(a) Is the process stationary? Why?
(b) Find the ACF of the above process.

3.14 (a) Find the AR representation of the MA(1) process

$$Z_t = a_t - .4a_{t-1}.$$

(b) Find the MA representation of the AR(2) process

$$Z_t = .2Z_{t-1} + .4Z_{t-2} + a_t.$$

3.15 For each of the following models:
(i) $(1-B)Z_t = (1-1.5B)a_t$,
(ii) $(1-.8B)Z_t = (1-.5B)a_t$,
(iii) $(1-1.1B+.8B^2)Z_t = (1-1.7B+.72B^2)a_t$,
(iv) $(1-.6B)Z_t = (1-1.2B+.2B^2)a_t$
(a) Verify whether it is stationary and/or invertible.
(b) Express the model in an MA representation if it exists.
(c) Express the model in an AR representation if it exists.

3.16 For each of the following processes:
(i) $(1-.6B)Z_t = (1-.9B)a_t$,
(ii) $(1-1.4B+.6B^2)Z_t = (1-.8B)a_t$
(a) Find the ACF ρ_k.
(b) Find the PACF ϕ_{kk} for $k = 1, 2, 3$.
(c) Find the autocovariance generating function.

3.17 Simulate a series of 100 observations from each of the models with $\sigma_a^2 = 1$ in Exercise 3.16. For each simulated series, plot the series, calculate, and study its sample ACF $\hat{\rho}_k$ and PACF $\hat{\phi}_{kk}$ for $k = 0, 1, \ldots, 20$.

4 NONSTATIONARY TIME SERIES MODELS

The time series processes we have discussed so far are all stationary processes. However, many applied time series, particularly those arising from economic and business areas, are nonstationary. With respect to the class of covariance stationary processes, nonstationary time series can occur in many different ways. They could have nonconstant means μ_t, time varying second moments such as nonconstant variance σ_t^2, or have both of these properties. For example, the monthly series of unemployed females between ages 16 and 19 in the United States from January 1961 to December 1985 plotted in Figure 4.1 clearly shows that the mean level changes with time. The plot of the yearly U. S. tobacco production between 1871 and 1979 shown in Figure 4.2 indicates both that the mean level depends on time and that the variance increases as the mean level increases.

In this chapter we illustrate the construction of a very useful class of homogeneous nonstationary time series models—the autoregressive integrated moving average (ARIMA) models. Some useful differencing and variance stabilizing transformations are introduced to connect the stationary and nonstationary time series models.

4.1 NONSTATIONARITY IN THE MEAN

A process nonstationary in the mean could pose a very serious problem for estimation of the time dependent mean function without multiple realizations. Fortunately, there are models that can be constructed from a single realization to describe this time dependent phenomenon. Two such classes of models that are very useful in modeling time series nonstationary in the mean are introduced in this section.

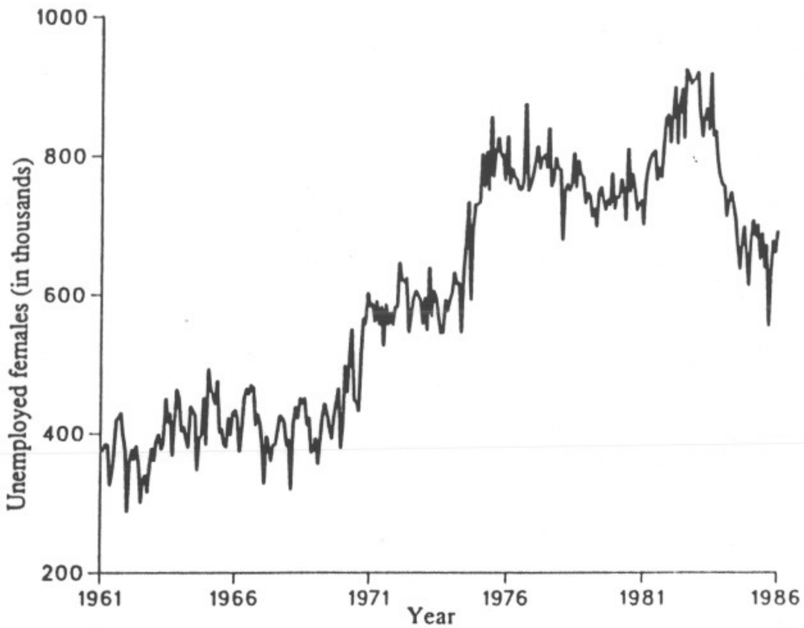

Fig. 4.1 Monthly series of unemployed females between ages 16 and 19 in the United States from January 1961 to December 1985.

4.1.1 Deterministic Trend Models

The mean function of a nonstationary process could be represented by a deterministic trend of time. In such a case, a standard regression model might be used to describe the phenomenon. For example, if the mean function μ_t follows a linear trend, $\mu_t = \alpha_0 + \alpha_1 t$, one can use the deterministic linear trend model

$$Z_t = \alpha_0 + \alpha_1 t + a_t, \tag{4.1.1}$$

with the a_t being a zero mean white noise series. For a deterministic quadratic mean function, $\mu_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2$, one can use

$$Z_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + a_t. \tag{4.1.2}$$

More generally, if the deterministic trend can be described by a kth order polynomial of time, one can model the process by

$$Z_t = \alpha_0 + \alpha_1 t + \cdots + \alpha_k t^k + a_t. \tag{4.1.3}$$

If the deterministic trend can be represented by a sine-cosine curve, one can use

$$Z_t = \nu_0 + \nu \cos(\omega t + \theta) + a_t \tag{4.1.4}$$

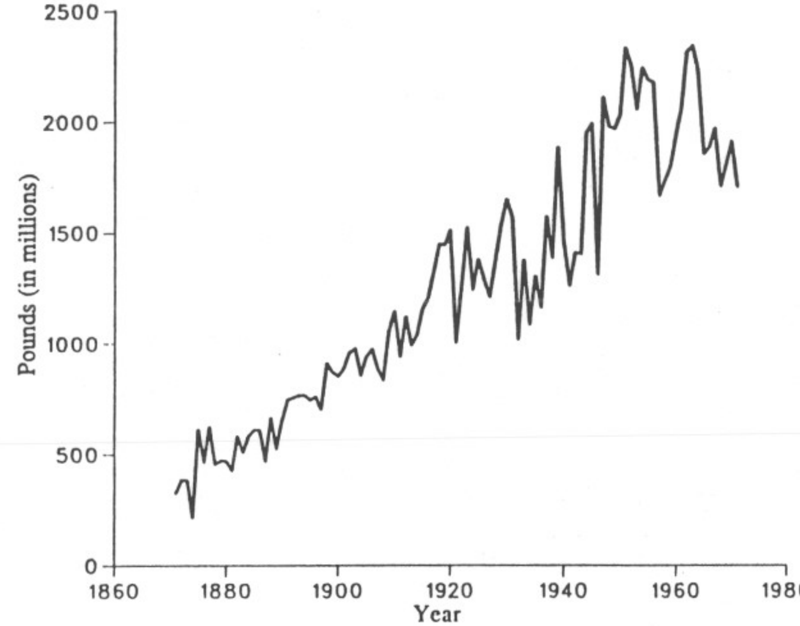

Fig. 4.2 Yearly U. S. tobacco production between 1871 and 1971.

$$= \nu_0 + \alpha \cos(\omega t) + \beta \sin(\omega t) + a_t, \tag{4.1.5}$$

where

$$\alpha = \nu \cos\theta, \qquad \beta = -\nu \sin\theta, \tag{4.1.6}$$

$$\nu = \sqrt{\alpha^2 + \beta^2}, \tag{4.1.7}$$

and

$$\theta = \tan^{-1}(-\beta/\alpha). \tag{4.1.8}$$

We call ν the amplitude, ω the frequency, and θ the phase of the curve. More generally, we can have

$$Z_t = \nu_0 + \sum_{j=1}^{m} (\alpha_j \cos \omega_j t + \beta_j \sin \omega_j t) + a_t, \tag{4.1.9}$$

which is often called the model of hidden periodicities. These models can be handled through standard regression analysis and are discussed again later in Chapter 11.

4.1.2 Stochastic Trend Models and Differencing

Although many time series are nonstationary, due to some equilibrium forces, different parts of these series behave very much alike except for their difference in the local mean levels. Box and Jenkins (1976, p. 85) refer to this kind of nonstationary behavior as homogeneous nonstationary. In terms of the ARMA models, the process is nonstationary if some roots of its AR polynomial do not lie outside the unit circle. However, by the nature of homogeneity, the local behavior of this kind of homogeneous nonstationary series is independent of its level. Hence, by letting $\psi(B)$ be the autoregressive operator describing the behavior, we have

$$\Psi(B)(Z_t + C) = \Psi(B)Z_t \tag{4.1.10}$$

for any constant C. This implies that $\Psi(B)$ must be of the form

$$\Psi(B) = \phi(B)(1-B)^d, \tag{4.1.11}$$

for some $d > 0$ where $\phi(B)$ is a stationary autoregressive operator. Thus, a homogeneous nonstationary series can be reduced to a stationary series by taking a suitable difference of the general series. In other words, the series $\{Z_t\}$ is nonstationary but its dth differenced series, $\{(1-B)^d Z_t\}$ for some integer $d \geq 1$, is stationary. For example, if the dth differenced series follows a white noise phenomenon, we have

$$(1-B)^d Z_t = a_t. \tag{4.1.12}$$

To see the implication of this kind of homogeneous nonstationary series, consider $d = 1$ in (4.1.12), i.e.,

$$(1-B)Z_t = a_t \tag{4.1.13a}$$

or

$$Z_t = Z_{t-1} + a_t. \tag{4.1.13b}$$

Given the past information $Z_{t-1}, Z_{t-2}, \ldots$, the level of the series at time t is

$$\mu_t = Z_{t-1}, \tag{4.1.14}$$

which is subject to the stochastic disturbance at time $(t-1)$. In other words, the mean level of the process Z_t in $(1-B)^d Z_t$ for $d \geq 1$ changes through time stochastically, and we characterize the process as having a stochastic trend. This is different from the deterministic trend model mentioned in the previous section, where the mean level of the process at time t is a pure deterministic function of time.

4.2 AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODELS

4.2.1 The General ARIMA Model

Obviously, the stationary process resulting from a properly differenced homogeneous nonstationary series is not necessarily white noise as in (4.1.12). More generally, the differenced series $(1-B)^d Z_t$ follows the general stationary ARMA(p, q) process discussed in (3.4.1) of Chapter 3. Thus, we have

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t, \tag{4.2.1}$$

where the stationary AR operator $\phi_p(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ and the invertible MA operator $\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$ share no common factors. The parameter θ_0 plays very different roles for $d = 0$ and $d > 0$. When $d = 0$, the original process is stationary, and we recall from (3.4.16) that θ_0 is related to the mean of the process, i.e., $\theta_0 = \mu(1 - \phi_1 - \cdots - \phi_p)$. However, when $d \geq 1$, θ_0 is called the deterministic trend term and, as shown in the next section, is often omitted from the model unless it is really needed.

The resulting homogeneous nonstationary model in (4.2.1) has been referred to as the Autoregressive Integrated Moving Average model of order (p, d, q) and is denoted as the ARIMA(p, d, q) model. When $p = 0$, the ARIMA(p, d, q) model is also called the Integrated Moving Average model of order (d, q) and is denoted as the IMA(d, q) model. In the following, we illustrate some commonly encountered ARIMA models.

4.2.2 The Random Walk Model

In (4.2.1), if $p = 0$, $d = 1$, and $q = 0$, we have the well-known random walk model,

$$(1-B)Z_t = a_t \tag{4.2.2a}$$

or

$$Z_t = Z_{t-1} + a_t. \tag{4.2.2b}$$

This model has been widely used to describe the behavior of the series of a stock price. In the random walk model the value of Z at time t is equal to its value at time $(t-1)$ plus a random shock. This behavior is similar to following a drunken man whose position at time t is his position at time $(t-1)$ plus a step in a random direction at time t.

Note that the random walk model is the limiting process of the AR(1) process $(1 - \phi B)Z_t = a_t$ with $\phi \to 1$. Because the autocorrelation function of the AR(1) process is $\rho_k = \phi^k$, as $\phi \to 1$, the random walk model phenomenon can be characterized by large nonvanishing spikes in the sample ACF of the original series $\{Z_t\}$ and insignificant zero ACF for the differenced series $\{(1-B)Z_t\}$.

Next, consider the following simple modification of (4.2.2a) with a nonzero constant term

$$(1-B)Z_t = \theta_0 + a_t \tag{4.2.3}$$

or

$$Z_t = Z_{t-1} + \theta_0 + a_t. \tag{4.2.4}$$

With reference to a time origin k, by successive substitution, we have

$$\begin{aligned} Z_t &= Z_{t-1} + \theta_0 + a_t \\ &= Z_{t-2} + 2\theta_0 + a_t + a_{t-1} \\ &\vdots \\ &= Z_k + (t-k)\theta_0 + \sum_{j=k+1}^{t} a_j, \qquad \text{for } t > k. \end{aligned} \tag{4.2.5}$$

Thus, it is clear that Z_t contains a deterministic trend with slope or drift θ_0. More generally, for the model involving the dth differenced series $\{(1-B)^d Z_t\}$ the nonzero θ_0 in (4.2.1) can be shown to correspond to the coefficient α_d of t^d in the deterministic trend, $\alpha_0 + \alpha_1 t + \cdots + \alpha_d t^d$. For this reason, when $d > 0$, θ_0 is referred to as the deterministic trend term. For large t, this term can become very dominating so that it forces the series to follow a deterministic pattern. Hence, in general, when $d > 0$, we assume that $\theta_0 = 0$, unless it is clear from the data or the nature of the problem that a deterministic component is really needed.

The process in (4.2.3) with $\theta_0 \neq 0$ is usually called the random walk model with drift. Given $Z_{t-1}, Z_{t-2}, \ldots$, by (4.2.4) the mean level of the series Z_t at time t is

$$\mu_t = Z_{t-1} + \theta_0, \tag{4.2.6}$$

which is influenced by the stochastic disturbance at time $(t-1)$ through the term Z_{t-1} as well as by the deterministic component through the slope θ_0. When $\theta_0 = 0$, we have a model with only a stochastic trend.

Example 4.1 To illustrate the results of the random walk model discussed in this section, we simulated 100 observations each from the model $(1-B)Z_t = a_t$ and the model $(1-B)Z_t = 4 + a_t$, where the a_t in both models are i.i.d. normal $N(0, 1)$ white noise. The sample ACF and PACF of the original series are shown in Table 4.1 and Figure 4.3. The fact that the ACF for both series decays very slowly indicates that it is nonstationary. To identify the model properly, we calculate the sample ACF and PACF of the differenced series $(1-B)Z_t$ as shown in Table 4.2 and Figure 4.4. As expected, both exhibit phenomena of a

Table 4.1 Sample ACF and sample PACF for the original series Z_t simulated from random walk models.

	(a) For Z_t from $(1-B)Z_t = a_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat\rho_k$	.94	.88	.83	.77	.71	.66	.60	.53	.46	.40
St.E.	.10	.17	.21	.24	.26	.28	.30	.31	.32	.32
$\hat\phi_{kk}$	.94	−.07	−.01	−.04	.02	−.07	−.04	−.15	.02	−.04
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
	(b) For Z_t from $(1-B)Z_t = 4 + a_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat\rho_k$	.94	.94	.91	.88	.85	.82	.79	.76	.73	.70
St.E.	.10	.17	.22	.25	.28	.30	.33	.34	.36	.38
$\hat\phi_{kk}$	.97	−.01	−.01	−.02	−.01	−.01	−.02	−.02	−.01	−.01
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10

white noise process. In fact, the values of $\hat\rho_k$ and $\hat\phi_{kk}$ are coincidently equal for the two models. Then how can we tell the difference between the regular random walk model and the random walk model with drift? We cannot if we rely only on their autocorrelation structures, although the sample ACF of the original series from a random walk model with drift generally decays more slowly. However, if we look at their behaviors as plotted in Figure 4.5, the difference is striking. The series of the random walk model with drift is clearly dominated by the deterministic linear trend with slope 4. On the other hand, the nonstationarity of the random walk model without drift is shown through a stochastic trend and its values are free to wander.

4.2.3 The ARIMA(0,1,1) or IMA(1,1) Model

When $p = 0$, $d = 1$, and $q = 1$, the model in (4.2.1) becomes

$$(1-B)Z_t = (1-\theta B)a_t \tag{4.2.7a}$$

or

$$Z_t = Z_{t-1} + a_t - \theta a_{t-1}, \tag{4.2.7b}$$

where $-1 < \theta < 1$. This IMA(1, 1) model is reduced to a stationary MA(1) process after taking the first difference. The random walk model is a special case of this IMA(1, 1) model with $\theta = 0$. Thus, the basic phenomenon of the IMA(1, 1) model is characterized by the facts that the sample ACF of the original series fails to die out and that the sample ACF of the first differenced series exhibits the pattern of a first order moving average phenomenon.

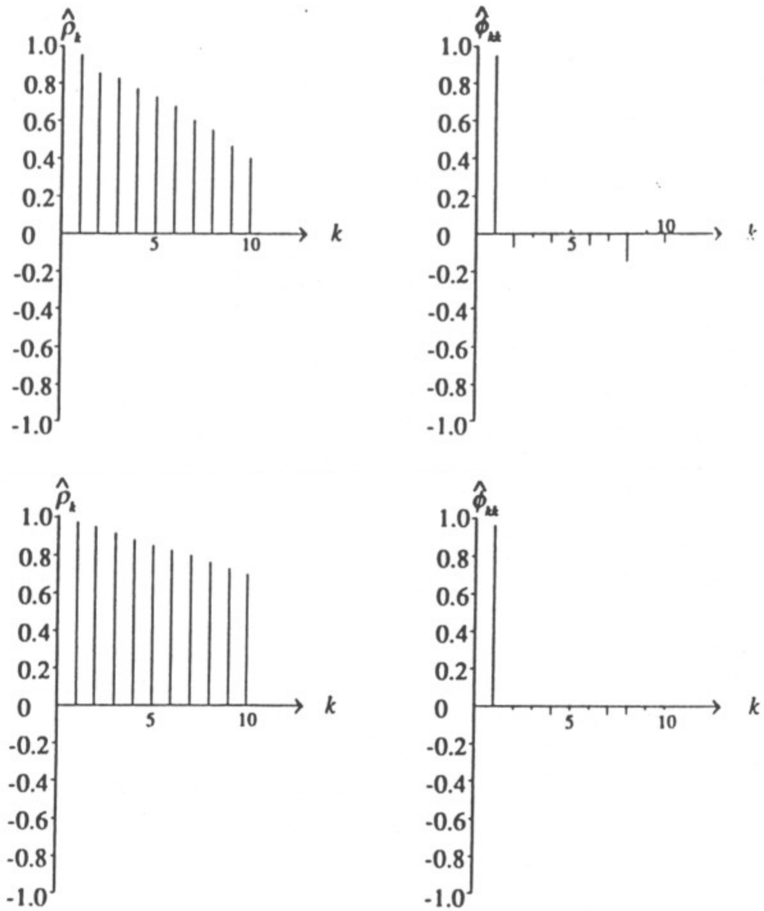

Fig. 4.3 Sample ACF and PACF of the random walk model. (a) For Z_t from $(1-B)Z_t = a_t$. (b) For Z_t from $(1-B)Z_t = 4 + a_t$.

For $-1 < \theta < 1$,

$$\frac{(1-B)}{(1-\theta B)} = (1-B)(1+\theta B+\theta^2B^2+\cdots)$$
$$= 1+\theta B+\theta^2B^2+\cdots$$
$$-B-\theta B^2-\cdots$$
$$= 1-(1-\theta)B-(1-\theta)\theta B^2-(1-\theta)\theta^2B^3-\cdots$$
$$= 1-\alpha B-\alpha(1-\alpha)B^2-\alpha(1-\alpha)^2B^3-\cdots \tag{4.2.8}$$

Table 4.2 Sample ACF and sample PACF for the differenced series $W_t = (1-B)Z_t$ simulated from random walk models.

	(a) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = a_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.11	.03	.00	.00	.11	.02	.06	.01	−.02	.06
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
$\hat{\phi}_{kk}$	.11	.02	.00	.00	.11	.00	.05	.00	−.03	.05
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
	(b) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = 4 + a_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.11	.03	.00	.00	.11	.02	.06	.01	−.02	.06
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
$\hat{\phi}_{kk}$	.11	.02	.00	.00	.11	.00	.05	.00	−.03	.05
St.E.	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10

where $\alpha = (1-\theta)$. Hence,

$$Z_t = \alpha\sum_{j=1}^{\infty}(1-\alpha)^{j-1}Z_{t-j} + a_t. \tag{4.2.9}$$

This is the AR representation of the model, and from the results of regression analysis the optimal forecast, $\hat{Z}_t$, of Z_t is given by

$$\hat{Z}_t = \alpha\sum_{j=1}^{\infty}(1-\alpha)^{j-1}Z_{t-j}. \tag{4.2.10}$$

In other words, the optimal forecast of Z_t at time t is an exponentially decreasing weighted moving average of its past values $Z_{t-1}, Z_{t-2}, \ldots$, etc. Moreover, (4.2.10) implies that

$$\hat{Z}_{t+1} = \alpha\sum_{j=1}^{\infty}(1-\alpha)^{j-1}Z_{t+1-j}$$
$$= \alpha Z_t + (1-\alpha)\alpha\sum_{j=2}^{\infty}(1-\alpha)^{j-2}Z_{t+1-j}$$
$$= \alpha Z_t + (1-\alpha)\alpha\sum_{i=1}^{\infty}(1-\alpha)^{i-1}Z_{t-i}$$
$$= \alpha Z_t + (1-\alpha)\hat{Z}_t,$$

which shows that the new forecast of Z at the next time period is equal to the weighted average of the newly available observation and the last forecast. The

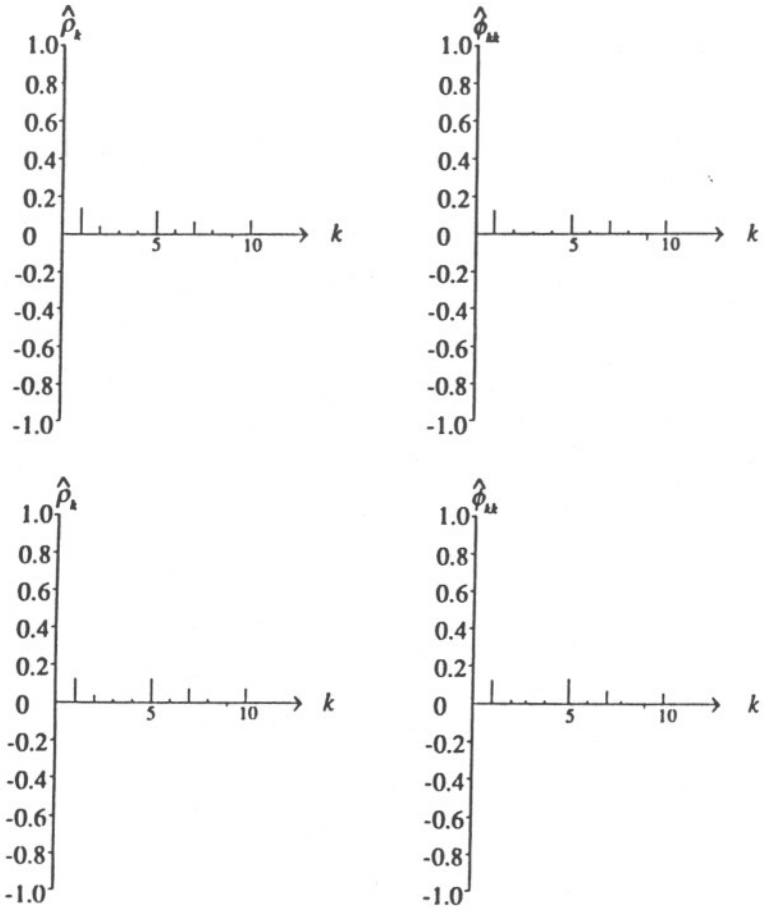

Fig. 4.4 Sample ACF and PACF of differences of random walk models. (a) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = a_t$. (b) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = 4 + a_t$.

coefficient α is often called the smoothing constant in the method of exponential smoothing. Thus, the general ARIMA(p, d, q) model contains many smoothing methods as special cases. See Abraham and Ledolter (1983) for a more detailed discussion of the relationship between exponential smoothing and the ARIMA models.

Example 4.2 We simulated 250 values for each of the following three ARIMA models: (1) the ARIMA(1, 1, 0) model, $(1-.8B)(1-B)Z_t = a_t$; (2) the ARIMA(0, 1, 1) model, $(1-B)Z_t = (1-.75B)a_t$; and (3) the ARIMA(1, 1, 1) model, $(1-.9B)(1-B)Z_t = (1-.5B)a_t$. The series a_t are independent Gaus-

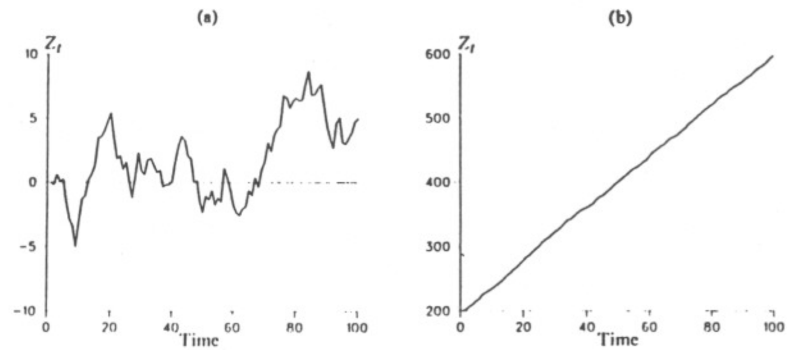

Fig. 4.5 Simulated series from random walk models. (a) A simulated random walk series: $(1-B)Z_t = a_t$. (b) A simulated random walk series with drift: $(1-B)Z_t = 4 + a_t$.

sian $N(0, 1)$ white noise. The sample ACF and PACF of the original three series are computed and shown in Table 4.3 and Figure 4.6. Each one shows the same phenomenon of the sustained large ACF and an exceptionally large first lag PACF. The dominating phenomenon of nonstationarity overshadows the fine details of the underlying characteristics of these models. To remove the shadow, we take differences for each of these series. The sample ACF and PACF for these differenced series are shown in Table 4.4 and Figure 4.7. Now, all the fine details are evident. The ACF $\hat{\rho}_k$ of $W_t = (1-B)Z_t$ from the ARIMA(1, 1, 0) model tails off while its PACF $\hat{\phi}_{kk}$ cuts off after lag 1; the ACF $\hat{\rho}_k$ of $W_t = (1-B)Z_t$ from the ARIMA(0, 1, 1) model cuts off after lag 1 while its PACF $\hat{\phi}_{kk}$ tails off; both the ACF $\hat{\rho}_k$ and PACF $\hat{\phi}_{kk}$ of $W_t = (1-B)Z_t$ from the ARIMA(1, 1, 1) model tail off as expected from their characteristics discussed in Chapter 3.

4.3 NONSTATIONARITY IN THE VARIANCE AND THE AUTOCOVARIANCE

4.3.1 Variance and Autocovariance of the ARIMA Models

A process that is stationary in the mean is not necessarily stationary in the variance and the autocovariance. However, a process that is nonstationary in the mean will also be nonstationary in the variance and the autocovariance. As we have shown in the previous section, the mean function of the ARIMA model is time dependent. We now show that the ARIMA model is also nonstationary in its variance and autocovariance functions.

First, we note a very fundamental phenomenon about the ARIMA model. That is, although the model is nonstationary, the complete characteristic of the process is determined for all time only by a finite number of parameters, i.e.,

Table 4.3 Sample ACF and sample PACF for the original series Z_t simulated from three ARIMA models.

(a) For Z_t from $(1-.8B)(1-B)Z_t = a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.97	.95	.92	.89	.86	.83	.80	.77	.74	.72
St.E.	.06	.11	.14	.16	.18	.19	.21	.22	.23	.24
$\hat{\phi}_{kk}$	.97	−.06	−.04	−.02	−.01	−.01	−.01	−.00	.02	.03
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
(b) For Z_t from $(1-B)Z_t = (1-.75)a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.96	.91	.86	.82	.79	.75	.72	.69	.67	.66
St.E.	.06	.11	.13	.15	.17	.19	.20	.21	.22	.22
$\hat{\phi}_{kk}$	.96	−.13	.04	.01	.03	.01	.00	.13	.07	.06
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
(c) For Z_t from $(1-.9B)(1-B)Z_t = (1-.5B)a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.98	.96	.94	.91	.89	.87	.84	.82	.80	.78
St.E.	.06	.11	.14	.16	.18	.20	.21	.23	.24	.25
$\hat{\phi}_{kk}$	.98	−.03	−.03	−.02	−.02	−.02	−.02	−.01	.01	.02
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

the ϕ_i, the θ_j, and σ_a^2. Thus, the complete future evolution of the process can be developed from the ARIMA model fitted to a given data set, $\{Z_1, Z_2, \ldots, Z_n\}$. For example, suppose we fit the IMA(1, 1) model

$$(1-B)Z_t = (1-\theta B)a_t \tag{4.3.1a}$$

or

$$Z_t = Z_{t-1} + a_t - \theta a_{t-1} \tag{4.3.1b}$$

to a series of n_0 observations. With reference to this time origin n_0, for $t > n_0$, we can write by successive substitutions,

$$\begin{aligned} Z_t &= Z_{t-1} + a_t - \theta a_{t-1} \\ &= Z_{t-2} + a_t + (1-\theta)a_{t-1} - \theta a_{t-2} \\ &\vdots \\ &= Z_{n_0} + a_t + (1-\theta)a_{t-1} + \cdots + (1-\theta)a_{n_0+1} - \theta a_{n_0}. \end{aligned} \tag{4.3.2}$$

Similarly, for $t-k > n_0$,

$$Z_{t-k} = Z_{n_0} + a_{t-k} + (1-\theta)a_{t-k-1} + \cdots + (1-\theta)a_{n_0+1} - \theta a_{n_0}. \tag{4.3.3}$$

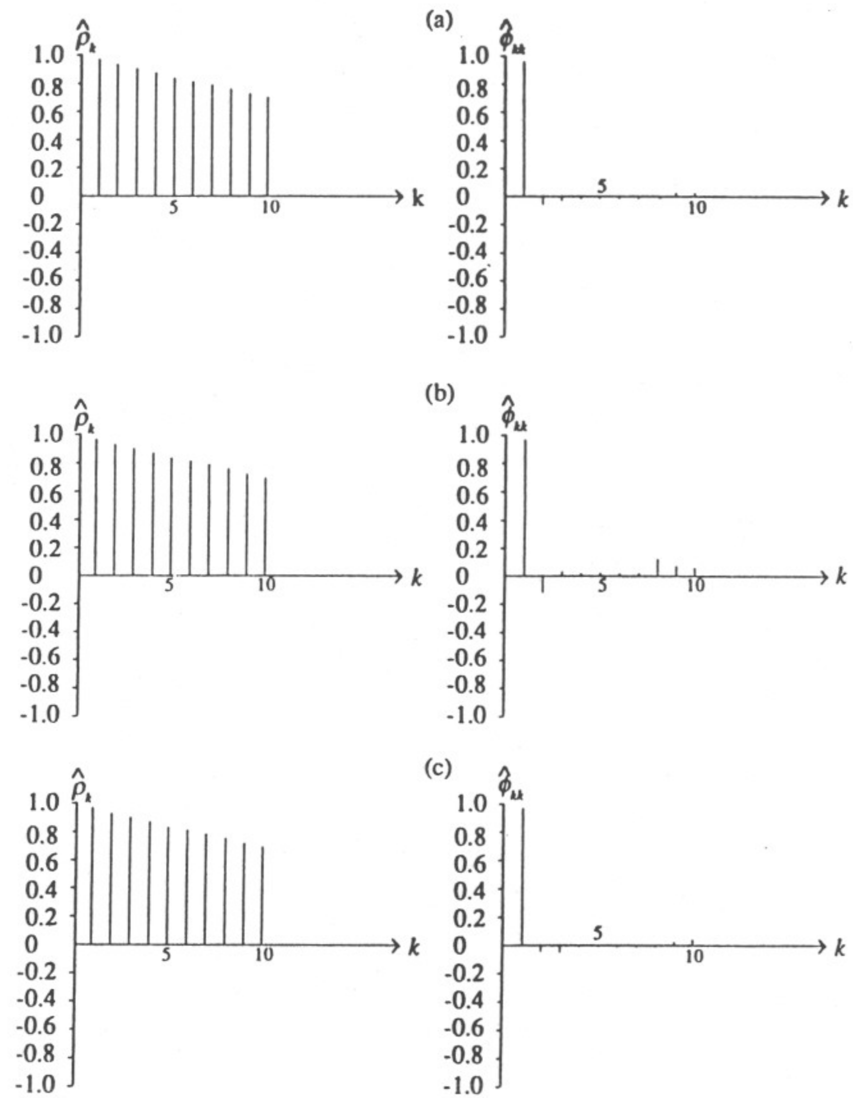

Fig. 4.6 Sample ACF and PACF from three ARIMA models. (a) For Z_t from $(1-.8B)(1-B)Z_t = a_t$. (b) For Z_t from $(1-B)Z_t = (1-.75B)a_t$. (c) For Z_t from $(1-.9B)(1-B)Z_t = (1-.5B)a_t$.

Table 4.4 Sample ACF and sample PACF for the differenced series $W_t = (1-B)Z_t$ simulated from three ARIMA models.

(a) For $W_t = (1-B)Z_t$ from $(1-.8B)(1-B)Z_t = a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.71	.50	.36	.20	.12	.05	.02	.00	−.03	−.01
St.E.	.06	.09	.10	.11	.11	.11	.11	.11	.11	.11
$\hat{\phi}_{kk}$	.71	.01	.01	−.12	.04	−.04	.00	.11	−.03	.05
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
(b) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = (1-.75B)a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.40	−.07	−.02	−.10	−.08	−.06	−.06	−.14	−.04	.04
St.E.	.06	.07	.07	.07	.07	.07	.07	.07	.07	.07
$\hat{\phi}_{kk}$	.40	−.28	.16	−.23	.11	−.16	.07	−.10	.02	.05
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
(c) For $W_t = (1-B)Z_t$ from $(1-.9B)(1-B)Z_t = (1-.5B)a_t$										
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.41	.34	.33	.17	.18	.12	.08	.11	.05	.09
St.E.	.06	.07	.08	.08	.09	.09	.09	.09	.09	.09
$\hat{\phi}_{kk}$	.41	.20	.17	−.06	.04	−.01	.00	.04	−.03	.06
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

Hence, with respect to the time origin n_0,

$$\text{Var}(Z_t) = [1 + (t - n_0 - 1)(1-\theta)^2]\sigma_a^2 \tag{4.3.4}$$

$$\text{Var}(Z_{t-k}) = [1 + (t - k - n_0 - 1)(1-\theta)^2]\sigma_a^2 \tag{4.3.5}$$

$$\text{Cov}(Z_{t-k}, Z_t) = [(1-\theta) + (t - k - n_0 - 1)(1-\theta)^2]\sigma_a^2, \tag{4.3.6}$$

where we note that Z_{n_0} and a_{n_0} are known values with respect to the time origin n_0, and

$$\text{Corr}(Z_{t-k}, Z_t) = \frac{\text{Cov}(Z_{t-k}, Z_t)}{\sqrt{\text{Var}(Z_{t-k})\,\text{Var}(Z_t)}} \tag{4.3.7}$$

$$= \frac{(1-\theta) + (t-k-n_0-1)(1-\theta)^2}{\sqrt{[1+(t-k-n_0-1)(1-\theta)^2][1+(t-n_0-1)(1-\theta)^2]}}.$$

Now, from (4.3.4) to (4.3.7), we have the following important observations:

1. The variance, $\text{Var}(Z_t)$, of the ARIMA process is time dependent and $\text{Var}(Z_t) \neq \text{Var}(Z_{t-k})$ for $k \neq 0$.
2. The variance $\text{Var}(Z_t)$ is unbounded as $t \to \infty$.
3. The autocovariance $\text{Cov}(Z_{t-k}, Z_t)$ and the autocorrelation $\text{Corr}(Z_{t-k}, Z_t)$ of the process are also time dependent, and hence not invariant with respect

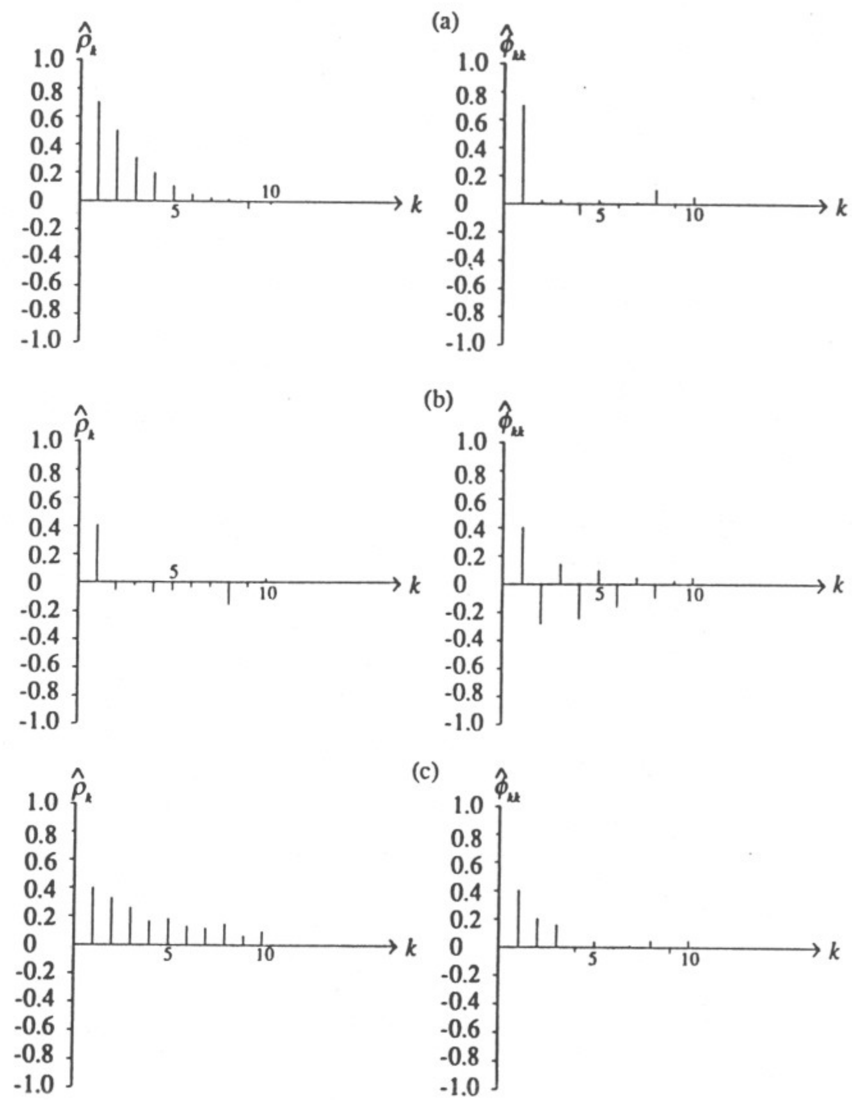

Fig. 4.7 Sample ACF and PACF for differences from three ARIMA models. (a) For $W_t = (1-B)Z_t$ from $(1-.8B)(1-B)Z_t = a_t$. (b) For $W_t = (1-B)Z_t$ from $(1-B)Z_t = (1-.75B)a_t$. (c) For $W_t = (1-B)Z_t$ from $(1-.9B)(1-B)Z_t = (1-.5B)a_t$.

to time translation. In other words, they are not only functions of the time difference k but also functions of both the time origin t and the original reference point n_0.

4. If n_0 is large with respect to k, from (4.3.7), $\text{Corr}(Z_{t-k}, Z_t) \simeq 1$. Because $|\text{Corr}(Z_{t-k}, Z_t)| \leq 1$, this implies that the autocorrelation function vanishes slowly as k increases.

In general, with only a single realization, it is difficult or impossible to make the statistical inferences of a process that is nonstationary in both the mean and the autocovariance or autocorrelation function. Fortunately, for the homogeneous nonstationary process, we can apply a proper differencing to reduce it to stationary process. That is, although the original series Z_t is nonstationary, its properly differenced series $W_t = (1-B)^d Z_t$ is stationary and can be represented as an ARMA process

$$\phi(B)W_t = \theta(B)a_t \tag{4.3.8}$$

where

$$\phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p) \quad \text{and} \quad \theta(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$$

have all their roots outside of the unit circle. Thus, the parameters ϕ_i, θ_j, and σ_a^2 that control the evolution of the nonstationary phenomenon of Z_t can be estimated from the differenced series W_t in exactly the same way as the stationary case discussed in Chapter 7.

4.3.2 Variance Stabilizing Transformations

Not all nonstationary problems can be cured by differencing. There are many time series that are stationary in the mean but nonstationary in the variance. To overcome this problem, we need a proper variance stabilizing transformation.

It is very common for the variance of a nonstationary process to change as its level changes. Thus,

$$\text{Var}(Z_t) = cf(\mu_t) \tag{4.3.9}$$

for some positive constant c and function f. How do we find a function T so that the transformed series, $T(Z_t)$, has a constant variance? To illustrate the method, we approximate the desired function by a first order Taylor series about the point μ_t. Let

$$T(Z_t) \simeq T(\mu_t) + T'(\mu_t)(Z_t - \mu_t) \tag{4.3.10}$$

where $T'(\mu_t)$ is the first derivative of $T(Z_t)$ evaluated at μ_t. Now

$$\begin{aligned}\text{Var}[T(Z_t)] &= [T'(\mu_t)]^2 \,\text{Var}(Z_t) \\ &= c[T'(\mu_t)]^2 f(\mu_t).\end{aligned} \tag{4.3.11}$$

Thus, in order that the variance of $T(Z_t)$ be constant the variance stabilizing transformation $T(Z_t)$ must be chosen so that

$$T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}}. \tag{4.3.12}$$

Equation (4.3.12) implies that

$$T(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t. \tag{4.3.13}$$

For example, if the standard deviation of a series is proportional to the level so that $\text{Var}(Z_t) = c^2\mu_t^2$, then

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t^2}} d\mu_t = \ln(\mu_t). \tag{4.3.14}$$

Hence, a logarithmic transformation (the base is irrelevant) of the series, $\ln(Z_t)$, will give a constant variance.

Next, if the variance of the series is proportional to the level so that $\text{Var}(Z_t) = c\mu_t$, then

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t}} d\mu_t = 2\sqrt{\mu_t}. \tag{4.3.15}$$

Thus, a square root transformation of the series, $\sqrt{Z_t}$, will give a constant variance.

Third, if the standard deviation of the series is proportional to the square of the level so that $\text{Var}(Z_t) = c^2\mu_t^4$, then

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t^4}} d\mu_t = -\frac{1}{\mu_t}. \tag{4.3.16}$$

Therefore, a desired transformation that gives a constant variance will be the reciprocal $1/Z_t$.

More generally, to stabilize the variance, we can use the power transformation

$$T(Z_t) = Z_t^{(\lambda)} = \frac{Z_t^\lambda - 1}{\lambda}, \tag{4.3.17}$$

introduced by Box and Cox (1964). λ is called the transformation parameter. Some commonly used values of λ and their associated transformations are

Values of λ (lambda)	Transformation
-1.0	$\frac{1}{Z_t}$
-0.5	$\frac{1}{\sqrt{Z_t}}$
0.0	$\ln Z_t$
0.5	$\sqrt{Z_t}$
1.0	Z_t (no transformation)

To see why $\lambda = 0$ corresponds to the logarithmic transformation, we note that

$$\lim_{\lambda \to 0} T(Z_t) = \lim_{\lambda \to 0} Z_t^{(\lambda)} = \lim_{\lambda \to 0} \frac{Z_t^{\lambda} - 1}{\lambda} = \ln(Z_t). \tag{4.3.18}$$

Some important remarks are in order.

1. The variance stabilizing transformations introduced above are defined only for positive series. However, this is not as restrictive as it seems because a constant can always be added to the series without affecting the correlation structure of the series.
2. A variance stabilizing transformation, if needed, should be performed before any other analysis such as differencing.
3. λ in the power transformation can be taken as a parameter in the model to be estimated from the observed series. The maximum likelihood estimate of λ is the one that minimizes the residual sum of squares, a term discussed further in Chapter 7. For any given value of λ, the residual sum of squares is calculated from the fitted model. The maximum likelihood estimate of λ is the one that gives the smallest residual sum of squares from among all values of λ. In actual applications, evaluations of these residual sum of squares are often based on a grid of λ values.
4. Frequently, the transformation not only stabilizes the variance, but also improves the approximation to normality.

Exercises

4.1 Consider the following model:

$$(1-B)^2 Z_t = (1 - .3B - .5B^2)a_t.$$

(a) Is the model for Z_t stationary? Why?
(b) Let $W_t = (1-B)^2 Z_t$. Is the model for W_t stationary? Why?
(c) Find the ACF for the second order differences W_t.

4.2 Consider the following processes:
(a) $(1-B)^2 Z_t = a_t - .81a_{t-1} + .38a_{t-2}$,
(b) $(1-B)Z_t = (1-.5B)a_t$.
Express each of the above processes in the AR representation by actually finding and plotting the π weights.

4.3 (a) Simulate a series of 100 observations from each of the following models:
(i) $(1-B)Z_t = (1-.6B)a_t$,
(ii) $(1-B)Z_t = 5 + (1-.6B)a_t$,
(iii) $(1-.9B)(1-B)Z_t = a_t$,
(iv) $(1-.9B)(1-B)Z_t = (1-.5B)a_t$.
(b) Plot the simulated series.
(c) Calculate and examine the sample ACF, $\hat{\rho}_k$, and PACF, $\hat{\phi}_{kk}$, at $k = 0, 1, \ldots, 20$ for each simulated series.

4.4 Suppose that Z_t is generated according to $Z_t = a_t + ca_{t-1} + \cdots + ca_1$, for $t \geq 1$, where c is a constant.
(a) Find the mean and covariance for Z_t. Is it stationary?
(b) Find the mean and covariance for $(1-B)Z_t$. Is it stationary?

4.5 Consider the stationary MA(1) process $Z_t = (1-\theta B)a_t$, where $|\theta| < 1$. If we take the first differencing of this stationary series, what will be the variance of the differenced series? Compare it with the variance of the original series Z_t.

4.6 Let $Z_1, Z_2, \ldots, Z_n$ be a random sample from a Poisson distribution with mean μ.
(a) Show that the variance of Z_t depends on its mean μ.
(b) Find a proper transformation so that the variance of the transformed variable is independent of μ.
(c) Find the variance of the transformed variable.

4.7 Let r_n be the Pearson correlation coefficient of a sample size of n. It is known that r_n is asymptotically distributed as $N\left(\rho, (1-\rho^2)^2/n\right)$. Show that Fisher's z-transformation $Z = 1/2\ln((1+r_n)/(1-r_n))$ is actually a kind of variance-stabilizing transformation.