6 MODEL IDENTIFICATION

In time series analysis, the most crucial steps are to identify and build a model based on the available data. This requires a good understanding of the processes discussed in Chapters 3 and 4, particularly the characteristics of these processes in terms of their ACF, ρ_k, and PACF, ϕ_{kk}. In practice, these ACF and PACF are unknown, and for a given observed time series $Z_1, Z_2, \ldots,$ and Z_n, they have to be estimated by the sample ACF, $\hat{\rho}_k$, and PACF, $\hat{\phi}_{kk}$ discussed in Section 2.5. Thus, in model identification, our goal is to match patterns in the sample ACF, $\hat{\rho}_k$, and sample PACF, $\hat{\phi}_{kk}$, with the known patterns of the ACF, ρ_k, and PACF, ϕ_{kk}, for the ARMA models. For example, since we know that the ACF cuts off at lag 1 for an MA(1) model, a large single significant spike at lag 1 for $\hat{\rho}_k$ will indicate an MA(1) model as a possible underlying process.

After introducing systematic and useful steps for model identification, we give illustrative examples of identifying models for a wide variety of actual time series data. We also discuss some recently introduced identification tools such as the inverse autocorrelation and extended sample autocorrelation functions.

6.1 STEPS FOR MODEL IDENTIFICATION

To illustrate the model identification, we consider the general ARIMA(p, d, q) model

$$(1-\phi_1 B - \cdots - \phi_p B^p)(1-B)^d Z_t = \theta_0 + (1-\theta_1 B - \cdots - \theta_q B^q)a_t. \quad (6.1.1)$$

Model identification refers to the methodology in identifying the required transformations, such as variance stabilizing transformations and differencing transformations, the decision to include the deterministic parameter θ_0 when $d \geq 1$, and the proper orders of p and q for the model. Given a time series, we use the following useful steps to identify a tentative model.

Step 1. Plot the time series data and choose proper transformations.

In any time series analysis, the first step is to plot the data. Through careful examination of the plot, we usually get a good idea about whether the series contains a trend, seasonality, outliers, nonconstant variances, and other nonnormal and nonstationary phenomena. This understanding often provides a basis for postulating a possible data transformation.

In time series analysis, the most commonly used transformations are variance-stabilizing transformations and differencing. Since differencing may create some negative values, we should always apply variance-stabilizing transformations before taking differences. A series with nonconstant variance often needs a logarithmic transformation. More generally, to stabilize the variance, we can apply Box–Cox's power transformation discussed in Section 4.3.2. Since a variance-stabilizing transformation, if necessary, is often chosen before we do any further analysis, we refer to this transformed data as the original series in the following discussion unless mentioned otherwise.

Step 2. Compute and examine the sample ACF and the sample PACF of the original series to further confirm a necessary degree of differencing. Some general rules are:

1. If the sample ACF decays very slowly (the individual ACF may not be large) and the sample PACF cuts off after lag 1, it indicates that differencing is needed. Try taking the first differencing $(1-B)Z_t$. One can also use the test proposed by Dickey and Fuller (1979). In a borderline case, differencing is generally recommended (see Dickey, Bell, and Miller [1986]).
2. More generally, to remove nonstationarity we may need to consider a higher order differencing $(1-B)^d Z_t$ for $d > 1$. In most cases, d is either 0, 1, or 2. Some authors argue that the consequences of unnecessary differencing are much less serious than those of underdifferencing. But do beware of the artifacts created by overdifferencing, so that unnecessary overparameterization can be avoided.

Step 3. Compute and examine the sample ACF and PACF of the properly transformed and differenced series to identify the orders of p and q (where we recall that p is the highest order in the autoregressive polynomial $(1-\phi_1 B - \cdots - \phi_p B^p)$, and q is the highest order in the moving average polynomial $(1-\theta_1 B - \cdots - \theta_q B^q)$. Usually, the needed orders of these p and q are less than or equal to 3. Table 6.1 summarizes the important results for selecting p and q.

It is useful and interesting to note that a strong duality exists between the AR and the MA models in terms of their ACFs and PACFs. To build a reasonable ARIMA model, ideally, we need a minimum of $n = 50$ observations, and the number of sample ACF and PACF to be calculated should be about $n/4$, although occasionally for data of good quality one may be able to identify an adequate model with a smaller sample size.

We identify the orders p and q by matching the patterns in the sample ACF and PACF with the theoretical patterns of known models. The art of a

Table 6.1 Characteristics of theoretical ACF and PACF for stationary processes.

Process	ACF	PACF
AR(p)	Tails off as exponential decay or damped sine wave	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off as exponential decay or damped sine wave
ARMA(p, q)	Tails off after lag $(q-p)$	Tails off after lag $(p-q)$

time series analyst's model identification is very much like the method of an FBI agent's criminal search. Most criminals disguise themselves to avoid being recognized. This is also true of ACF and PACF. The sampling variation and the correlation among the sample ACF and PACF as shown in Section 2.5 often disguise the theoretical ACF and PACF patterns. Hence, in the initial model identification we always concentrate on the general broad features of these sample ACF and PACF without focusing on the fine details. Model improvement can be easily achieved at a later stage of diagnostic checking.

The estimated variances of both the sample ACF and PACF given in (2.5.21) and (2.5.27) are very rough approximations. Some authors recommend that a conservative threshold of 1.5 standard deviations be used in checking the significance of the short-term lags of these ACF and PACF at the initial model identification phase. This is especially true for a relatively short series.

Step 4. Test the deterministic trend term θ_0 when $d > 0$.

As discussed in Section 4.2, for a nonstationary model, $\phi(B)(1-B)^d Z_t = \theta_0 + \theta(B)a_t$, the parameter θ_0 is usually omitted so that it is capable of representing series with random changes in the level, slope, or trend. However, if there is reason to believe that the differenced series contains a deterministic trend mean, we can test for its inclusion by comparing the sample mean $\overline{W}$ of the differenced series $W_t = (1-B)^d Z_t$ with its approximate standard error $S_{\overline{W}}$. To derive $S_{\overline{W}}$, we note from Section 2.5.1 that $\lim_{n\to\infty} n \operatorname{Var}(\overline{W}) = \sum_{j=-\infty}^{\infty} \gamma_j$, and hence

$$\sigma_{\overline{W}}^2 = \frac{\gamma_0}{n} \sum_{j=-\infty}^{\infty} \rho_j = \frac{1}{n} \sum_{j=-\infty}^{\infty} \gamma_j = \frac{1}{n}\gamma(1), \tag{6.1.2}$$

where $\gamma(B)$ is the autocovariance generating function defined in (2.6.8) and $\gamma(1)$ is its value at $B = 1$. Thus, the variance and hence the standard error for $\overline{W}$ is model dependent. For example, for the ARIMA(1, d, 0) model, $(1 - \phi B)W_t = a_t$, we have, from (2.6.9),

$$\gamma(B) = \frac{\sigma_a^2}{(1-\phi B)(1-\phi B^{-1})}$$

so that

$$\begin{aligned}\sigma_{\overline{W}}^2 &= \frac{\sigma_a^2}{n}\frac{1}{(1-\phi)^2} = \frac{\sigma_W^2}{n}\frac{1-\phi^2}{(1-\phi)^2} \\ &= \frac{\sigma_W^2}{n}\left(\frac{1+\phi}{1-\phi}\right) = \frac{\sigma_W^2}{n}\left(\frac{1+\rho_1}{1-\rho_1}\right),\end{aligned} \tag{6.1.3}$$

where we note that $\sigma_W^2 = \sigma_a^2/(1-\phi^2)$. The required standard error is

$$S_{\overline{W}} = \sqrt{\frac{\hat{\gamma}_0}{n}\left(\frac{1+\hat{\rho}_1}{1-\hat{\rho}_1}\right)}. \tag{6.1.4}$$

Expressions of $S_{\overline{W}}$ for other models can be derived similarly. However, at the model identification phase, since the underlying model is unknown, most available software use the approximation

$$S_{\overline{W}} = \left[\frac{\hat{\gamma}_0}{n}(1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \cdots + 2\hat{\rho}_k)\right]^{1/2} \tag{6.1.5}$$

where $\hat{\gamma}_0$ is the sample variance and $\hat{\rho}_1, \ldots, \hat{\rho}_k$ are the first k significant sample ACFs of $\{W_t\}$. Under the null hypothesis $\rho_k = 0$ for $k \geq 1$, Equation (6.1.5) reduces to

$$S_{\overline{W}} = \sqrt{\hat{\gamma}_0/n}. \tag{6.1.6}$$

Alternatively, one can include θ_0 initially and discard it at the final model estimation if the preliminary estimation result is not significant. Parameter estimation is discussed in the next chapter.

6.2 EMPIRICAL EXAMPLES

In this section we present a variety of real-world examples to illustrate the method of model identification. Several mainframe computer programs such as BMDP, MINITAB, SAS, SCA, and SPSS, and PC software such as AUTOBOX and RATS are available to facilitate the methods. Some programs are available for both the mainframe and personal computers.

Example 6.1 Figure 6.1 shows Series W1, which is the daily average number of defects per truck found in the final inspection at the end of the assembly line of a truck manufacturing plant. The data consist of 45 daily observations of consecutive business days between November 4 to January 10, as reported in Bun (1976, p. 134). The figure suggests a stationary process with constant mean and variance. The sample ACF and sample PACF are calculated in Table 6.2 and plotted in Figure 6.2. The fact that the sample ACF decays exponentially and the sample PACF has a single spike at lag 1 indicates that the series is likely to be generated by an AR(1) process,

$$(1-\phi B)(Z_t - \mu) = a_t. \tag{6.2.1}$$

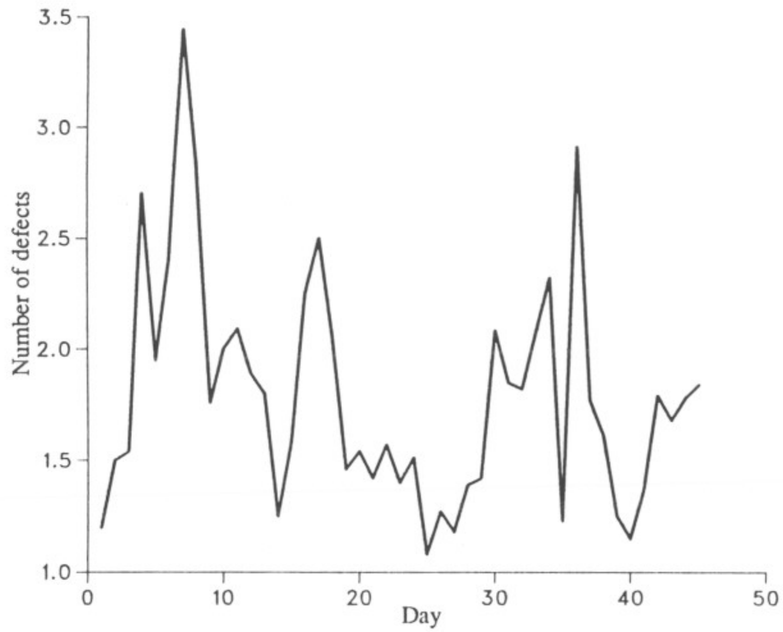

Fig. 6.1 Daily average number of truck manufacturing defects (Series W1).

Example 6.2 Series W2 is the classic series of the Wolf yearly sunspot numbers between 1700 and 1984, giving a total of $n = 285$ observations. Scientists believe that the sunspot numbers affect the weather of the earth and hence human activities such as agriculture, telecommunications, and others. The Wolf sunspot numbers have been discussed extensively in time series literature, e.g., Yule (1927), Bartlett (1950), Whittle (1954), Brillinger and Rosenblatt (1967), and others. This series is also known as the Wolfer sunspot numbers, named after Wolfer, who was a student of Wolf's. For an interesting account of the history of the series, see Izenman (1985). The series between 1700 and 1955 is

Table 6.2 Sample ACF and sample PACF for the daily average number of truck manufacturing defects (Series W1).

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.43	.26	.14	.08	−.09	−.07	−.21	−.11	−.05	−.01
St.E.	.15	.15	.17	.18	.19	.19	.19	.19	.19	.19
$\hat{\phi}_{kk}$	.43	.09	.00	.00	−.16	.00	−.18	.07	.05	.01
St.E.	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15

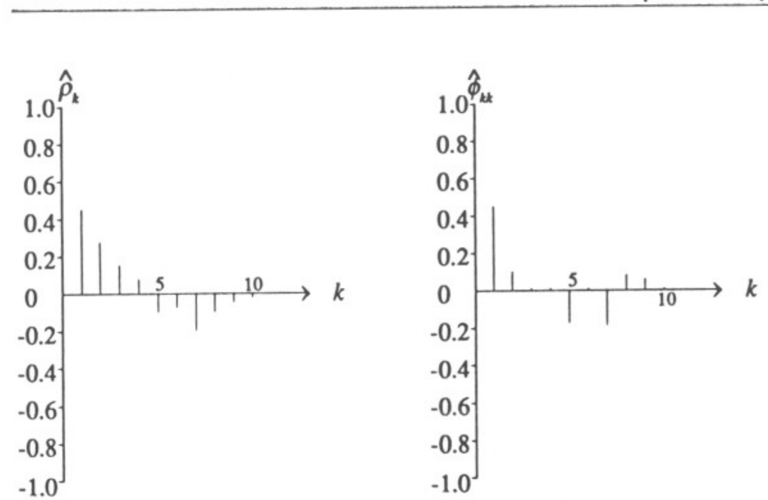

Fig. 6.2 Sample ACF and PACF of the daily average number of truck manufacturing defects (Series W1).

from Waldmeirer (1961), and the remaining observations are calculated from the monthly sunspot numbers in Andrews and Herzberg (1985).

The plot of the data is given in Figure 6.3. It indicates that the series is stationary in the mean. To investigate whether the series is also stationary in the variance, we calculated the following preliminary residual sum of squares:

$$S(\lambda) = \sum_{t=1}^{n} (Z_t(\lambda) - \hat{\mu})^2 \tag{6.2.2}$$

for various values of the power transformation parameter as shown in Table 6.3, where $\hat{\mu}$ is the corresponding sample mean of the transformed series. These calculations suggest that a square root transformation be applied to the data.

The sample ACF and sample PACF of the transformed data are computed as shown in Table 6.4 and Figure 6.4. The sample ACF shows a damping sine-cosine wave, and the sample PACF has relatively large spikes at lags 1, 2, and 9, suggesting that a tentative model may be an AR(2).

$$(1 - \phi_1 B - \phi_2 B^2)(\sqrt{Z_t} - \mu) = a_t, \tag{6.2.3}$$

or an AR(9)

$$(1 - \phi_1 B - \cdots - \phi_9 B^9)(\sqrt{Z_t} - \mu) = a_t. \tag{6.2.4}$$

By ignoring the values of $\hat{\phi}_{kk}$ beyond lag 3, Box and Jenkins (1976) suggest that an AR(3) model, $(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(\sqrt{Z_t} - \mu) = a_t$, is also possible, even though their analysis was based on the nontransformed data between 1770

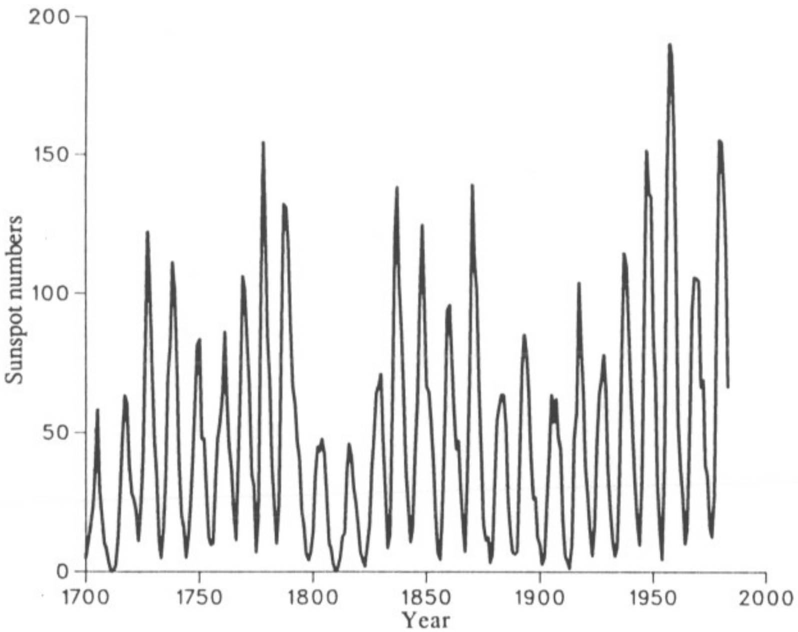

Fig. 6.3 Wolf yearly sunspot numbers, 1700–1983 (Series W2).

Table 6.3 Residual sum of squares in the power transformation.

λ	Residual sum of squares
1.0	13.82
0.5	9.77
0.0	11.06
−0.5	25.86
−1.0	147.68

and 1869. Because of the large autocorrelation .65 at lag 11, many scientists believe that the series contains a cycle of eleven years. We examine this series more carefully in later chapters.

Example 6.3 Series W3 is a laboratory series of blowfly data taken from Nicholson (1950). A fixed number of adult blowflies with balanced sex ratios were kept inside a cage and given a fixed amount of food daily. The blowfly population was then enumerated every other day for approximately two years,

Table 6.4 Sample ACF and sample PACF for the square root transformed sunspot numbers (Series W2).

k	$\hat{\rho}_k$									
1–10	.81	.45	.06	−.25	−.41	−.40	−.21	.08	.40	.61
St.E.	.06	.09	.10	.10	.10	.11	.11	.11	.11	.11
11–20	.65	.50	.22	−.06	−.28	−.37	−.34	−.21	−.00	.20
St.E.	.12	.14	.15	.15	.15	.15	.15	.16	.16	.16
21–30	.35	.37	.26	.05	−.15	−.30	−.37	−.32	−.20	−.03
St.E.	.16	.16	.16	.17	.17	.17	.17	.17	.17	.17
k	$\hat{\phi}_{kk}$									
1–10	.81	−.62	−.16	−.02	−.05	.13	.24	.16	.28	.03
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
11–20	−.01	−.05	−.06	.09	−.01	−.08	−.09	−.11	−.01	−.05
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
21–30	.05	−.03	−.10	−.08	.02	−.03	.01	.05	−.05	.01
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

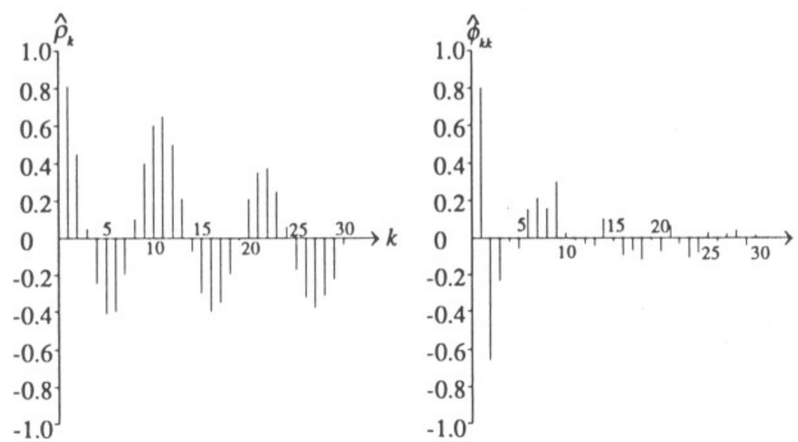

Fig. 6.4 Sample ACF and sample PACF for the square root transformed sunspot numbers (Series W2).

giving a total of $n = 364$ observations. Brillinger, Guckenheimer, Guttorp, and Oster (1980) first applied time series analysis on the data set. Later Tong (1983) considered the following two subseries:

Blowfly A: for Z_t between $20 \leq t \leq 145$,

Blowfly B: for Z_t between $218 \leq t \leq 299$,

and argued that the series Blowfly A is nonlinear. Series W3 used in our analysis is the series Blowfly B of 82 observations as shown in Figure 6.5.

The data plot suggests that the series is stationary in the mean. However, the power transformation analysis indicates that a square root or a logarithmic transformation is needed as shown in Table 6.5.

The sample ACF and PACF are calculated for the square root transformed data as shown in Table 6.6 and Figure 6.6. The sample $\hat{\rho}_k$ tails off exponentially and $\hat{\phi}_{kk}$ cuts off after lag 1. Thus, the following AR(1) model is entertained:

$$(1-\phi B)(\sqrt{Z_t}-\mu) = a_t \qquad (6.2.5)$$

Example 6.4 Recall the monthly series of 300 unemployed females between ages 16 and 19 in the United States from January 1961 to December 1985 as shown in Figure 4.1. The series is labeled as Series W4. It is clearly nonstationary in the mean, suggesting differencing is needed. This is further confirmed by the sustained large spikes of the sample ACF shown in Table 6.7 and Figure 6.7. The differenced series is now stationary and is shown in Figure 6.8. The sample ACF and sample PACF are computed for this differenced series in Table 6.8 and also plotted in Figure 6.9. The sample ACF now cuts off af-

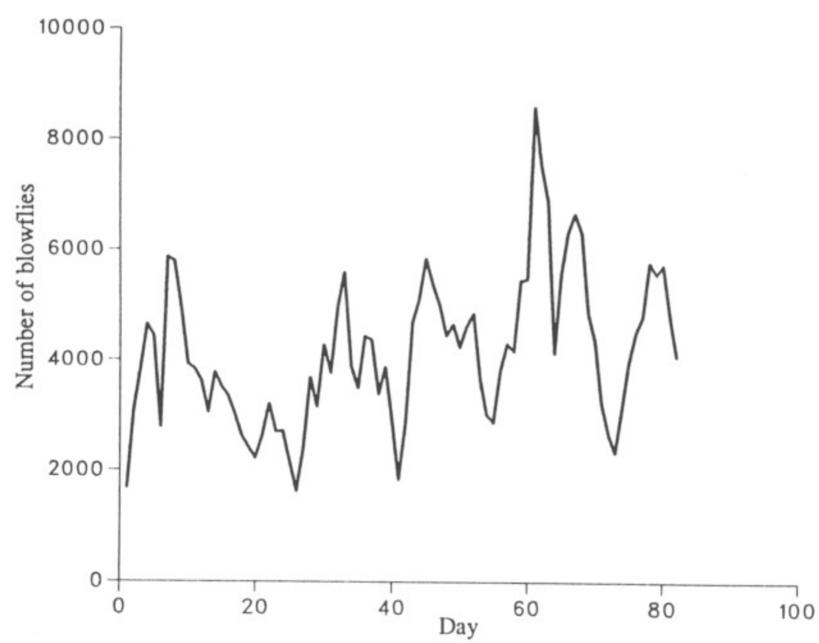

Fig. 6.5 Blowfly data (Series W3).

Table 6.5 Results of the power transformation on blowfly data.

λ	Residual sum of squares
1.0	55.31
0.5	50.09
0.0	50.38
−0.5	56.67
−1.0	71.50

Table 6.6 Sample ACF and sample PACF for the blowfly data (Series W3).

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.73	.49	.30	.20	.12	.02	−.01	−.04	−.01	−.03
St.E.	.11	.16	.18	.18	.19	.19	.19	.19	.19	.19
$\hat{\phi}_{kk}$	.73	−.09	−.04	.04	−.03	−.12	.07	−.05	.07	−.08
St.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

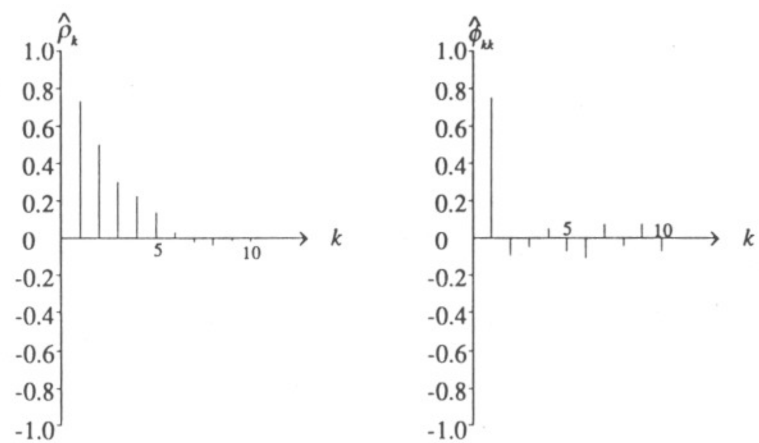

Fig. 6.6 Sample ACF and sample PACF for the blowfly data (Series W3).

Table 6.7 Sample ACF and sample PACF of Series W4 — the U.S. monthly series of unemployed females between ages 16 and 19 from January 1961 to December 1985.

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.97	.96	.95	.94	.93	.93	.92	.91	.90	.90
St.E.	.06	.10	.13	.15	.17	.18	.20	.21	.22	.24
$\hat{\phi}_{kk}$	.97	.37	.10	.08	.02	.08	.02	.02	−.04	.00
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

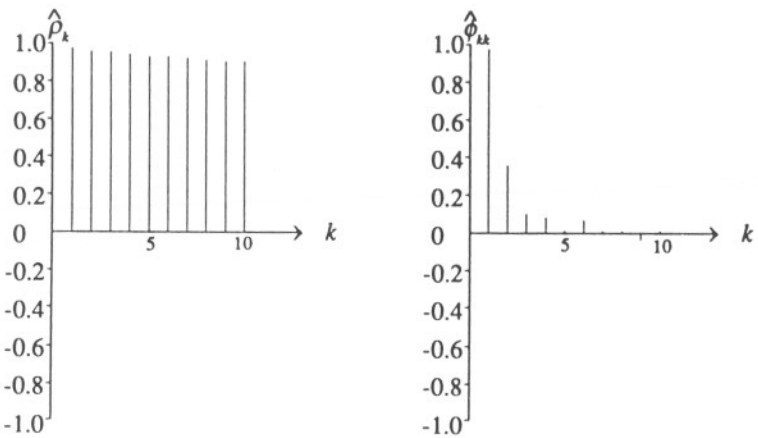

Fig. 6.7 Sample ACF and sample PACF of Series W4 — the U.S. monthly series of unemployed females between ages 16 and 19 from January 1961 to December 1985.

ter lag 1 while the sample PACF tails off. This pattern is very similar to the ACF and PACF for MA(1) with positive θ_1 as shown in Figure 3.10, suggesting an MA(1) model for the differenced series and hence an IMA(1, 1) model for the original series. To test whether a deterministic trend parameter θ_0 is needed, we calculate the t-ratio $\overline{W}/S_{\overline{W}} = 1.0502/2.4223 = .4335$, which is not significant. Thus, our proposed model is

$$(1-B)Z_t = (1-\theta_1 B)a_t. \tag{6.2.6}$$

Example 6.5 The accidental death rate is a vital statistic for many state and federal governments. Figure 6.10 shows Series W5, which is the yearly accidental death rate (per 100,000 population) of Pennsylvania between 1950 and 1984 published in the 1984 Pennsylvania Vital Statistics Annual Report by the Pennsylvania State Health Data Center. The series is clearly nonstation-

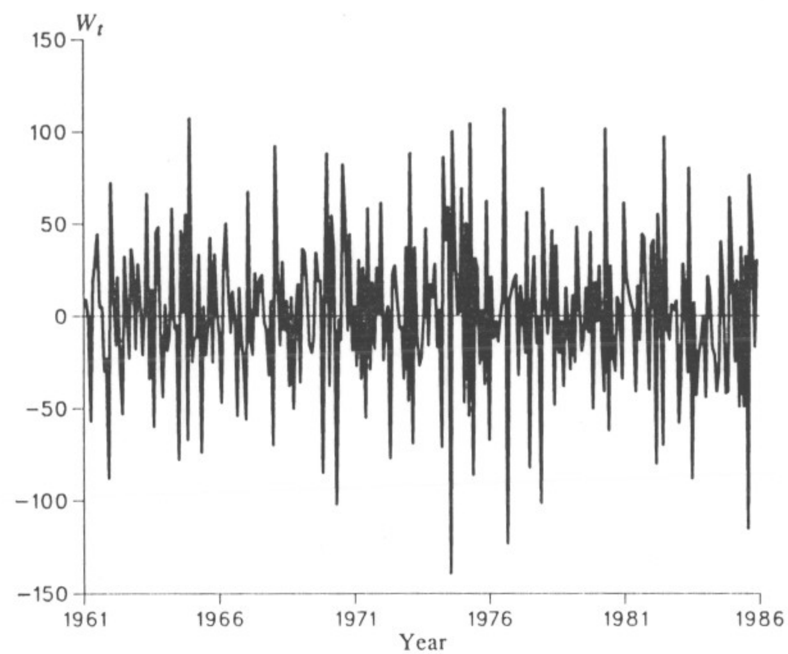

Fig. 6.8 The differenced monthly series, $W_t = (1-B)Z_t$, of the U.S. monthly series of unemployed females between ages 16 and 19 from January 1961 to December 1985 (Series W4).

ary with a decreasing trend. This nonstationarity is also shown by the slowly decaying ACF in Table 6.9 and Figure 6.11. Both Figure 6.11 and the evaluation of power transformations indicate no transformations other than differencing are needed. The sample ACF and PACF of the differenced data shown in Table 6.10 and Figure 6.12 suggest a white noise phenomenon. The t-ratio, $\overline{W}/S_{\overline{W}} = -.5618/.2507 = -2.24$, implies that a deterministic trend term is

Table 6.8 Sample ACF and sample PACF of the differenced U.S. monthly Series $W_t = (1-B)Z_t$, of unemployed females between ages 16 and 19 (Series W4) $\overline{W} = 1.0502$, $S_{\overline{W}} = 2.4223$.

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.41	.06	−.08	.06	−.09	.07	−.03	.07	−.06	−.08
St.E.	.06	.07	.07	.07	.07	.07	.07	.07	.07	.07
$\hat{\phi}_{kk}$	−.41	−.14	.14	−.04	−.11	−.02	−.02	.05	−.01	.06
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

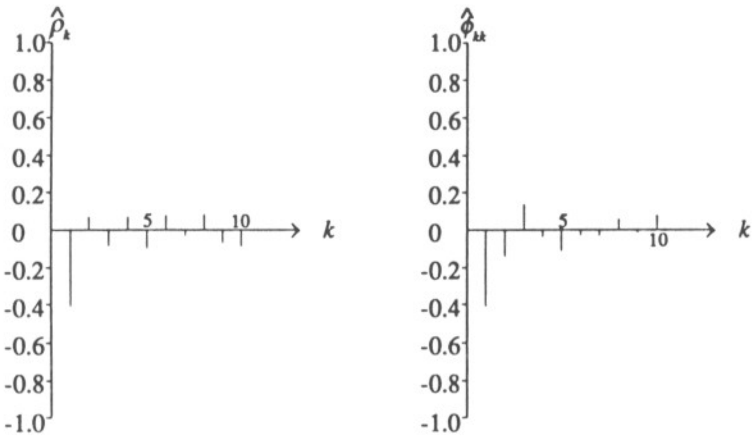

Fig. 6.9 Sample ACF and sample PACF of the differenced U.S. monthly series, $W_t = (1-B)Z_t$, of unemployed females between the ages of 16 and 19 (Series W4).

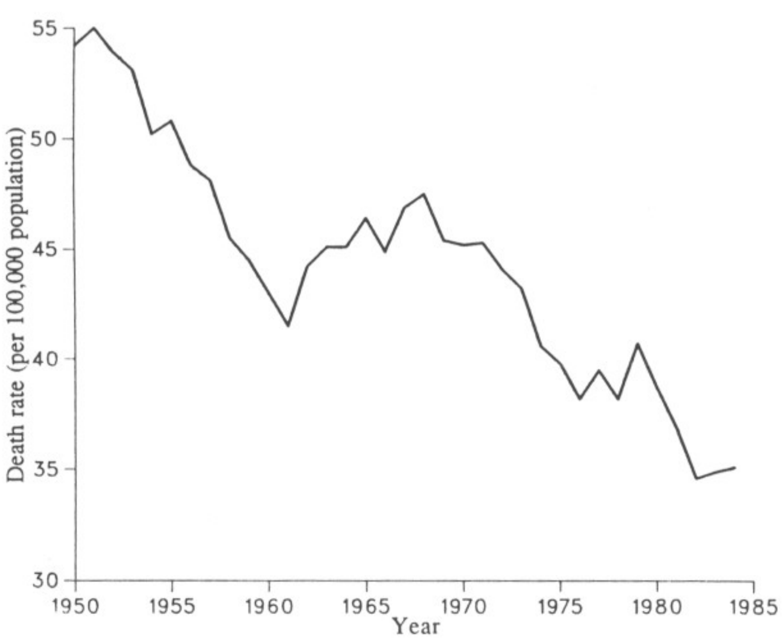

Fig. 6.10 Series W5—The yearly Pennsylvania accidental death rate between 1950 and 1984.

Table 6.9 Sample ACF and sample PACF of the Pennsylvania accidental death rate between 1950 and 1984 (Series W5).

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.87	.73	.57	.44	.34	.25	.15	.08	.01	−.02
St.E.	.17	.27	.32	.35	.36	.37	.38	.38	.38	.38
$\hat{\phi}_{kk}$	.87	−.13	−.15	.03	.06	−.08	−.14	.04	−.02	.03
St.E.	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17

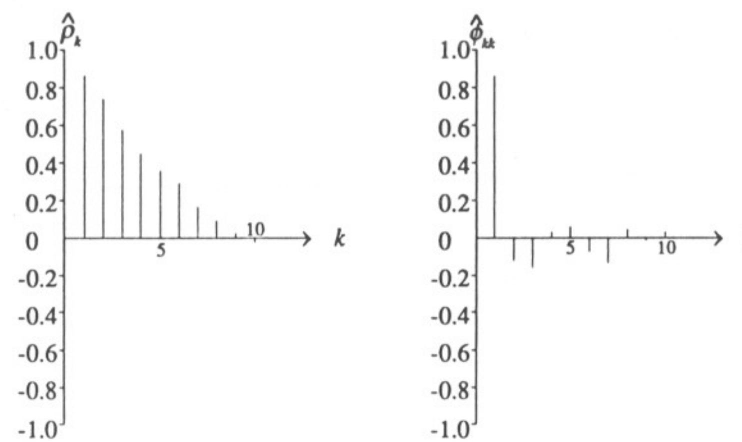

Fig. 6.11 Sample ACF and sample PACF of Series W5—the yearly Pennsylvania accidental death rate between 1950 and 1984.

recommended. Hence, the following random walk model with drift is entertained:

$$(1-B)Z_t = \theta_0 + a_t. \tag{6.2.7}$$

Based on the ACF and PACF of the original nondifferenced data in Table 6.9, one may also suggest an alternative AR(1) model

$$(1-\phi_1 B)(Z_t - \mu) = a_t. \tag{6.2.8}$$

However, the clear downward trend should give an estimate of ϕ_1 close to 1. We investigate both models in Chapter 7 when we discuss parameter estimation.

Example 6.6 We now examine Series W6, which is the yearly U.S. tobacco production from 1871 to 1984 published in the 1985 Agricultural Statistics by the United States Department of Revenue and shown in Figure 4.2. The plot indicates that the series is nonstationary both in the mean and the variance. In fact, the standard deviation of the series is roughly proportional to the level

Table 6.10 Sample ACF and sample PACF for the differenced series of the Pennsylvania accidental death rate from 1950 to 1984 (Series W5) $\overline{W} = -.5618$, $S_{\overline{W}} = .2507$.

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.10	.27	−.04	.09	.02	.05	−.08	.01	−.05	−.12
St.E.	.17	.17	.18	.18	.18	.18	.18	.19	.19	.19
$\hat{\phi}_{kk}$	−.10	.27	.01	.02	.05	.03	−.10	−.02	−.01	−.14
St.E.	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17

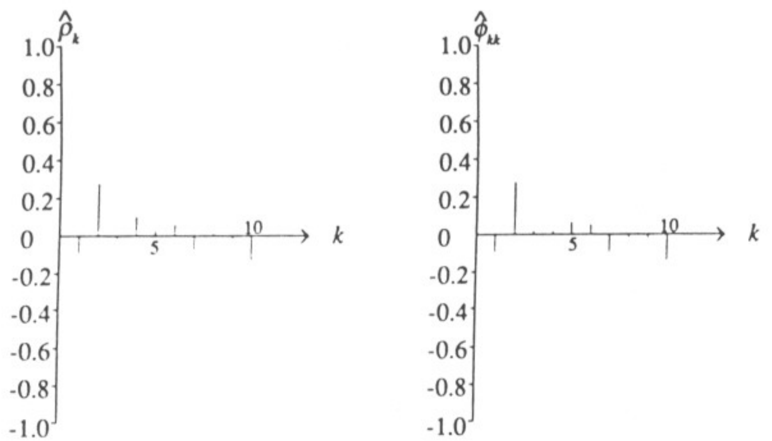

Fig. 6.12 Sample ACF and sample PACF for the differenced series of the Pennsylvania accidental death rate (Series W5).

of the series. Hence, from the results of Section 4.3.2, a logarithmic transformation is suggested, which is also confirmed by the λ value of the power transformation calculated in Table 6.11. These transformed data are plotted in Figure 6.13 and show an upward trend with a constant variance.

The very slowly decaying ACF as shown in Table 6.12 and Figure 6.14 further supports the need for differencing. Hence, the sample ACF and PACF for the differenced data, $W_t = (1-B)\ln Z_t$, are calculated in Table 6.13 with their plots in Figure 6.15. The ACF cuts off after lag 1, and the PACF tails off exponentially, which looks very similar to Figure 3.10 with $\theta_1 > 0$. It suggests that IMA(1, 1) is a possible model. To determine whether a deterministic trend

Table 6.11 Result of the power transformation on the tobacco production data.

λ	Residual sum of squares
1.0	7.88
0.5	5.95
0.0	5.11
−0.5	5.55
−1.0	7.92

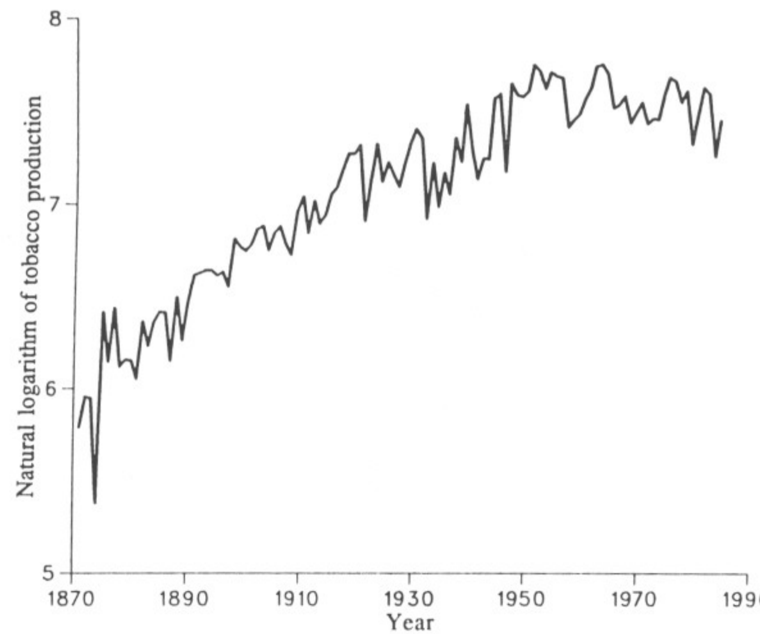

Fig. 6.13 Natural logarithms of the U.S. yearly tobacco production in million pounds (Series W6).

term θ_0 is needed, we examine the t-ratio, $t = \overline{W}/S_{\overline{W}} = .0147/.0186 = .7903$, which is not significant. Hence, we entertain the following IMA(1, 1) model as our tentative model:

$$(1-B)\ln Z_t = (1-\theta_1 B)a_t. \tag{6.2.9}$$

Table 6.12 Sample ACF and sample PACF for natural logarithms of the U.S. yearly tobacco production (Series W6).

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.90	.88	.84	.79	.78	.76	.75	.72	.69	.66
St.E.	.15	.19	.22	.24	.27	.28	.30	.32	.33	.34
$\hat{\phi}_{kk}$	.90	.37	.05	−.11	.15	.14	.08	−.11	−.12	.00
St.E.	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09

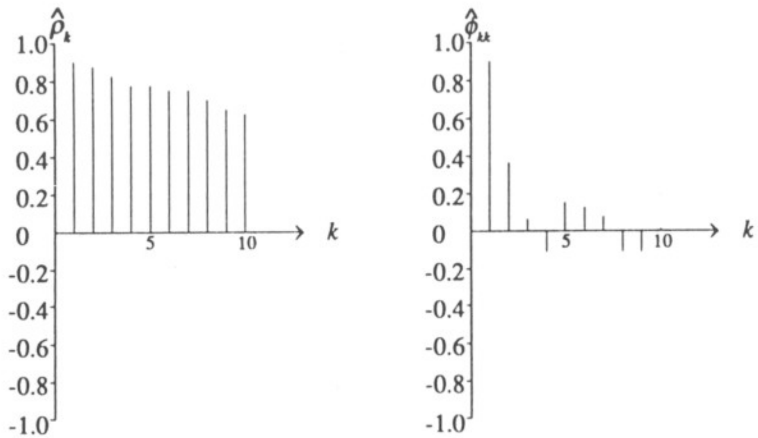

Fig. 6.14 Sample ACF and sample PACF for natural logarithms of the U.S. yearly tobacco production (Series W6).

Table 6.13 Sample ACF and sample PACF for the differenced series of natural logarithms of the U.S. yearly tobacco production (Series W6).

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.51	.11	−.09	.02	−.03	.00	.04	.04	−.05	−.01
St.E.	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12
$\hat{\phi}_{kk}$	−.51	−.20	−.17	−.14	−.13	−.12	−.04	.06	.02	−.03
St.E.	.09	.09	.09	.09	.09	.09	.09	.09	.09	.09

$\overline{W} = .0147$, $S_{\overline{W}} = .0186$

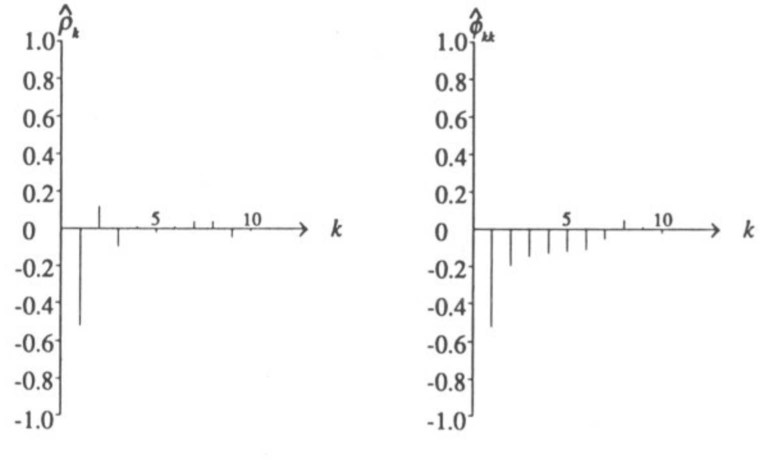

Fig. 6.15 Sample ACF and sample PACF for the differenced natural logarithms of the tobacco data (Series W6).

Example 6.7 Figure 6.16(a) shows Series W7—the yearly number of lynx pelts sold by the Hudson's Bay Company in Canada between 1857 and 1911 as reported in Andrews and Herzberg (1985). The result of the power transformation in Table 6.14 indicates that a logarithmic transformation is required. The natural logarithm of the series is stationary and is plotted in Figure 6.16(b).

The sample ACF in Table 6.15 and Figure 6.17 show a clear sine-cosine phenomenon indicating an AR(p) model with $p \geq 2$. The three significant PACF strongly suggest $p = 3$. Thus, our entertained model is

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(\ln Z_t - \mu) = a_t. \quad (6.2.10)$$

A related series that was studied by many time series analysts is the number of Canadian lynx trapped for the years from 1821 to 1934. References include Campbell and Walker (1977), Tong (1977), Priestly (1981, Section 5.5), and Lin (1987). The series of lynx pelt sales that we analyze here is much shorter and has received much less attention in the literature. We use this series extensively for various illustrations in this book.

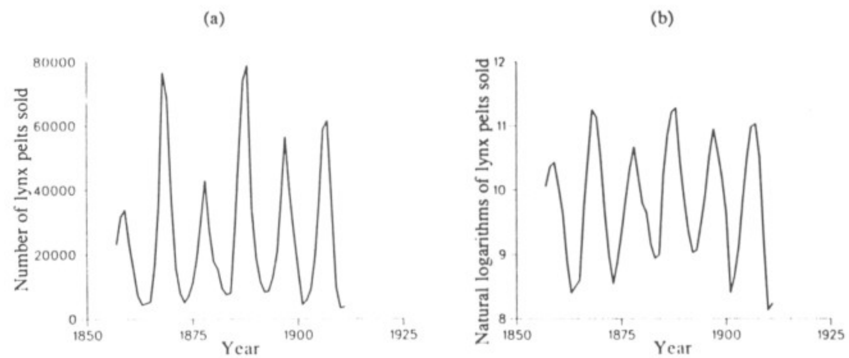

Fig. 6.16 The Canadian lynx pelt sales data (Series W7). (a) The yearly number of lynx pelts sold in Canada between 1857 and 1911. (b) The natural logarithms of yearly numbers of lynx pelts sold in Canada between 1857 and 1911.

Table 6.14 Result of the power transformation on the lynx pelt sales.

λ	Residual sum of squares
1.0	107.1642
0.5	61.1585
0.0	51.6603
−0.5	68.1996
−1.0	129.2240

Table 6.15 Sample ACF and PACF for natural logarithms of the yearly number of lynx pelts sold (Series W7).

(a) ACF, $\hat{\rho}_k$

1–10	.73	.22	−.32	−.69	−.76	−.53	−.08	.35	.61	.59
St.E.	.13	.19	.20	.21	.25	.29	.30	.30	.31	.33
11–20	.31	−.06	−.41	−.58	−.49	−.21	.16	.44	.54	.40
St.E.	.35	.36	.37	.37	.39	.40	.40	.41	.41	.43

(b) PACF, $\hat{\phi}_{kk}$

1–10	.73	−.68	−.36	−.20	−.09	−.08	.13	−.08	.06	−.07
St.E.	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13
11–20	−.13	.05	−.19	.07	−.02	−.04	.14	−.04	.10	−.09
St.E.	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13

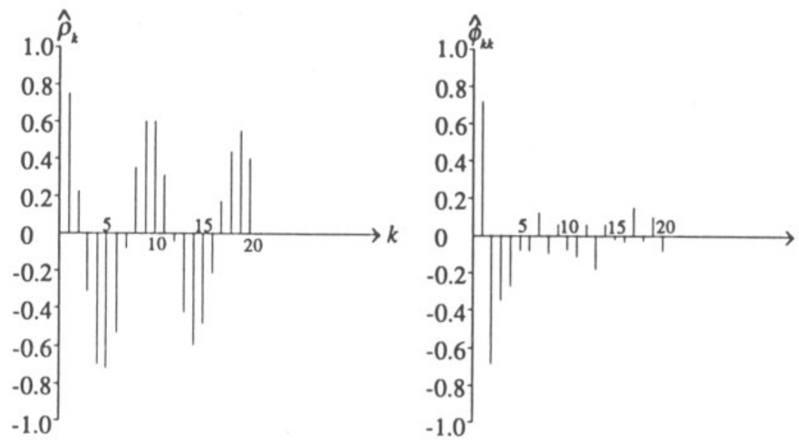

Fig. 6.17 Sample ACF and PACF for natural logarithms of the yearly number of lynx pelts sold (Series W7).

6.3 INVERSE AUTOCORRELATION FUNCTION (IACF)

Let

$$\phi_p(B)(Z_t - \mu) = \theta_q(B)a_t \tag{6.3.1}$$

be an ARMA(p, q) model where $\phi_p(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ is a stationary autoregressive operator, $\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$ is an invertible moving average operator, and the a_t is a white noise series with a zero mean and a constant variance σ_a^2. We can rewrite Equation (6.3.1) as the moving average representation

$$(Z_t - \mu) = \frac{\theta_q(B)}{\phi_p(B)}a_t = \psi(B)a_t, \tag{6.3.2}$$

where $\psi(B) = \theta_q(B)/\phi_p(B)$. From (2.6.9), the autocovariance generating function of this model is given by

$$\begin{aligned}\gamma(B) &= \sum_{k=-\infty}^{\infty} \gamma_k B^k = \sigma_a^2 \psi(B)\psi(B^{-1}) \\ &= \sigma_a^2 \frac{\theta_q(B)\theta_q(B^{-1})}{\phi_p(B)\phi_p(B^{-1})}.\end{aligned} \tag{6.3.3}$$

Assume that $|\gamma(B)| > 0$ for all $|B| \le 1$ and let

$$\gamma^{(I)}(B) = \frac{1}{\gamma(B)} = \sum_{k=-\infty}^{\infty} \gamma_k^{(I)} B^k$$
$$= \frac{1}{\sigma_a^2} \frac{\phi_p(B)\phi_p(B^{-1})}{\theta_q(B)\theta_q(B^{-1})}. \tag{6.3.4}$$

Clearly, from the assumptions on $\phi_p(B)$ and $\theta_q(B)$, the coefficients $\{\gamma_k^{(I)}\}$ in (6.3.4) form a positive definite, square summable sequence that is symmetric about the origin. Hence, they are proper autocovariances for a process. With respect to the model given in (6.3.1), the process having $\gamma^{(I)}(B)$ as its autocovariance generating function is referred to as the inverse process. Hence, $\gamma^{(I)}(B)$ is also referred to as the inverse autocovariance generating function of $\{Z_t\}$. Naturally, the inverse autocorrelation generating function of $\{Z_t\}$, from (2.6.10), is given by

$$\rho^{(I)}(B) = \frac{\gamma^{(I)}(B)}{\gamma_0^{(I)}}. \tag{6.3.5}$$

The kth lag inverse autocorrelation is defined as

$$\rho_k^{(I)} = \frac{\gamma_k^{(I)}}{\gamma_0^{(I)}}, \tag{6.3.6}$$

which, of course, is equal to the coefficient of B^k or B^{-k} in $\rho^{(I)}(B)$. As a function of k, $\rho_k^{(I)}$ is called the inverse autocorrelation function (IACF).

From Equations (6.3.3) and (6.3.4), it is clear that if $\{Z_t\}$ is an ARMA (p, q) process, then the inverse process will be an ARMA(q, p) process. Specifically, if $\{Z_t\}$ is an AR(p) process with autocorrelations tailing off, then the inverse process will be an MA(p) process with its autocorrelations cutting off at lag p. In other words, the inverse autocorrelation function of an AR(p) process will cut off at lag p. Similarly, an MA(q) process with autocorrelations cutting off at lag q will have its inverse autocorrelations tailing off. Hence, the inverse autocorrelation function of a process exhibits characteristics similar to the partial autocorrelation function and can be used as an identification tool in model identification.

For an AR(p) process, it is easily seen that the inverse autocorrelation function is given by

$$\rho_k^{(I)} = \begin{cases} \dfrac{-\phi_k + \phi_1\phi_{k+1} + \cdots + \phi_{p-k}\phi_p}{1 + \phi_1^2 + \cdots + \phi_p^2}, & k = 1, \ldots, p, \\ 0, & k > p. \end{cases} \tag{6.3.7}$$

Since one can approximate a series by an AR(p) model with a sufficiently large p, one method to obtain a sample inverse autocorrelation $\hat{\rho}_k^{(I)}$ is to replace the AR parameters by their estimates, i.e.,

$$\hat{\rho}_k^{(I)} = \begin{cases} \dfrac{-\hat{\phi}_k + \hat{\phi}_1\hat{\phi}_{k+1} + \cdots + \hat{\phi}_{p-k}\hat{\phi}_p}{1 + \hat{\phi}_1^2 + \cdots + \hat{\phi}_p^2}, & k = 1, \ldots, p, \\ 0, & k > p. \end{cases} \tag{6.3.8}$$

The parameter estimation is discussed in Chapter 7. Under the null hypothesis of a white noise process, the standard error of $\hat{\rho}_k^{(I)}$ is given by

$$S_{\hat{\rho}_k^{(I)}} = \sqrt{\frac{1}{n}}. \tag{6.3.9}$$

Thus, one can use the limits $\pm 2/\sqrt{n}$ to assess the significance of the sample inverse autocorrelations.

Example 6.8 As illustrations, Table 6.16 shows sample inverse autocorrelations for two time series examined in Section 6.2. Table 6.16(a) is the sample inverse autocorrelation for the daily average series of truck manufacturing defects that we discussed in Example 6.1 in which we entertained an AR(1) model for the series. Note that although the inverse autocorrelation function seems to cut off at lag 1, it is not statistically significant at $\alpha = .05$. Table 6.16(b) is the sample inverse autocorrelation for the natural logarithms of the Canadian lynx pelt sales that was identified as an AR(3) model based on the three significant PACF examined in Example 6.7. However, in terms of the IACF, the maximum AR order will be 2. In fact, because the standard error of the sample IACF is .14, the model implied by the IACF may be an AR(1), which is clearly not appropriate from the ACF in Table 6.15(a).

It is not unusual that the values of the IACF are smaller than those of PACF at lower lags in the above examples. If the underlying model is an AR(p), we

Table 6.16 Sample inverse autocorrelation functions (SIACF).

(a) The daily average number of truck manufacturing defects (Series W1)

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k^{(I)}$	−.27	.05	−.03	−.07	.10	−.11	.22	.01	−.05	.02
St.E.	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15

(b) The natural logarithms of Canadian lynx pelt sales (Series W7)

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k^{(I)}$	−.67	.23	.01	−.03	.09	−.11	.15	−.11	.07	−.13
St.E.	.14	.14	.14	.14	.14	.14	.14	.14	.14	.14

know that $\phi_{kk} = \rho_k^{(I)} = 0$ for $k > p$. But, Equation (6.3.7) implies that

$$\rho_p^{(I)} = \frac{-\phi_p}{1 + \phi_1^2 + \cdots + \phi_p^2}, \tag{6.3.10}$$

and from the discussion in the closing paragraph of Section 2.3, we have

$$\phi_{pp} = \phi_p. \tag{6.3.11}$$

Hence, $|\phi_{pp}| > |\rho_p^{(I)}|$, and sample IACF in general are smaller than sample PACF, particularly at lower lags. In a recent study, Abraham and Ledolter (1984) conclude that, as an identification aid, the PACF generally outperforms the IACF. Some computer programs such as SAS and SCA provide both PACF and IACF options for analysis.

The inverse autocorrelation function was first introduced by Cleveland (1972) through the inverse of a spectrum and the relationship between the spectrum and the autocorrelation function. This leads to another method to obtain a sample inverse autocorrelation function. We return to this point in Chapter 11.

6.4 EXTENDED SAMPLE AUTOCORRELATION FUNCTION AND OTHER IDENTIFICATION PROCEDURES

6.4.1 Extended Sample Autocorrelation Function (ESACF)

From the previous empirical examples, it seems clear that, due to the cutting off property of the PACF and IACF for AR models and the same cutting off property of the ACF for MA models, the identification of the order p of an AR model and the order q of an MA model through the sample ACF, PACF, and IACF are relatively simple. However, for a mixed ARMA process, the ACF, PACF, and IACF all exhibit tapering off behavior, which makes the identification of the orders p and q much more difficult. One commonly used method is based on the fact that if Z_t follows an ARMA(p, q) model

$$(1 - \phi_1 B - \cdots - \phi_p B^p) Z_t = \theta_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) a_t, \tag{6.4.1a}$$

or equivalently

$$Z_t = \theta_0 + \sum_{i=1}^{p} \phi_i Z_{t-1} - \sum_{i=1}^{q} \theta_i a_{t-1} + a_t, \tag{6.4.1b}$$

then

$$\begin{aligned} Y_t &= (1 - \phi_1 B - \cdots - \phi_p B^p) Z_t \\ &= Z_t - \sum_{i=1}^{p} \phi_i Z_{t-i} \end{aligned} \tag{6.4.2}$$

follows an MA(q) model

$$Y_t = (1 - \theta_1 B - \cdots - \theta_q B^q) a_t \tag{6.4.3}$$

where without loss of generality we assume $\theta_0 = 0$. Thus, some authors such as Tiao and Box (1981) suggest using the sample ACF of the estimated residuals

$$\hat{Y}_t = (1 - \hat{\phi}_1 B - \cdots - \hat{\phi}_p B^p) Z_t \tag{6.4.4}$$

from an ordinary least squares (OLS) AR fitting to identify q and hence obtaining the orders p and q for the ARMA(p, q) model. For example, an MA(2) residual process from an AR(1) fitting implies a mixed ARMA(1, 2) model. But, as we show in Section 7.4, because the OLS estimates of the AR parameter ϕ_is in (6.4.4) are not consistent when the underlying model is a mixed ARMA(p, q) with $q > 0$, the procedure may lead to incorrect identification.

To derive consistent estimates of ϕ_i, suppose that n observations adjusted for mean are available from the ARMA(p, q) process in (6.4.1a). If an AR(p) is fitted to the data, i.e.,

$$Z_t = \sum_{i=1}^{p} \phi_i Z_{t-i} + e_t, \qquad t = p+1, \ldots, n \tag{6.4.5}$$

where e_t represents the error term, then the OLS estimates $\hat{\phi}_i^{(0)}$ of ϕ_i, $i = 1, \ldots, p$, will be inconsistent and the estimated residuals

$$\hat{e}_t^{(0)} = Z_t - \sum_{i=1}^{p} \hat{\phi}_i^{(0)} Z_{t-i}$$

will not be white noise. In fact, if $q \geq 1$, the lagged values $\hat{e}_{t-i}^{(0)}$, $i = 1, \ldots, q$, will contain some information about the process Z_t. This leads to the following iterated regressions. First, we consider an AR(p) regression plus an added term $\hat{e}_{t-1}^{(0)}$, i.e.,

$$Z_t = \sum_{i=1}^{p} \phi_i^{(1)} Z_{t-i} + \beta_1^{(1)} \hat{e}_{t-1}^{(0)} + e_t^{(1)}, \qquad t = p+2, \ldots, n \tag{6.4.6}$$

where the superscript (1) refers to the first iterated regression and $e_t^{(1)}$ represents the corresponding error term. The OLS estimates $\hat{\phi}_i^{(1)}$ will be consistent if $q = 1$. However, if $q > 1$, the $\hat{\phi}_i^{(1)}$ will again be inconsistent, the estimated residuals $\hat{e}_t^{(1)}$ are not white noise, and the lagged values $\hat{e}_{t-i}^{(1)}$ will contain some information about Z_t. We thus consider the second iterated AR(p) regression

$$Z_t = \sum_{i=1}^{p} \phi_i^{(2)} Z_{t-i} + \beta_1^{(2)} \hat{e}_{t-1}^{(1)} + \beta_2^{(2)} \hat{e}_{t-2}^{(0)} + e_t^{(2)}, \qquad t = p+3, \ldots, n. \quad (6.4.7)$$

The OLS estimates $\hat{\phi}_i^{(2)}$ will be consistent if $q = 2$. For $q > 2$, the $\hat{\phi}_i^{(2)}$ will again be inconsistent. However, consistent estimates can be obtained by repeating the above iteration. That is, the OLS estimates $\hat{\phi}_i^{(q)}$ obtained from the following qth iterated AR(p) regression will be consistent:

$$Z_t = \sum_{i=1}^{p} \phi_i^{(q)} Z_{t-i} + \sum_{i=1}^{q} \beta_i^{(q)} \hat{e}_{t-i}^{(q-i)} + e_t^{(q)}, \qquad t = p+q+1, \ldots, n \quad (6.4.8)$$

where $\hat{e}_t^{(j)} = Z_t - \sum_{i=1}^{p} \hat{\phi}_i^{(j)} Z_{t-i} - \sum_{i=1}^{q} \hat{\beta}_i^{(j)} \hat{e}_{t-i}^{(j-i)}$ is the estimated residual of the jth iterated AR(p) regression and the $\hat{\phi}_i^{(j)}$ and $\hat{\beta}_i^{(j)}$ are the corresponding least square estimates.

In practice, the true order p and q of the ARMA(p, q) model are usually unknown and have to be estimated. However, based on the preceding consideration, Tsay and Tiao (1984) suggest a general set of iterated regressions and introduce the concept of the extended sample autocorrelation function (ESACF) to estimate the orders p and q. Specifically, for $m = 0, 1, 2, \ldots$, let $\hat{\phi}_i^{(j)}, i = 1, \ldots, m$, be the OLS estimates from the jth iterated AR(m) regression of the ARMA process Z_t. They define the mth ESACF $\hat{\rho}_j^{(m)}$ of Z_t as the sample autocorrelation function for the transformed series

$$Y_t^{(j)} = (1 - \hat{\phi}_1^{(j)} B - \cdots - \hat{\phi}_m^{(j)} B^m) Z_t. \quad (6.4.9)$$

It is useful to arrange $\hat{\rho}_j^{(m)}$ in a two-way table as shown in Table 6.17 where the first row corresponding to $\hat{\rho}_j^{(0)}$ gives the standard sample ACF of Z_t, the second row gives the first ESACF $\hat{\rho}_j^{(1)}$, and so on. The rows are numbered 0, 1, ... to specify the AR order, and the columns are numbered in a similar way for the MA order.

Table 6.17 The ESACF Table.

AR	MA 0	1	2	3	4	...
0	$\hat{\rho}_1^{(0)}$	$\hat{\rho}_2^{(0)}$	$\hat{\rho}_3^{(0)}$	$\hat{\rho}_4^{(0)}$	$\hat{\rho}_5^{(0)}$	...
1	$\hat{\rho}_1^{(1)}$	$\hat{\rho}_2^{(1)}$	$\hat{\rho}_3^{(1)}$	$\hat{\rho}_4^{(1)}$	$\hat{\rho}_5^{(1)}$	...
2	$\hat{\rho}_1^{(2)}$	$\hat{\rho}_2^{(2)}$	$\hat{\rho}_3^{(2)}$	$\hat{\rho}_4^{(2)}$	$\hat{\rho}_5^{(2)}$	...
3	$\hat{\rho}_1^{(3)}$	$\hat{\rho}_2^{(3)}$	$\hat{\rho}_3^{(3)}$	$\hat{\rho}_4^{(3)}$	$\hat{\rho}_5^{(3)}$	...
⋮			⋮			

Table 6.18 The asymptotic ESACF for an ARMA(1, 1) model.

AR	MA 0	1	2	3	4	...
0	X	X	X	X	X	...
1	X	0	0	0	0	...
2	X	X	0	0	0	...
3	X	X	X	0	0	...
4	X	X	X	X	0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋱

Note that the ESACF $\hat{\rho}_j^{(m)}$ is a function of n, the number of observations, even though it is not explicitly shown. In fact, it can be shown (see Tsay and Tiao [1984]) that for an ARMA(p, q) process, we have the following convergence in probability, i.e., for $m = 1, 2, \ldots$, and $j = 1, 2, \ldots$, we have

$$\hat{\rho}_j^{(m)} \xrightarrow{P} \begin{cases} 0, & 0 \le m - p < j - q, \\ X \ne 0, & \text{otherwise.} \end{cases} \quad (6.4.10)$$

Thus, by (6.4.10), the asymptotic ESACF table for an ARMA(1, 1) model becomes the one shown in Table 6.18. The zeros can be seen to form a triangle with the vertex at the (1, 1) position. More generally, for an ARMA(p, q) process, the vertex of the zero triangle in the asymptotic ESACF will be at the (p, q) position. Hence, the ESACF can be a useful tool in model identification, particularly for a mixed ARMA model.

Of course, in practice, we have finite samples, and the limit of $\hat{\rho}_j^{(m)}$ for $0 \le m - p < j - q$ may not be exactly zero. However, the asymptotic variance of $\hat{\rho}_j^{(m)}$ can be approximated using Bartlett's formula or more crudely by $(n - m - j)^{-1}$ on the hypothesis that the transformed series $Y_t^{(j)}$ of (6.4.9) is white noise. The ESACF table can then be constructed using indicator symbols with X referring to values greater than or less than ± 2 standard deviations and 0 for values within ± 2 standard deviations.

Example 6.9 To illustrate the method, we use SCA to compute the ESACF for the natural logarithms of Canadian lynx pelt sales discussed in Example 6.7. Table 16.19(a) shows the ESACF and Table 16.19(b) the corresponding indicator symbols for the ESACF of the series. The vertex of the triangle suggests a mixed ARMA(2, 2) model. This is different from an AR(3) model that we en-

Table 6.19 (a) The ESACF for natural logarithms of Canadian lynx pelt sales.

AR	MA 0	1	2	3	4	5	6	7	8	9
	(a) The ESACF									
0	.73	.22	−.32	−.69	−.76	−.53	−.08	.33	.61	.59
1	.68	.22	−.28	−.60	−.65	.51	−.09	.31	.53	.57
2	−.54	.01	.20	−.15	.17	−.26	.25	−.15	−.05	.10
3	−.53	.06	.32	.09	−.02	−.10	.15	−.16	−.20	.06
4	.00	.13	.37	−.12	−.02	−.09	−.01	−.18	−.16	−.03
	(b) The ESACF with indicator symbols									
0	X	0	X	X	X	X	0	X	X	X
1	X	0	0	X	X	X	0	X	X	X
2	X	0	0	0	0	0	0	0	0	0
3	X	0	X	0	0	0	0	0	0	0
4	0	X	X	0	0	0	0	0	0	0

tertained earlier based on the characteristic of the PACF. We further examine the series in Chapter 7.

Example 6.10 As another example, Table 6.20 shows the ESACF for Series C of 226 temperature readings in Box and Jenkins (1976). Because the vertex of the zero triangle occurs at the (2, 0) position, the ESACF approach suggests an AR(2) model for this series.

It should be noted that the OLS estimation is used in the iterated regression. Because the OLS regression estimation can always be applied to a regression model regardless of whether a series is stationary or nonstationary, invertible or noninvertible, Tsay and Tiao (1984) allow the roots of AR and MA polynomials to be on or outside the unit circle in their definition of the ESACF.

Table 6.20 The ESACF of Series C in Box and Jenkins.

AR	MA 0	1	2	3	4	5	6	7	8
0	X	X	X	X	X	X	X	X	X
1	X	X	X	X	X	X	X	X	X
2	0	0	0	0	0	0	0	0	0
3	X	0	0	0	0	0	0	0	0
4	X	X	0	0	0	0	0	0	0
5	X	X	X	0	0	0	0	0	0

Hence, the ESACF is computed from the original nondifferenced data. As a result, the ARIMA(p, d, q) process is identified as the ARMA(P, q) process with $P = p + d$. For example, based on the sample ACF and PACF of Series C as given in Table 6.21, Box and Jenkins (1976) suggested an ARIMA(1, 1, 0) model or an ARIMA(0, 2, 0) model. However, both these ARIMA(1, 1, 0) and ARIMA(0, 2, 0) models are identified as an ARMA(2, 0) model in terms of the ESACF.

Because the ESACF proposed by Tsay and Tiao (1984) is defined on the original series, they argue that the use of ESACF eliminates the need for differencing and provides a unified identification procedure to both stationary and nonstationary processes. To see whether a series is nonstationary, they suggest that for given specified values of p and q, the iterated AR estimates can be examined to see whether the AR polynomial contains a nonstationary factor with roots on the unit circle. For example, in the above Series C in Box and Jenkins (1976), the iterated AR estimates are $\hat{\phi}_1^{(0)} = 1.81$ and $\hat{\phi}_2^{(0)} = -.82$. Because $\hat{\phi}^{(0)}(B) \simeq (1-B)(1-.8B)$, it indicates that the AR polynomial contains the factor $(1-B)$. However, other than a few nice exceptions, the task of identifying nonstationarity through this approach is generally difficult. The real advantage of the ESACF is for the identification of p and q for the mixed ARMA mod-

Table 6.21 Sample ACF and PACF of Series C in Box and Jenkins (1976).

	(a) Z_t									
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.98	.94	.90	.85	.80	.75	.69	.64	.58	.52
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
$\hat{\phi}_{kk}$	.98	−.81	−.03	−.02	−.10	−.07	−.01	−.03	.04	−.04
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
	(b) $(1-B)Z_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.80	.65	.53	.44	.38	.32	.26	.19	.14	.14
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
$\hat{\phi}_{kk}$	.80	−.01	−.01	.06	.03	−.03	−.01	−.08	.00	.10
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
	(c) $(1-B)^2Z_t$									
k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.08	−.07	−.12	−.06	.01	−.02	.05	−.05	−.12	.12
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
$\hat{\phi}_{kk}$	−.08	−.08	−.14	−.10	−.03	−.05	.02	−.06	−.16	.09
St.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07

els. This advantage, I think, can be much better utilized if the ESACF is used for a properly transformed stationary series. This is particularly true because a tentatively identified model will be subjected to more efficient estimation procedures (such as the maximum likelihood estimation), which generally require stationarity.

Due to sampling variations and correlations among sample ACF, the pattern in the ESACF table from most time series may not be as clear-cut as those shown in the above examples. But from the author's experience, models can usually be identified without much difficulty through a joint study of ACF, PACF, and ESACF.

Some computer programs such as AUTOBOX and SCA provide the option to compute the ESACF in the model identification phase.

6.4.2 Other Identification Procedures

Other model identification procedures include the information criterion (AIC) proposed by Akaike (1974); the R-and-S-array introduced by Gray, Kelley, and McIntire (1978); and the corner method suggested by Beguin, Gourieroux, and Monfort (1980). The statistical properties of the statistics used in the R-and-S-array approach and the corner method are still largely unknown, and the software needed for these methods is not easily available. Interested readers are referred to their original research papers listed in the reference section of this book. The information criterion is discussed later in Chapter 7.

At this point, it is appropriate to say that model identification is both a science and an art. One should not use one method to the exclusion of others. Through careful examination of the ACF, PACF, IACF, ESACF, and other properties of time series, model identification truly becomes the most interesting aspect of time series analysis.

Exercises

6.1 Identify appropriate time series models from the sample ACF below. Justify your choice using the knowledge of the theoretical ACF for ARIMA models.

(a) $n = 121$, data $= Z_t$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.15	−.08	.04	.08	.08	.03	.02	.05	.04	−.11

(b) $n = 250$, data $= Z_t$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.63	.36	−.17	.09	−.07	.06	−.08	.10	−.11	.06

(c) $n = 250$, data $= Z_t$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	−.35	−.17	.09	−.06	.01	−.01	−.04	.07	−.07	.09

(d) $n = 100$, data $= Z_t$, $W_t = (1-B)Z_t$, $\overline{W} = 2.5$, $S_W^2 = 20$.

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_Z(k)$	.99	.98	.98	.97	.94	.91	.89	.86	.85	.83
$\hat{\rho}_W(k)$	.45	−.04	.12	.06	−.18	.16	−.07	.05	.10	.09

(e) $n = 100$, data $= Z_t$, $W_t = (1-B)Z_t$, $\overline{W} = 35$, $S_W^2 = 15$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_Z(k)$	.94	.93	.90	.89	.87	.86	.84	.81	.80	.80
$\hat{\rho}_W(k)$	.69	.50	.33	.19	.10	.08	.03	.01	.01	.00

6.2 Identify proper models for the following data sets (read across):

(a)

−2.401	−.574	.382	−.535	−1.639	−.960	−1.118
−.719	−1.236	.117	−.493	−2.282	−1.823	.645
−.179	.589	1.413	.370	.082	−.531	−1.891
−.961	−.865	−.790	−1.476	−2.491	−4.479	−2.809
−2.154	−1.532	−2.119	−3.349	−1.588	.740	.907
1.540	.557	2.259	2.622	.701	2.463	2.714
2.089	3.750	4.322	3.186	3.192	2.939	3.263
3.279	.295	.227	1.356	1.912	1.060	.370
−.195	.340	1.084	1.237	.610	2.126	3.960
3.317	2.167	1.292	.595	.140	−.082	−.769
.870	1.551	2.610	2.193	1.353	−.600	−.455
.203	1.472	1.367	1.875	2.082	1.604	2.033
3.746	2.954	.676	1.163	1.368	.343	−.334
1.041	1.328	1.325	.968	1.970	2.296	2.896
1.918	1.569					

(b)

−1.453	.867	.727	−.765	−1.317	.024	−.542
−.048	−.805	.858	−.563	−1.986	−.454	1.738
−.566	.697	1.060	−.478	−.140	−.581	−1.572
.174	−.289	−.270	−1.002	−1.605	−2.984	−.122
.469	−.239	−1.200	−2.077	.421	1.693	.463
.996	−.367	1.925	1.267	−.872	2.043	1.236
.461	2.497	2.072	.593	1.281	1.023	1.500
1.321	−1.673	.050	1.219	1.098	−.087	−.266
−.417	.457	.880	.586	−.132	1.760	2.684
.941	.177	−.008	−.180	−.217	−.165	−.720
1.332	1.029	1.679	.627	.038	−1.412	−.095
.476	1.350	.484	1.055	.957	.355	1.071
2.526	.707	−1.096	.757	.670	−.477	−.540
1.241	.704	.528	.173	1.389	1.115	1.519
.180	.419					

(c)

3.485	5.741	5.505	3.991	3.453	4.773	4.142
4.598	3.796	5.430	3.960	2.541	4.054	6.155
3.778	5.066	5.422	3.908	4.302	3.876	2.888
4.613	4.075	4.054	3.288	2.654	1.215	3.979
3.452	3.569	2.523	1.584	3.998	5.135	3.842