1.5 = 0.167$

Since $\alpha_2 + \alpha_1 = 0.5 < 1$, using equation (33), we can say that the process is stationary. From equations (43a and 44a), we have

pacf (1) = $\alpha_1 = 1.0$, pacf (2) = $\alpha_2 = -0.5$

E5) For the model $X_t = 1.5X_{t-1} - 0.6X_{t-2} + a_t$

we have $\alpha_1 = 1.5$ and $\alpha_2 = -0.6$.

Multiplying the model by X_{t-1} and X_{t-2}, taking expectation and dividing by σ_a^2, we get

$$\rho_1 = 1.5 - 0.6\rho_1 \quad \text{and} \quad \rho_2 = 1.5\rho_1 - 0.6$$

Solving the above equations for ρ_1 and ρ_2, we get

$$\rho_1 = 1.5/1.6 = 0.937 \text{ and } \rho_2 = 0.805$$

E6) We are given the model $X_t = X_{t-1} - 0.25X_{t-2} + a_t$

and we have $\alpha_1 = 1.0$ and $\alpha_2 = -0.25$.

Multiplying by X_{t-k}, taking expectations and dividing by σ_a^2, we get the autocorrelation function as:

$$\rho_k = \rho_{k-1} - 0.25\rho_{k-2}$$

On putting the values α_1=1, $\alpha_2 = -0.25$ in equation (36a), we obtain
$\rho_1 = \alpha_1/(1-\alpha_2) = 1/1.25 = 0.8, \quad \rho_2 = \alpha_2 + \alpha_1^2/(1-\alpha_2) = 0.55$

E7) We first draw a time plot for the given time series:

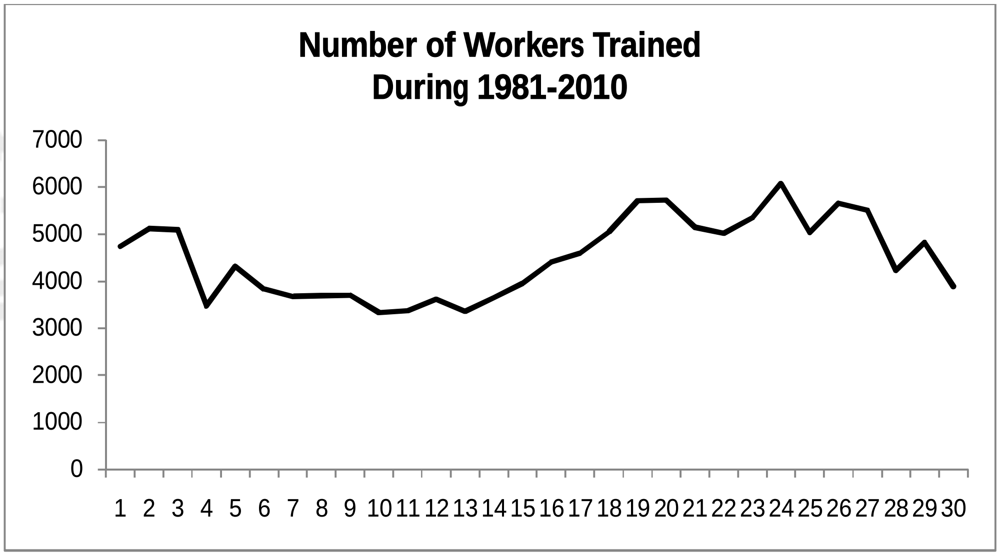

Fig. 16.3: Time plot of the series of number of workers trained during 1981 to 2010.

We are given the following values:

$r_1 = 0.732$, $r_2 = 0.661$, $r_3 = 0.557$, $r_4 = 0.385$, $r_5 = 0.272$, $r_6 = 0.119$,

$r_7 = 0.019$, $r_8 = -0.139$, $r_9 = -0.268$, $r_{10} = -0.375$, $\overline{y} = 4503.00$ and $\sigma_y = 836.74$

The correlogram for the process is shown in Fig. 16.4.

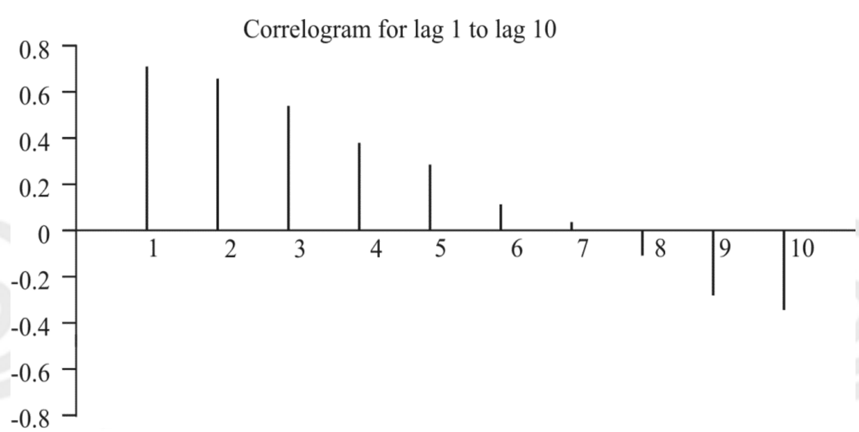

Fig. 16.4: Correlogram for the given series.

Using equation (44a), we get

pacf (1) = $r_1 = 0.732$,

and pacf (2) $= \left(r_2 - r_1^2\right)/\left(1 - r_1^2\right) = 0.125/0.464 = 0.269$

and range = $\pm 2/\sqrt{N} = \pm 2/\sqrt{30} = \pm 0.365$

Since p.a.c.f. (1) lies outside the range, while pacf (2) lies within the range, AR (1) will be suitable for this time series.

On putting the value of $\alpha_1' = r_1 = 0.732$, we get the fitted model as

$$X_t = 0.7332\, X_{t-1} + a_t$$

or $$\left(Y_t - \overline{Y}\right) = 0.732\left(Y_{t-1} - \overline{Y}\right) + a_t$$

or $$\left(Y_t - 4503.0\right) = 0.732\ \left(Y_{t-1} - 4503.0\right) + a_t$$

E8) We are given the ARMA (1,1) model

$$X_t - 0.5\ X_{t-1} = a_t - 0.5\ a_{t-1}$$

Now we start in reverse order and take

$$X_t = a_t \qquad \text{… (i)}$$

Next we take

$$X_{t-1} = a_{t-1} \qquad \text{… (ii)}$$

We multiply equation (ii) by 0.5. Then we have

$$0.5\, X_{t-1} = 0.5\ a_{t-1} \qquad \text{… (iii)}$$

On subtracting (iii) from (i), we get the given model as

$$X_t - 0.5\ X_{t-1} = a_t - 0.5\ a_{t-1}$$

Another way is to write the given model in B form as:

$$\left(1 - 0.5\,B\right)\ X_t = \left(1 - 0.5\,B\right)\ a_t$$

or $$X_t = \left(1 - 0.5B\right)^{-1}\left(1 - 0.5B\right)a_t \Rightarrow X_t = a_t$$

E9) We are given the time series

$$X_t = \ \beta_1 + \beta_2 t + a_t \qquad \text{… (i)}$$

Taking expectation of equation (i), we get

$$E\left(X_t\right) = \beta_1 + \beta_2\ t \qquad \text{… (ii)}$$

Since equation (ii) depends on t and it changes with time, X_t is not a stationary process. Now we consider the first difference series as

$$Y_t = \nabla X_t = (1 - B)X_t = X_t - X_{t-1}$$

$$= \beta_1 + \beta_2 t + a_t - \beta_1 - \beta_2\left(t - 1\right) - a_{t-1}$$

$$= \beta_2 + a_t - a_{t-1}$$

Now the modified series Y_t has a constant mean β_2 and is a stationary MA(1) process.

E10) We are given a time series consisting of 100 observations, which has

r_1=0.31, r_2=0.37, r_3= – 0.05, r_4= 0.06, r_5= – 0.21, r_6=0.11, r_7=0.08, r_8 = 0.05, r_9=0.12, r_{10}= – 0.01

The range,

$$\pm\, 2\big/\sqrt{N}\ = \pm\, 2\big/\sqrt{100} = \pm\, 2/10 = \pm\, 0.20$$

We can see that only r_1 and r_2 are significantly different from zero, and r_5 is marginally significant, which can be ignored as the series is too small (N = 100). In this case only r_1 and r_2 are significant and an MA(2) model is suggested. The correlogram for the process is shown in Fig. 16.5.

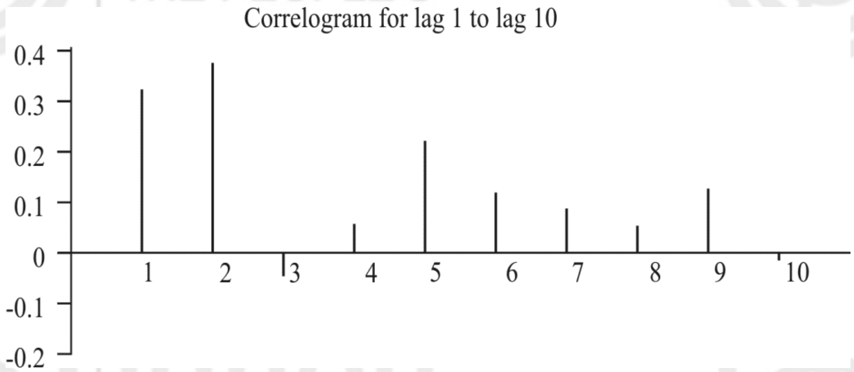

Fig. 16.5: Correlogram of the given time series.