9 INTERVENTION ANALYSIS AND OUTLIER DETECTION

Time series are frequently affected by certain external events such as holidays, strikes, sales, promotions, and other policy changes. We call these external events *interventions*. In this chapter we introduce the technique, called *intervention analysis*, to evaluate the effect of these external events. Intervention analysis has been successfully used to study the impact of air pollution control and economic policies (Box and Tiao, 1975), the impact of the Arab oil embargo (Montgomery and Weatherby, 1980), the impact of the New York blackout (Izenman and Zabell, 1981), and many other events. We first discuss the analysis when the timing of the interventions is known. The method is then generalized to study the impact of the events when the timing of interventions is unknown and hence leads to the general time series outlier analysis.

9.1 INTERVENTION MODELS

Given that a known intervention occurs at time T, is there any evidence of a change in the time series (such as the increase of the mean level), and if so, by how much? One may initially think that the traditional two-sample t-test could be used to analyze this problem in terms of comparing the pre-intervention data with the postintervention data. However, the t-test assumes both normality and independence. Even though the t-test is known to be robust with respect to the normality assumption, it is extremely sensitive to the violation of the independence assumption as shown by Box and Tiao (1965), who developed the intervention analysis to study a time series structural change due to external events (Box and Tiao, 1975).

There are two common types of intervention variables. One represents an intervention occurring at time T that remains in effect thereafter. That is, the intervention is a step function,

$$S_t^{(T)} = \begin{cases} 0, & t < T, \\ 1, & t \geq T. \end{cases} \tag{9.1.1}$$

The other one represents an intervention taking place at only one time period. Thus, it is a pulse function,

$$P_t^{(T)} = \begin{cases} 1, & t = T, \\ 0, & t \neq T. \end{cases} \tag{9.1.2}$$

Note that the pulse function can be produced by differencing the step function $S_t^{(T)}$. That is, $P_t^{(T)} = S_t^{(T)} - S_{t-1}^{(T)} = (1-B)S_t^{(T)}$. Therefore, an intervention model can be represented equally well with the step function or with the pulse function. The use of a specific form is usually based on the convenience of interpretation.

There are many possible responses to the step and pulse interventions. We illustrate some commonly encountered ones.

1. A fixed unknown impact of an intervention is felt b periods after the intervention. Thus, depending on the type of intervention, the impact is

$$\omega B^b S_t^{(T)} \tag{9.1.3}$$

or

$$\omega B^b P_t^{(T)}. \tag{9.1.4}$$

2. An impact of an intervention is felt b periods after the intervention, but the response is gradual. For a step input, we have

$$\frac{\omega B^b}{(1-\delta B)} S_t^{(T)}, \tag{9.1.5}$$

and for a pulse input,

$$\frac{\omega B^b}{(1-\delta B)} P_t^{(T)} \tag{9.1.6}$$

where $0 \leq \delta \leq 1$. For $\delta = 0$, (9.1.5) and (9.1.6) reduce to (9.1.3) and (9.1.4), respectively. If $\delta = 1$, the impact increases linearly without bound. For most cases, we have $0 < \delta < 1$, and the response is gradual.

For illustration, we plot the above interventions with $b = 1$ and $0 < \delta < 1$ in Figure 9.1.

Note that various responses can be produced by different combinations of step and pulse inputs. For example, we may have the response

$$\frac{\omega_0 B}{(1-\delta B)} P_t^{(T)} + \omega_1 B S_t^{(T)} \tag{9.1.7}$$

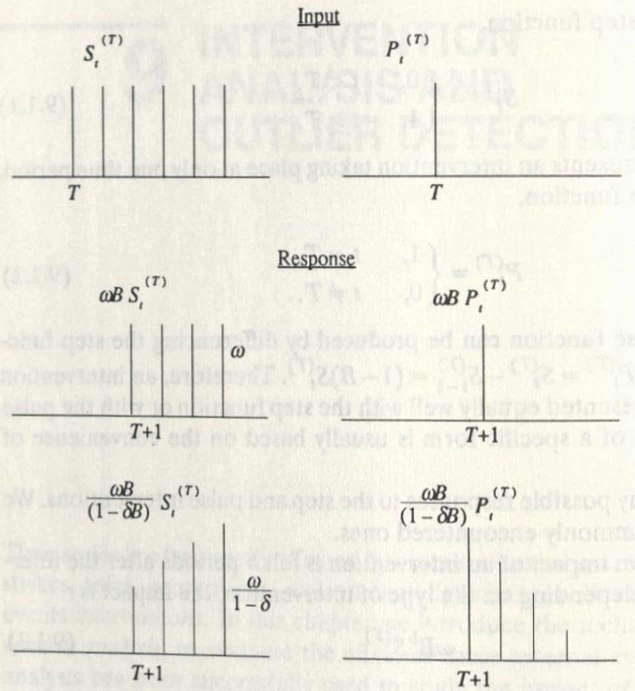

Fig. 9.1 Responses to step and pulse inputs.

as represented in Figure 9.2. However, as mentioned earlier, because $P_t^{(T)} = (1-B)S_t^{(T)}$, response (9.1.7) can also be written as

$$\left[\frac{\omega_0 B}{(1-\delta B)}+\frac{\omega_1 B}{(1-B)}\right] P_t^{(T)}. \tag{9.1.8}$$

This model is useful to represent the phenomenon in which an intervention produces a response that tapers off gradually but leaves a permanent residue effect in the system. The impact of an intervention such as advertising on sales can be represented as shown in Figure 9.2(a), and the effect of a price or a tax increase on imports may be represented in Figure 9.2(b).

More generally, a response may be represented as a rational function

$$\frac{\omega(B)B^b}{\delta(B)} \tag{9.1.9}$$

where $\omega(B) = \omega_0 - \omega_1 B - \cdots - \omega_s B^s$ and $\delta(B) = 1 - \delta_1 B - \cdots - \delta_r B^r$ are polynomials in B, b is the time delay for the intervention effect, and the weights ω_j's in the polynomial $\omega(B)$ often represent the expected initial effects of the intervention. The polynomial $\delta(B)$, on the other hand, measures the behavior of the permanent effect of the intervention. The roots of $\delta(B) = 0$ are assumed

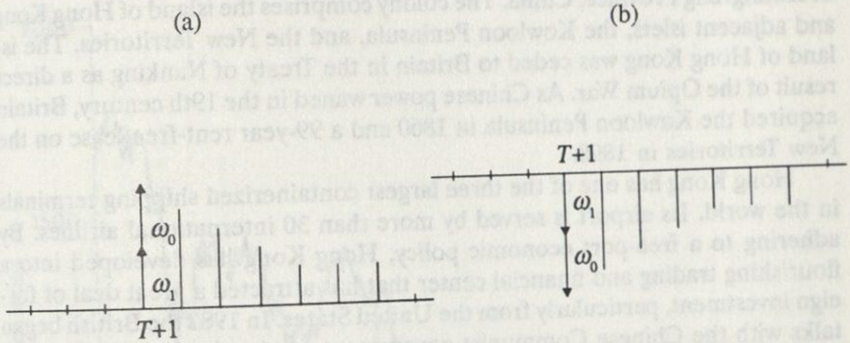

Fig. 9.2 Response to combined inputs, $\frac{\omega_0 B}{1-\delta B}P_t^{(T)} + \omega_1 BS_t^{(T)}$. (a) $\omega_0 > 0$ and $\omega_1 > 0$. (b) $\omega_0 < 0$ and $\omega_1 < 0$.

to be on or outside the unit circle. The unit root represents an impact that increases linearly, and the root outside the unit circle represents a phenomenon that has a gradual response.

For multiple intervention inputs, we have the following general class of models:

$$Z_t = \sum_{j=1}^{k} \frac{\omega_j(B)B^{b_j}}{\delta_j(B)} I_{jt} + \frac{\theta(B)}{\psi(B)} a_t \tag{9.1.10}$$

where I_{jt}, $j = 1, 2, \ldots, k$ are intervention variables. These intervention variables can be either step or pulse functions. More generally, they can be proper indicator variables, as shown later in Example 9.4. The form $\omega_j(B)B^{b_j}/\delta_j(B)$ for the jth intervention is postulated based on the expected form of the response given knowledge of the intervention. The main purpose of intervention models is to measure the effect of interventions. Thus, with respect to the intervention variables I_{jt}, the time series free of intervention is called the *noise series* and denoted by N_t, and its model is hence known as the *noise model*. The noise model $[\theta(B)/\Psi(B)]a_t$ is usually identified using the univariate model identification procedure based on the time series Z_t before the date of intervention, i.e., $\{Z_t : t < T\}$. If diagnostic checking of the model reveals no model inadequacy, then we can make appropriate inferences about the intervention. Otherwise, appropriate modifications must be made to the model, and estimation and diagnostic checking repeated.

9.2 EXAMPLES OF INTERVENTION ANALYSIS

Example 9.1 Hong Kong, a British Crown Colony since the first Opium War (1839–1842) between China and Britain, is located off the southern coast of Kwangtung Province, China. The colony comprises the island of Hong Kong and adjacent islets, the Kowloon Peninsula, and the New Territories. The island of Hong Kong was ceded to Britain in the Treaty of Nanking as a direct result of the Opium War. As Chinese power waned in the 19th century, Britain acquired the Kowloon Peninsula in 1860 and a 99-year rent-free lease on the New Territories in 1898.

Hong Kong has one of the three largest containerized shipping terminals in the world. Its airport is served by more than 30 international airlines. By adhering to a free-port economic policy, Hong Kong has developed into a flourishing trading and financial center that has attracted a great deal of foreign investment, particularly from the United States. In 1982 the British began talks with the Chinese Communist government concerning the 1997 expiration of the lease on the New Territories. Arguing that the treaties ceding the entire territory of Hong Kong were unequal and unjust, the Chinese Communist government announced on July 16, 1982, its first proposal for the eventual reversion of Hong Kong to its sovereignty and administration. The announcement caused tremendous anxiety to residents of and investors in Hong Kong. The Hong Kong dollar plummeted to an all-time low on foreign exchange markets. The property market became bearish. The stock market fell lower than at any time since the global economic crisis in 1973. In the following we perform an intervention analysis to assess the impact of this announcement on Hong Kong's stock market.

Figure 9.3 shows Series W11, the daily closing Hong Kong index of stock prices between July 16, 1981, and September 31, 1983, issued by Hong Kong Economic and Research Center at the end of each trading day. The index is the sum of more than thirty stocks that trade in Hong Kong. The period between July 16, 1981, and July 15, 1982, is regarded as the noise series containing no major intervention that would affect the index. The sample ACF and PACF of the original and the first differenced noise series as shown in Table 9.1 suggest the following random walk model:

$$(1-B)N_t = a_t. \tag{9.2.1}$$

By assuming the effect of the Chinese Communist government proposal on Hong Kong stocks is to cause an immediate level change in stock prices, one could propose the following response function:

$$\omega_0 I_t \tag{9.2.2}$$

where ω_0 represents the impact of the announcement and

$$I_t = \begin{cases} 0, & t < 260 \text{ (July 16, 1982)}, \\ 1, & t \geq 260 \text{ (July 16, 1982)}. \end{cases}$$

The intervention model becomes

$$Z_t = \omega_0 I_t + \frac{a_t}{(1-B)}. \tag{9.2.3}$$

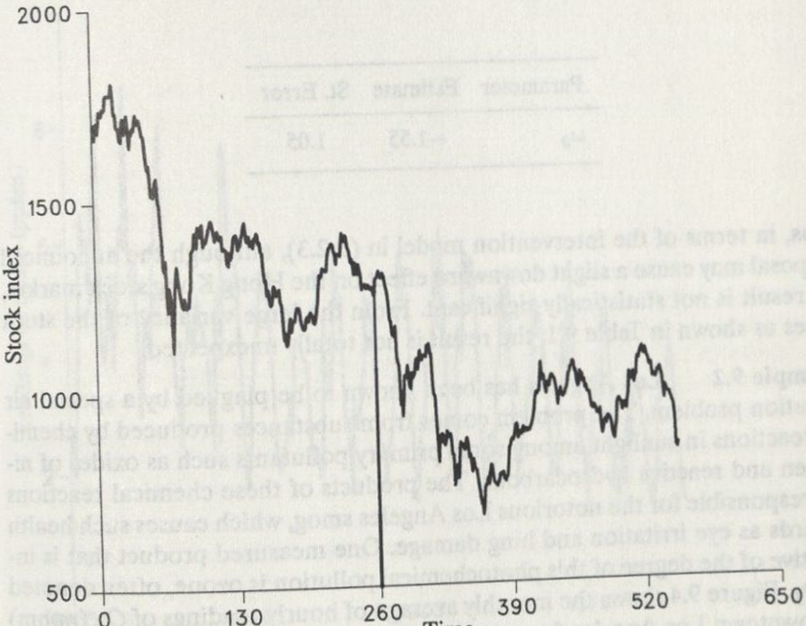

Fig. 9.3 Hang Seng index of Hong Kong stock prices from July 16, 1981, to September 31, 1983 (Series W11).

Table 9.1 Sample ACF and PACF of Hong Kong stock prices.

(a) $\{N_t\}$ $(\bar{N} = 1397.32, S_N = 174.46, n = 260)$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.98	.97	.95	.93	.91	.88	.86	.84	.82	.80
St.E.	.06	.11	.14	.16	.18	.28	.21	.22	.24	.25
$\hat{\phi}_{kk}$	98	−.10	.04	−.05	.04	−.06	.13	−.09	.09	.07
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

(b) $\{W_t = (1-B)N_t\}$ $(\bar{W} = -1.51, S_W = 25.33, n = 259)$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_k$	.09	−.06	.11	.03	.05	−.16	.03	.05	−.01	.05
St.E.	.06	.06	.06	.06	.06	.06	.07	.07	.07	.07
$\hat{\phi}_{kk}$	.09	−.07	.12	.01	.07	−.19	.08	.00	.04	.03
St.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

The estimation results are:

Parameter	Estimate	St. Error
ω_0	-1.55	1.05

Thus, in terms of the intervention model in (9.2.3), although the announced proposal may cause a slight downward effect on the Hong Kong stock market, the result is not statistically significant. From the large variance of the stock prices as shown in Table 9.1, the result is not totally unexpected.

Example 9.2 Los Angeles has been known to be plagued by a special air pollution problem. The problem comes from substances produced by chemical reactions in sunlight among some primary pollutants such as oxides of nitrogen and reactive hydrocarbons. The products of these chemical reactions are responsible for the notorious Los Angeles smog, which causes such health hazards as eye irritation and lung damage. One measured product that is indicative of the degree of this photochemical pollution is ozone, often denoted by O_3. Figure 9.4 shows the monthly average of hourly readings of O_3 (pphm) in downtown Los Angeles from 1955 to 1972.

To ease the air pollution problem, different methods were instituted. These include the diversion of traffic in early 1960 by the opening of the Golden State Freeway and the inception of a new law (Rule 63) that reduced the allowable proportion of reactive hydrocarbons in the gasoline sold locally. Also, after 1966 special regulations were implemented to require engine design changes in new cars in order to reduce the production of O_3. It was through the study of the effect of these events on the pollution problem that Box and Tiao (1975) introduced the intervention analysis.

The period from 1955 to 1960 is assumed to be free of intervention effects and is used to estimate the noise model for N_t. The sample ACF within this period suggest nonstationary and highly seasonal behavior. The ACF of the seasonally differenced series $(1-B)^{12}N_t$ have significant spikes only at lags 1 and 12, which implies the following noise model:

$$(1-B^{12})N_t = (1-\theta B)(1-\Theta B^{12})a_t. \tag{9.2.4}$$

Box and Tiao (1975) suggest that the opening of the Golden State Freeway and the implementation of Rule 63 in 1960 represent an intervention I_1, which might be expected to produce a step change in the O_3 level at the beginning of 1960. Intervention I_2 would be represented by the regulations implemented in 1966 requiring engine changes in new cars. The effect of I_2 would be most accurately measured by the proportion of new cars having specified engine changes in the car population over time. Unfortunately, no such data are available.

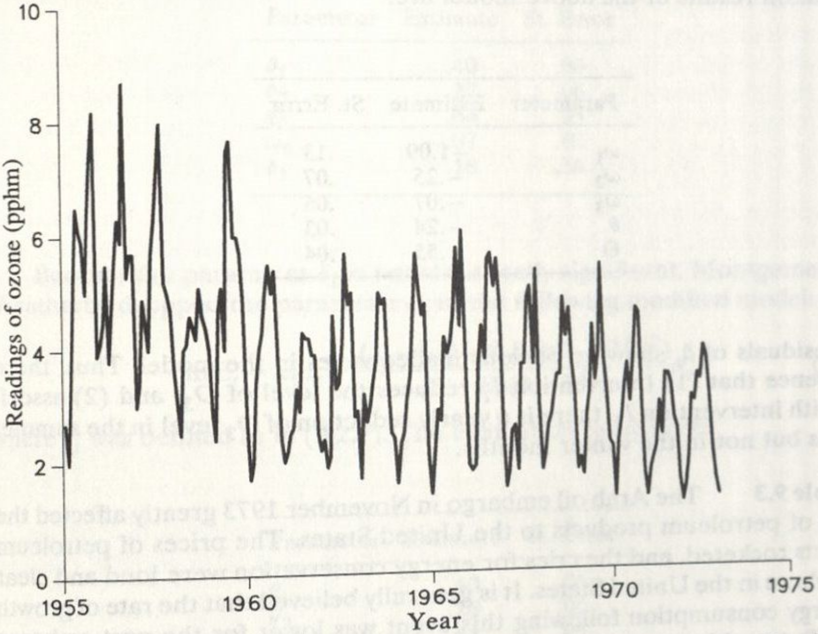

Fig. 9.4 Monthly average of hourly readings of O_3 (pphm) in downtown Los Angeles, 1955–1972.

However, one might represent the effect of I_2 as an annual trend reflecting the effect of the increased proportion of new design vehicles in the population. Because of the differences in the intensity of sunlight and other meteorological conditions between the summer months and the winter months, the effect of I_2 would be different in these two seasons. Thus, Box and Tiao (1975) proposed the following model:

$$Z_t = \omega_1 I_{1t} + \frac{\omega_2}{1-B^{12}} I_{2t} + \frac{\omega_3}{1-B^{12}} I_{3t} + \frac{(1-\theta B)(1-\Theta B^{12})}{(1-B^{12})} a_t \tag{9.2.5}$$

where

$$I_{1t} = \begin{cases} 0, & t < \text{January 1960,} \\ 1, & t \geq \text{January 1960,} \end{cases}$$

$$I_{2t} = \begin{cases} 1, & \text{“summer” months June–October beginning in 1966,} \\ 0, & \text{otherwise,} \end{cases}$$

$$I_{3t} = \begin{cases} 1, & \text{“winter” months November–May beginning in 1966,} \\ 0, & \text{otherwise.} \end{cases}$$

Estimation results of the above model are:

Parameter	Estimate	St. Error
ω_1	-1.09	.13
ω_2	$-.25$	.07
ω_3	$-.07$	.06
θ	$-.24$	.03
Θ	.55	.04

Residuals of $\hat{a}_t$ show no obvious inadequacies in the model. Thus, there is evidence that (1) Intervention I_1 reduces the level of O_3, and (2) associated with intervention I_2, there is a yearly reduction of O_3 level in the summer months but not in the winter months.

Example 9.3 The Arab oil embargo in November 1973 greatly affected the supply of petroleum products to the United States. The prices of petroleum products rocketed, and the cries for energy conservation were loud and clear everywhere in the United States. It is generally believed that the rate of growth of energy consumption following this event was lower for the post-embargo years. To test this assumption, Montgomery and Weatherby (1980) applied an intervention model to the natural logarithm of monthly electricity consumption from January 1951 to April 1977 using a total of 316 observations.

Although the embargo started in November 1973, Montgomery and Weatherby (1980) assumed that the effect of this embargo was not felt until December 1973. Thus, they used 275 months of data from January 1951 to November 1973 to model the noise series and obtained the following ARIMA(0, 1, 2) $\times$ (0, 1, 1)$_{12}$ model:

$$(1-B)(1-B^{12})\ln N_t = (1-\theta_1 B-\theta_2 B^2)(1-\Theta B^{12})a_t. \tag{9.2.6}$$

By assuming that the effect of the embargo is to cause a gradual change in consumption, one could propose the following intervention model

$$\ln Z_t = \frac{\omega_0}{(1-\delta_1 B)} I_t + \frac{(1-\theta_1 B-\theta_2 B^2)(1-\Theta B^{12})}{(1-B)(1-B^{12})} a_t \tag{9.2.7}$$

where ω_0 represents the initial impact of the oil embargo and

$$I_t = \begin{cases} 0, & t \le 275, \\ 1, & t > 275. \end{cases}$$

The estimates and the associated standard errors are:

Parameter	Estimate	St. Error
θ_1	.40	.06
θ_2	.27	.06
Θ	.64	.05
ω_0	−.07	.03
δ_1	.18	.36

Because the parameter δ_1 is not statistically significant, Montgomery and Weatherby dropped the parameter δ_1 in the following modified model:

$$\ln Z_t = \omega_0 I_t + \frac{(1-\theta_1 B - \theta_2 B^2)(1-\Theta B^{12})}{(1-B)(1-B^{12})} a_t \tag{9.2.8}$$

where I_t was defined as in (9.2.7). The estimation results are:

Parameter	Estimate	St. Error
θ_1	.40	.06
θ_2	.28	.06
Θ	.64	.05
ω_0	−.07	.02

The parameters are all statistically significant. The residual ACF do not exhibit any model inadequacy. Thus, the intervention model in (9.2.8) is satisfactory. The result implies that the embargo induced a permanent level change on electricity consumption. Because the model was built using the natural logarithm of electricity consumption, the estimate of the intervention effect in terms of the original MWH metric is $e^{\hat{\omega}_0} = e^{-.07} = .93$. Therefore, the postintervention level of electricity consumption is 93% of the pre-intervention level, or equivalently, the effect of the Arab oil embargo has been to reduce the growth of electricity consumption by 7%.

Example 9.4 At exactly 5:27 P.M., November 9, 1965, most of New York City was plunged into darkness due to a massive power failure. The blackout was long, and most of the City remained dark during the night. On Wednesday, August 10, 1966, the *New York Times* carried a front-page article with the headline "Births Up 9 Months After the Blackout." Numerous articles were published afterward in newspapers and magazines both inside and outside of the United States alleging a sharp increase in the city's birthrate. A number of medical and demographic articles then appeared with contradictory statements regarding the blackout effect. Using a total of 313 weekly births in New York City from

1961 to 1966, which is plotted in Figure 9.5, Izenman and Zabell (1981) applied the intervention technique to the phenomenon.

Since the blackout, which occurred on November 9, 1965, falls in the middle of the 254th week, the first 254 weekly birth totals are used to model the noise series with the following process:

$$(1-B)(1-B^{52})N_t = (1-\theta B)(1-\Theta B^{52})a_t. \tag{9.2.9}$$

Furthermore, the obstetrical and gynecological studies show that the mode of the gestational interval, which is defined as the time from onset of the last menstrual period (LMP) to birth, occurs at 40 or 41 weeks after the LMP. Izenman and Zabell (1981) suggested the intervention form

$$\omega_0 I_t \tag{9.2.10}$$

where ω_0 is used to represent the effect of the blackout and

$$I_t = \begin{cases} 1, & t = 292,\ 293,\ 294,\ 295, \\ 0, & \text{otherwise.} \end{cases} \tag{9.2.11}$$

The above intervention variable I_t was introduced to take the value 1 for weeks 38 through 41 after the blackout to give the model the maximum possible

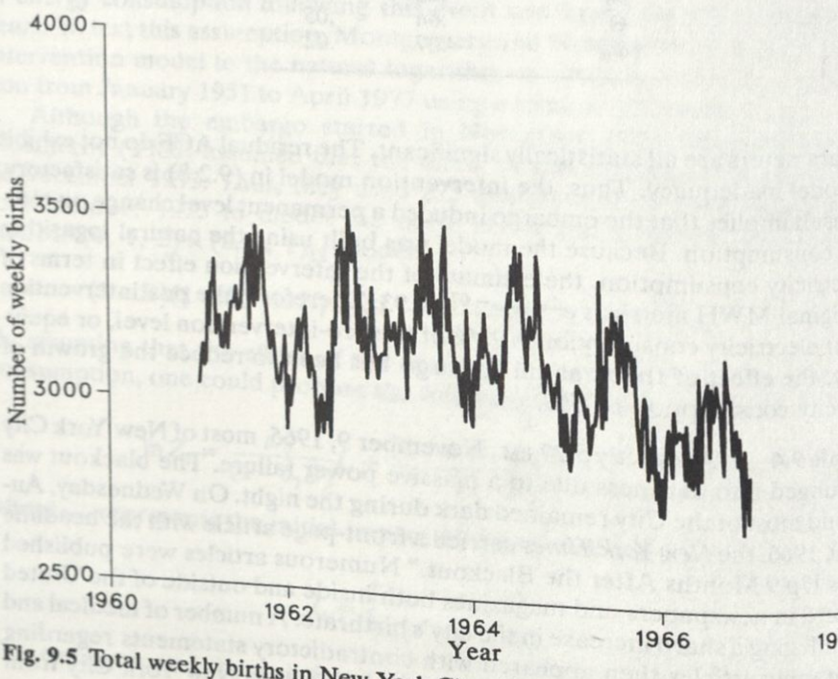

Fig. 9.5 Total weekly births in New York City, 1961–1966.

chance of detecting the effect. Thus, we have the following intervention model:

$$Z_t = \omega_0 I_t + \frac{(1-\theta B)(1-\Theta B^{52})}{(1-B)(1-B^{52})} a_t \qquad (9.2.12)$$

where I_t is defined as in (9.2.11).

The estimates and associated standard errors are:

Parameter	Estimate	St. Error
θ	.74	.04
Θ	.82	.02
ω_0	28.63	47.36

The residual ACF show no signs of model inadequacy. Because the parameter ω_0 is not statistically significant, we conclude that intervention analysis of the New York City birth data does not detect a significant increase in births that can be ascribed to the blackout. Here, we note that the model estimated in (9.2.12) is slightly different from the model given in Izenman and Zabell (1981).

9.3 TIME SERIES OUTLIERS

Time series observations are sometimes influenced by interruptive events, such as strikes, outbreaks of war, sudden political or economic crises, unexpected heat or cold waves, or even unnoticed errors of typing and recording. The consequences of these interruptive events create spurious observations that are inconsistent with the rest of the series. Such observations are usually referred to as *outliers*. When the timing and causes of interruptions are known, their effects can be accounted for by using the intervention model discussed in Sections 9.1 and 9.2. In practice, however, the timing of interruptive events are usually unknown. Because outliers are known to wreak havoc in data analysis, making the resultant inference unreliable or even invalid, it is important to have procedures that will detect and remove such outliers effects. The detection of time series outliers was first studied by Fox (1972), where two statistical models, additive and innovational, were introduced. Other references on this topic include Abraham and Box (1979), Martin (1980), Chang and Tiao (1983), Hillmer, Bell, and Tiao (1983), Tsay (1986), Chang, Tiao, and Chen (1988).

9.3.1 Additive and Innovational Outliers

For a given stationary or properly deduced stationary process, let Z_t be the observed series and X_t be the outlier-free series. Assume that $\{X_t\}$ follows a general ARMA(p, q) model

$$\phi(B)X_t = \theta(B)a_t \tag{9.3.1}$$

where $\phi(B) = 1-\phi_1B-\cdots-\phi_pB^p$ and $\theta(B) = (1-\theta_1B-\cdots-\theta_qB^q)$ are stationary and invertible operators sharing no common factors, and $\{a_t\}$ is a sequence of white noise, identically and independently distributed as $N(0, \sigma_a^2)$. An additive outlier (AO) is defined as

$$Z_t = \begin{cases} X_t, & t \neq T \\ X_t + \omega, & t = T \end{cases} \tag{9.3.2a}$$

$$= X_t + \omega I_t^{(T)} \tag{9.3.2b}$$

$$= \frac{\theta(B)}{\phi(B)}a_t + \omega I_t^{(T)} \tag{9.3.2c}$$

where

$$I_t^{(T)} = \begin{cases} 1, & t = T, \\ 0, & t \neq T, \end{cases}$$

is the indicator variable representing the presence or absence of an outlier at time T. An innovational outlier (IO) model is defined as

$$Z_t = X_t + \frac{\theta(B)}{\phi(B)}\omega I_t^{(T)} \tag{9.3.3a}$$

$$= \frac{\theta(B)}{\phi(B)}(a_t + \omega I_t^{(T)}). \tag{9.3.3b}$$

Hence, an additive outlier affects only the level of the Tth observation, whereas an innovational outlier affects all observations $Z_T, Z_{T+1}, \ldots,$ beyond time T through the memory of the system described by $\theta(B)/\phi(B)$.

More generally, a time series might contain several, say k, outliers of different types, and we have the following general outlier model:

$$Z_t = \sum_{j=1}^{k} \omega_j \nu_j(B) I_t^{(T_j)} + X_t \tag{9.3.4}$$

where $X_t = \frac{\theta(B)}{\phi(B)}a_t$, $\nu_j(B) = 1$ for an AO and $\nu_j(B) = \theta(B)/\phi(B)$ for an IO at time $t = T_j$.

9.3.2 Estimation of the Outlier Effect When the Timing of the Outlier Is Known

To motivate the procedure for detecting AO and IO, we consider a simpler case when T and all parameters in (9.3.1) are known. Letting

$$\pi(B) = \frac{\phi(B)}{\theta(B)} = (1-\pi_1B-\pi_2B^2-\cdots) \tag{9.3.5}$$

and defining

$$e_t = \pi(B)Z_t, \tag{9.3.6}$$

we have from (9.3.2c) and (9.3.3a) that

$$\text{AO:} \quad e_t = \omega\pi(B)I_t^{(T)} + a_t, \tag{9.3.7}$$

and

$$\text{IO:} \quad e_t = \omega I_t^{(T)} + a_t. \tag{9.3.8}$$

For n available observations, the AO model in (9.3.7) can be written as

$$\begin{bmatrix} e_1 \\ \vdots \\ e_{T-1} \\ e_T \\ e_{T+1} \\ e_{T+2} \\ \vdots \\ e_n \end{bmatrix} = \omega \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -\pi_1 \\ -\pi_2 \\ \vdots \\ -\pi_{n-T} \end{bmatrix} + \begin{bmatrix} a_1 \\ \vdots \\ a_{T-1} \\ a_T \\ a_{T+1} \\ a_{T+2} \\ \vdots \\ a_n \end{bmatrix}. \tag{9.3.9}$$

Let $\hat{\omega}_{AT}$ be the least square estimator of ω for the AO model. Because $\{a_t\}$ is white noise, from the least squares theory, we have that

$$\begin{aligned} \text{AO:} \quad \hat{\omega}_{AT} &= \frac{e_T - \sum_{j=1}^{n-T} \pi_j e_{T+j}}{\sum_{j=0}^{n-T} \pi_j^2} \\ &= \frac{\pi^*(F)e_T}{\tau^2} \end{aligned} \tag{9.3.10}$$

where $\pi^*(F) = (1 - \pi_1 F - \pi_2 F^2 - \cdots - \pi_{n-T}F^{n-T})$, F is the forward shift operator such that $Fe_t = e_{t+1}$ and $\tau^2 = \sum_{j=0}^{n-T} \pi_j^2$. The variance of the estimator is

$$\begin{aligned} \text{Var}(\hat{\omega}_{AT}) &= \text{Var}\left(\frac{\pi^*(F)e_T}{\tau^2}\right) \\ &= \frac{1}{\tau^4}\text{Var}[\pi^*(F)a_T] \\ &= \frac{\sigma_a^2}{\tau^2}. \end{aligned} \tag{9.3.11}$$

Similarly, letting $\hat{\omega}_{IT}$ be the least squares estimator of ω for the IO model, we have

$$\text{IO:} \quad \hat{\omega}_{IT} = e_T, \tag{9.3.12}$$

and

$$\begin{aligned}\text{Var}(\hat{\omega}_{IT}) &= \text{Var}(e_T) = \text{Var}(\omega I_t^{(T)} + a_T) \\ &= \sigma_a^2.\end{aligned} \tag{9.3.13}$$

Thus, the best estimate of the effect of an IO at time T is the residual e_T, whereas the best estimate of the effect of an AO is a linear combination of $e_T, e_{T+1}, \dots$ and e_n with weight depending on the structure of the time series process. It is easily seen that $\text{Var}(\hat{\omega}_{AT}) \le \text{Var}(\hat{\omega}_{IT}) = \sigma_a^2$ and in some cases $\text{Var}(\hat{\omega}_{AT})$ can be much smaller than σ_a^2.

Various tests can be performed for the hypotheses,

$$\begin{aligned} H_0:&\quad Z_T \text{ is neither an AO nor an IO} \\ H_1:&\quad Z_T \text{ is an AO} \\ H_2:&\quad Z_T \text{ is an IO.} \end{aligned}$$

The likelihood ratio test statistics for AO and IO are

$$H_1 \text{ vs. } H_0: \quad \lambda_{1,T} = \tau\hat{\omega}_{AT}/\sigma_a \tag{9.3.14}$$

and

$$H_2 \text{ vs. } H_0: \quad \lambda_{2,T} = \hat{\omega}_{IT}/\sigma_a. \tag{9.3.15}$$

Under the null hypothesis H_0, both $\lambda_{1,T}$ and $\lambda_{2,T}$ are distributed as N(0, 1).

9.3.3 Detection of Outliers Using an Iterative Procedure

If T is unknown but the time series parameters are known, we can proceed to calculate $\lambda_{1,t}$ and $\lambda_{2,t}$ for each $t = 1, 2, \dots, n$, and then make decisions based on the above sampling results. However, in practice, the time series parameters ϕ_j, θ_j, π_j, and σ_a^2 are usually unknown and have to be estimated. It is known that existence of outliers makes the estimates of the parameters seriously biased. In particular, σ_a^2 will tend to be overestimated, as shown earlier. Chang and Tiao (1983) proposed an iterative detecting procedure to handle the situation when an unknown number of AO or IO may exist.

Step 1. Model the series $\{Z_t\}$ by assuming that there are no outliers. Compute the residuals from the estimated model, i.e.,

$$\begin{aligned}\hat{e}_t &= \hat{\pi}(B)Z_t \\ &= \frac{\hat{\phi}(B)}{\hat{\theta}(B)}Z_t\end{aligned} \tag{9.3.16}$$

where $\hat{\phi}(B) = (1 - \hat{\phi}_1 B - \dots - \hat{\phi}_p B^p)$ and $\hat{\theta}(B) = (1 - \hat{\theta}_1 B - \dots - \hat{\theta}_q B^q)$. Let

$$\hat{\sigma}_a^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{e}_t^2$$

be the initial estimate of σ_a^2.

Step 2. Calculate $\hat{\lambda}_{1,t}$ and $\hat{\lambda}_{2,t}$ for $t = 1, 2, \ldots, n$, using the estimated model. Define

$$\hat{\lambda}_T = \max_t \max_i \left\{ |\hat{\lambda}_{i,t}| \right\}, \tag{9.3.17}$$

where T denotes the time when the maximum occurs. If $\hat{\lambda}_T = |\hat{\lambda}_{1,T}| > C$, where C is a predetermined positive constant usually taken to be some value between 3 and 4, then there is an AO at time T with its effect estimated by $\hat{\omega}_{AT}$. One can modify the data using (9.3.2b) as follows:

$$\tilde{Z}_t = Z_t - \hat{\omega}_{AT} I_t^{(T)}, \tag{9.3.18}$$

and define the new residuals using (9.3.7)

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{AT} \hat{\pi}(B) I_t^{(T)}. \tag{9.3.19}$$

If $\hat{\lambda}_T = |\hat{\lambda}_{2,T}| > C$, then there is an IO at time T with its effect being $\hat{\omega}_{IT}$. This IO effect can be removed by modifying the data using (9.3.3a), i.e.,

$$\tilde{Z}_t = Z_t - \frac{\hat{\theta}(B)}{\hat{\phi}(B)} \hat{\omega}_{IT} I_t^{(T)}, \tag{9.3.20}$$

and defining the new residuals using (9.3.8)

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{IT} I_t^{(T)}. \tag{9.3.21}$$

A new estimate $\tilde{\sigma}_a^2$ is then computed from the modified residuals.

Step 3. Recompute $\hat{\lambda}_{1,t}$ and $\hat{\lambda}_{2,t}$ based on the modified residuals and $\tilde{\sigma}_a^2$, and repeat Step 2 until all outliers are identified. The initial estimates in $\pi(B)$ remain unchanged.

Step 4. Suppose that Step 3 terminated and k outliers have been tentatively identified at times $T_1, T_2, \ldots$, and T_k. Treat these times as if they are known, and estimate the outlier parameters $\omega_1, \omega_2, \ldots, \omega_k$ and the time series parameters simultaneously using the model

$$Z_t = \sum_{j=1}^{k} \omega_j \nu_j(B) I_j^{(T_j)} + \frac{\theta(B)}{\phi(B)} a_t \tag{9.3.22}$$

where $\nu_j(B) = 1$ for an AO and $\nu_j(B) = \theta(B)/\phi(B)$ for an IO at $t = T_j$. This leads to the new residuals

$$\hat{e}_t^{(1)} = \hat{\pi}^{(1)}(B) \left[Z_t - \sum_{j=1}^{k} \hat{\omega}_j \nu_j(B) I_t^{(T_j)} \right]. \tag{9.3.23}$$

A revised estimate of σ_a^2 can then be calculated.

Step 2 through Step 4 are repeated until all outliers are identified and their impacts simultaneously estimated. Thus, we have the following fitted outlier model:

$$Z_t = \sum_{j=1}^{k} \hat{\omega}_j \nu_j(B) I_t^{(T_j)} + \frac{\hat{\theta}(B)}{\hat{\phi}(B)} a_t \tag{9.3.24}$$

where $\hat{\omega}_j$, $\hat{\phi}(B) = (1 - \hat{\phi}_1 B - \cdots - \hat{\phi}_p B^p)$ and $\hat{\theta}(B) = (1 - \hat{\theta}_1 B - \cdots - \hat{\theta}_q B^q)$ are obtained in the final iteration.

9.4 EXAMPLES OF OUTLIER ANALYSIS

The above outlier detection procedure can be easily implemented in any existing intervention analysis software or linear regression package. Both time series software SCA and AUTOBOX have implemented the procedure and make the analysis easier. A computer program based on a modification of the above procedure was also written by Bell (1984). We illustrate the following examples by using AUTOBOX and a standard regression package.

Example 9.5 The outlier detection procedure discussed in Section 9.3 has been applied using software AUTOBOX to the U.S. quarterly series of beer production between 1975 and 1982, which was fitted earlier by a seasonal ARIMA(0, 0, 0) × (0, 1, 1)$_4$ model in Chapter 8. The result indicates that there are no outliers evident with a significance level .05 in the series. To check the efficiency of the proposed procedure, we artificially contaminated the series by replacing the original observation $Z_{12} = 36.54$ with a new observation $Z_{12} = 56.54$, which could be due to a typing error. We then apply the procedure to this outlier contaminated series, which is plotted in Figure 9.6.

The analysis produces the following result:

Iteration	Detected Outliers Time	Type	Magnitude ($\hat{\omega}$)
1	12	AO	16.26
2	27	IO	−2.31

Thus, the procedure correctly identifies the AO at $t = 12$ in the first iteration. Although Z_{27} has also been detected as an IO in the second iteration, the effect is much smaller.

Example 9.6 Series W1 was analyzed in Sections 6.2 and 7.6 resulting in an AR(1) model

$$(1 - .43B)Z_t = .89 + a_t \tag{9.4.1}$$

with $\hat{\sigma}_a^2 = .21$. Inspection of the residuals from the fitted model indicates the possible existence of a number of outliers. The series is the daily average num-

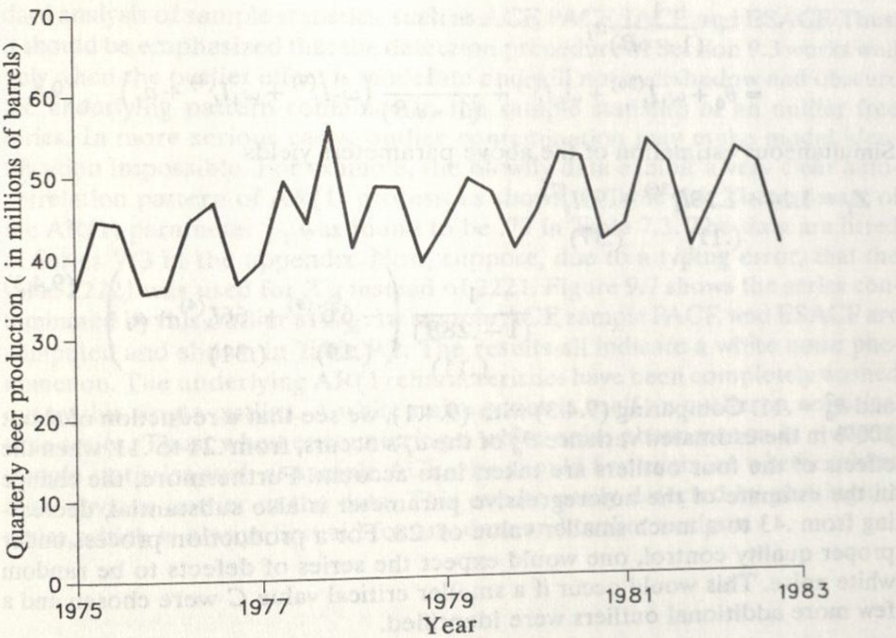

Fig. 9.6 Outlier contaminated U.S. quarterly series of beer production between 1975 and 1982.

ber of defects per truck found in the final inspection at the end of the assembly line of a truck manufacturing plant. To ensure the maintenance of quality, the detection of outliers is always an important task in quality control. In this example, we apply the above outlier detection procedure to the data set and obtain the following results:

Detected Outliers		
Iteration	Time	Type
1	36	AO
2	9	IO
3	7	AO
4	4	IO

Thus, we consider the following outlier model:

$$Z_t = \theta_0 + \omega_1 I_t^{(36)} + \omega_2 \frac{1}{(1-\phi B)} I_t^{(9)} + \omega_3 I_t^{(7)} + \omega_4 \frac{1}{(1-\phi B)} I_t^{(4)}$$

$$+\frac{1}{(1-\phi B)}a_t$$
$$= \theta_0 + \omega_1 I_t^{(36)} + \omega_3 I_t^{(7)} + \frac{1}{(1-\phi B)}\left(\omega_2 I_t^{(9)} + \omega_4 I_t^{(4)} + a_t\right) \qquad (9.4.2)$$

Simultaneous estimation of the above parameters yields

$$Z_t = \underset{(.11)}{1.14} + \underset{(.37)}{1.39} I_t^{(36)} + .99 I_t^{(7)}$$
$$+ \frac{1}{(1-\underset{(.11)}{.28}B)}\left(-\underset{(.19)}{.61}I_t^{(9)} + \underset{(.31)}{.66}I_t^{(4)} + a_t\right) \qquad (9.4.3)$$

and $\hat{\sigma}_a^2 = .11$. Comparing (9.4.3) with (9.4.1), we see that a reduction of about 100% in the estimated variance $\hat{\sigma}_a^2$ of the a_t's occurs, from .21 to .11, when the effects of the four outliers are taken into account. Furthermore, the change in the estimate of the autoregressive parameter is also substantial, decreasing from .43 to a much smaller value of .28. For a production process under proper quality control, one would expect the series of defects to be random white noise. This would occur if a smaller critical value C were chosen and a few more additional outliers were identified.

9.5 REMARKS ON OUTLIER AND INTERVENTION PROBLEMS

1. After outliers are identified, one can adjust data using (9.3.18) or (9.3.20) and then pursue the analysis based on the adjusted data. However, a more fruitful approach is possibly to search for the causes of the identified outliers and to further fine-tune the fitted model in (9.3.24). This is true not only for parameter estimation but also for model checking and forecasting. In searching for the causes of an outlier, one may find the nature of the disturbance. For example, some outliers may turn out to be important intervention variables due to some policy changes with which the analyst was unfamiliar and hence that were overlooked at the preliminary stage of data analysis. Thus, instead of using the adjusted data by removing the effects of outliers, the analyst should incorporate the information into the model by introducing proper intervention variables and response functions as were discussed in Sections 9.1 and 9.2. This explicit form of a combined intervention-outlier model is usually more useful in forecasting and control than is the univariate fitted model based on the outlier adjusted data.

2. It should be noted that the outlier detection procedure introduced in Section 9.3 is based on the assumption that the underlying model for the outlier free series is either known or can be identified. However, in practice, the underlying model is usually unknown and has to be identified through the stan-

dard analysis of sample statistics, such as ACF, PACF, IACF, and ESACF. Thus, it should be emphasized that the detection procedure of Section 9.3 works well only when the outlier effect is moderate and will not overshadow and obscure the underlying pattern contained in the sample statistics of an outlier free series. In more serious cases, outlier contamination may make model identification impossible. For example, the blowfly data exhibit a very clear autocorrelation pattern of AR(1) process, as shown in Table 6.6. The estimate of the AR(1) parameter ϕ_1 was found to be .73 in Table 7.3. The data are listed as Series W3 in the appendix. Now, suppose, due to a typing error, that the value 22221 was used for Z_{20} instead of 2221. Figure 9.7 shows the series contaminated by this outlier at Z_{20}. Its sample ACF, sample PACF, and ESACF are computed and shown in Table 9.2. The results all indicate a white noise phenomenon. The underlying AR(1) characteristics have been completely washed out by this single outlier. A white noise series is itself an outlier in empirical time series. Thus, when encountering a white noise phenomenon in studying sample statistics such as sample ACF, one should first examine whether there is an obvious outlier in the data. This can be easily detected by plotting the series, which is always first aid for any data and outlier analysis.

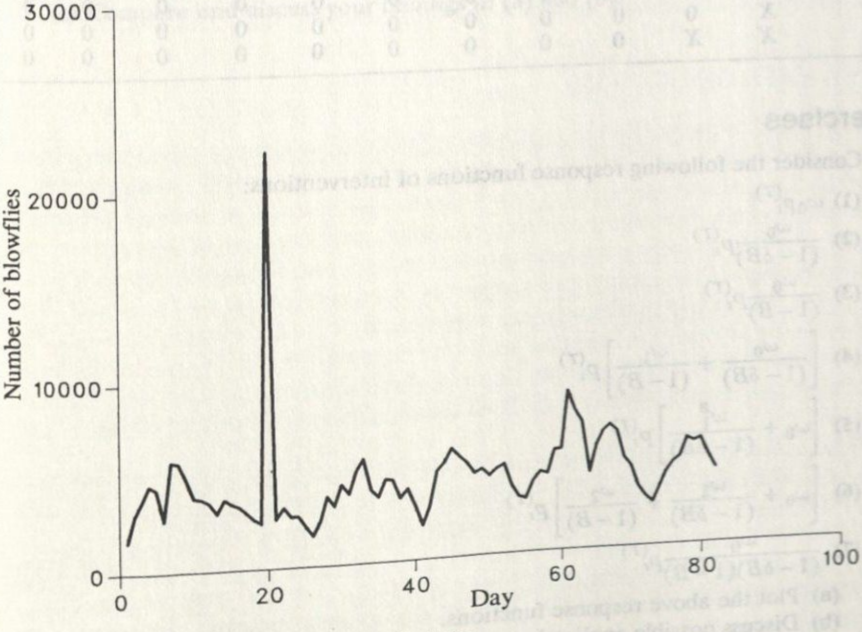

Fig. 9.7 Outlier contaminated blowfly data.

Table 9.2 Sample ACF, sample PACF, and ESACF for an outlier contaminated AR(1) series.

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

1–10	.09	.05	−.03	−.05	−.09	.12	.14	−.02	−.06	−.01
St.E.	.11	.11	.11	.11	.11	.11	.11	.11	.12	.12
11–20	.02	.06	.10	.04	.03	.07	.06	.00	−.00	−.08
St.E.	.12	.12	.12	.12	.12	.12	.12	.12	.12	.12

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

1–10	.09	.04	−.04	−.04	−.08	−.11	.12	−.04	−.06	−.02
St.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
11–20	.01	.02	.06	−.01	.01	.06	.06	.01	.03	−.04
St.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

(c) ESACF

MA / AR	0	1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	0	0	0
1	X	0	0	0	0	0	0	0	0	0	0
2	X	X	0	0	0	0	0	0	0	0	0
3	X	0	0	0	0	0	0	0	0	0	0
4	X	X	0	0	0	0	0	0	0	0	0

Exercises

9.1 Consider the following response functions of interventions:

(1) $\omega_0 P_t^{(T)}$

(2) $\dfrac{\omega_0}{(1-\delta B)} P_t^{(T)}$

(3) $\dfrac{\omega_0}{(1-B)} P_t^{(T)}$

(4) $\left[\dfrac{\omega_0}{(1-\delta B)} + \dfrac{\omega_1}{(1-B)}\right] P_t^{(T)}$

(5) $\left[\omega_0 + \dfrac{\omega_1^{B}}{(1-\delta B)}\right] P_t^{(T)}$

(6) $\left[\omega_0 + \dfrac{\omega_1}{(1-\delta B)} + \dfrac{\omega_2}{(1-B)}\right] P_t^{(T)}$

(7) $\dfrac{\omega_0}{(1-\delta B)(1-B)} P_t^{(T)}$

(a) Plot the above response functions.
(b) Discuss possible applications of the various interventions.

9.2 Find a time series that was affected by some external events. Carry out intervention analysis and submit a written report of your analysis.

9.3 Perform and report an iterative outlier analysis for the following time series (read across):

.561	.664	.441	.635	1.083	.961	.057
1.349	1.100	.544	−.132	−1.567	−1.277	−1.192
−1.346	1.401	.037	−.272	−.591	−.542	−.574
−.742	−1.416	.549	−1.446	1.883	1.050	1.134
1.947	−1.839	.803	.321	.470	−.279	1.913
−.785	.236	.147	−.690	.667	−.270	.221
−.633	−.245	−1.705	−1.648	−.723	−1.316	−.642
−.510	−.065	−.553	−1.058	−14.960	−.764	−.556
−.079	.047	−.203	.244	−.407	−.438	−1.616
−.231	−.371	−1.643	.203	−.338	−.830	−1.749
−1.025	−2.218	.360	−1.332	.199	−.034	.621
2.008	−.154	.308				

9.4 (a) Find a time series of your interest, which was likely contaminated by some outliers and/or interventions. Fit a best ARIMA model for the series.
(b) Detect all possible outlier and/or intervention variables, and build an outlier-intervention model.
(c) Compare and discuss your findings in (a) and (b).