CHAPTER 1

TIME SERIES DATA: EXAMPLES AND BASIC CONCEPTS

1.1 INTRODUCTION

In many fields of study, data is collected from a system (or as we would also like to call it a *process*) over time. This sequence of observations generates a time series such as the closing prices of the stock market, a country's unemployment rate, temperature readings of an industrial furnace, sea level changes in coastal regions, number of flu cases in a region, inventory levels at a production site, and so on. These are only a few examples of a myriad of cases where time series data is used to better understand the dynamics of a system and to make sensible forecasts about its future behavior.

Most physical processes exhibit inertia and do not change that quickly. This, combined with the sampling frequency, often makes consecutive observations correlated. Such correlation between consecutive observations is called *autocorrelation*. When the data is autocorrelated, most of the standard modeling methods based on the assumption of independent observations may become misleading or sometimes even useless. We therefore need to consider alternative methods that take into account the serial dependence in the data. This can be fairly easily achieved by employing time series models such as autoregressive integrated moving average (ARIMA) models. However, such models are usually difficult to understand from a practical point of view. What exactly do they mean? What are the practical implications of a given model and a specific set of parameters? In this book, our goal is to provide intuitive understanding of seemingly complicated time series models and their implications. We employ only the necessary amount of theory and attempt to present major concepts in time series analysis via numerous examples, some of which are quite well known in the literature.

1.2 EXAMPLES OF TIME SERIES DATA

Examples of time series can be found in many different fields such as finance, economics, engineering, healthcare, and operations management, to name a few.

Time Series Analysis and Forecasting by Example, First Edition. Søren Bisgaard and Murat Kulahci.

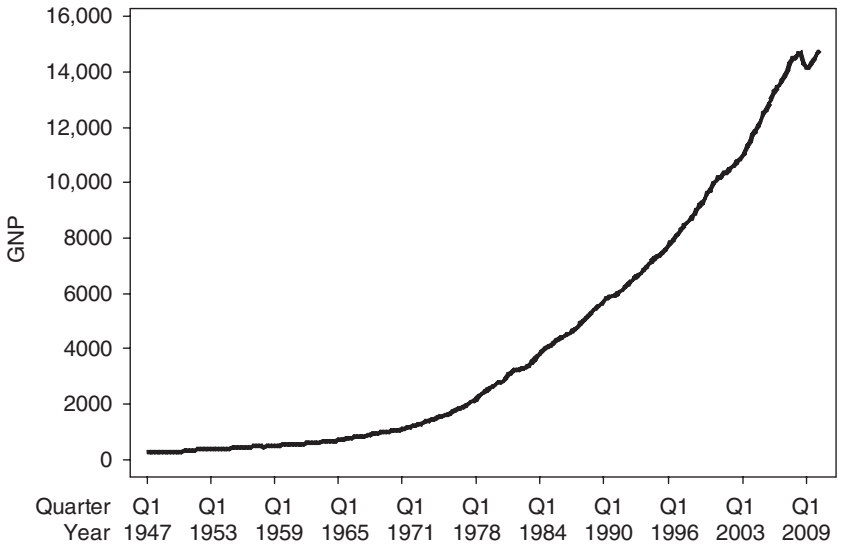

Figure 1.1 GNP (nominal) of the United States from 1947 to 2010 (in billion dollars). *Source:* US Department of Commerce, http://research.stlouisfed.org/fred2/data/GNP.txt.

Consider, for example, the gross national product (GNP) of the United States from 1947 to 2010 in Figure 1.1 where GNP shows a steady exponential increase over the years. However, there seems to be a "hiccup" toward the end of the period starting with the third quarter of 2008, which corresponds to the financial crisis that originated from the problems in the real estate market. Studying such macroeconomic indices, which are presented as time series, is crucial in identifying, for example, general trends in the national economy, impact of public policies, or influence of global economy.

Speaking of problems with the real estate market, Figure 1.2 shows the median sales prices of houses in the United States from 1988 to the second quarter of 2010. One can argue that the signs of the upcoming crisis could be noticed as early as in 2007. However, the more crucial issue now is to find out what is going to happen next. Homeowners would like to know whether the value of their properties will fall further and similarly the buyers would like to know whether the market has hit the bottom yet. These forecasts may be possible with the use of appropriate models for this and many other macroeconomic time series data.

Businesses are also interested in time series as in inventory and sales data. Figure 1.3 shows the well-known number of airline passengers data from 1949 to 1960, which will be discussed in greater detail in Chapter 5. On the basis of the cyclical travel patterns, we can see that the data exhibits a seasonal behavior. But we can also see an upward trend, suggesting that air travel is becoming more and more popular. Resource allocation and investment efforts in a company can greatly benefit from proper analysis of such data.

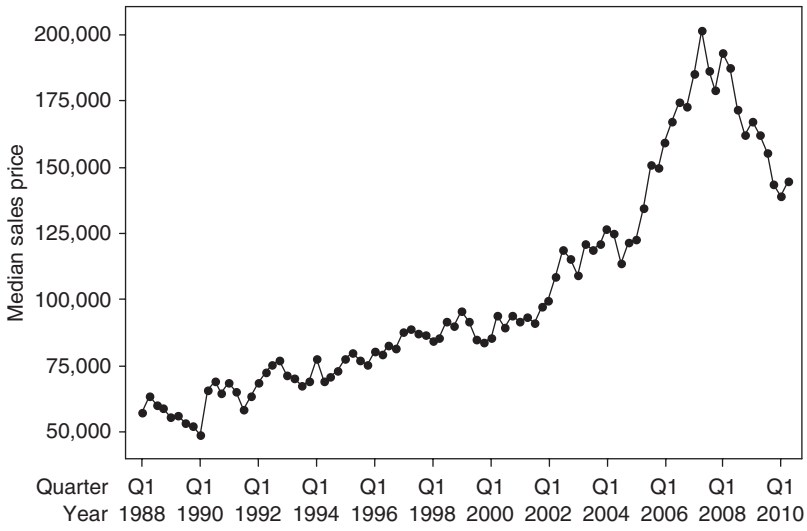

Figure 1.2 Median sales prices of houses in the United States. *Source:* US Bureau of the Census, http://www.census.gov/hhes/www/housing/hvs/historic/index.html.

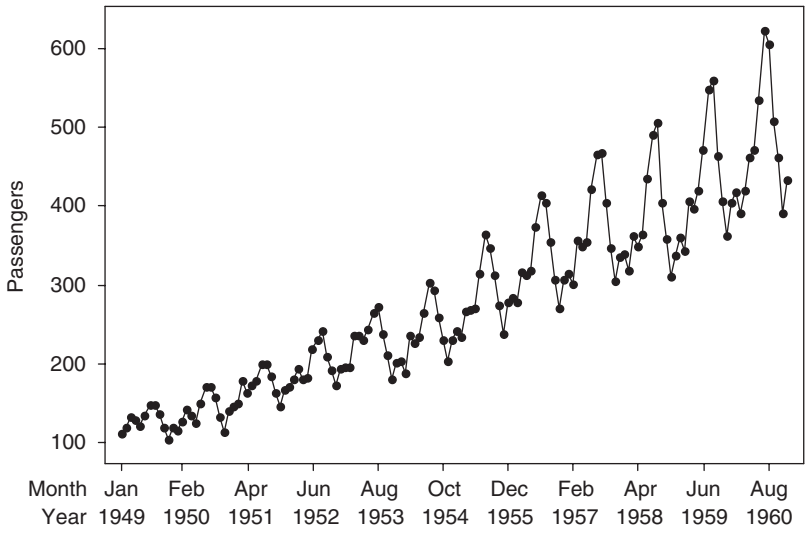

Figure 1.3 The number of airline passengers from 1949 to 1960.

In Figure 1.4, the quarterly dollar sales (in $1000) data of Marshall Field & Company for the period 1960 through 1975 also shows a seasonal pattern. The obvious increase in sales in the fourth quarter can certainly be attributed to Christmas shopping sprees. For inventory problems, for example, this type of data contains invaluable information. The data is taken from George Foster's

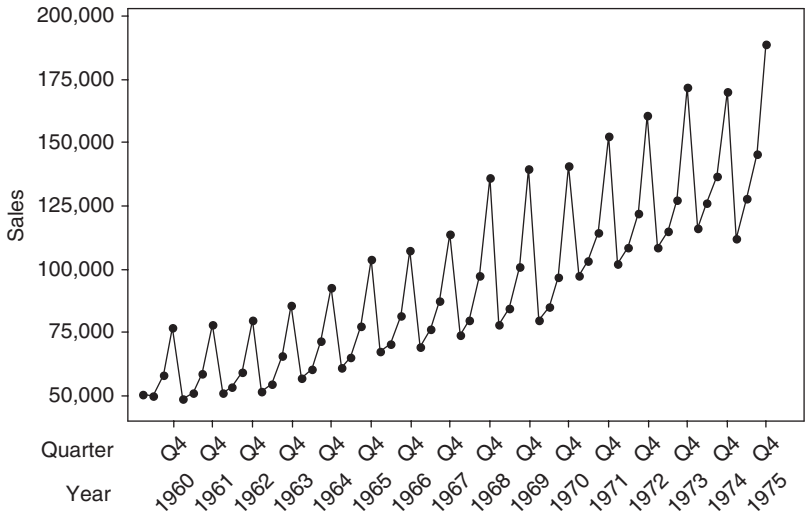

Figure 1.4 Quarterly dollar sales (in $1000) of Marshall Field & Company for the period 1960 through 1975.

Financial Statement Analysis (1978), where Foster uses this dataset in Chapter 4 to illustrate a number of statistical tools that are useful in accounting.

In some cases, it may also be possible to identify certain leading indicators for the variables of interest. For example, building permit applications is a leading indicator for many sectors of the economy that are influenced by construction activities. In Figure 1.5, the leading indicator is shown in the top panel whereas

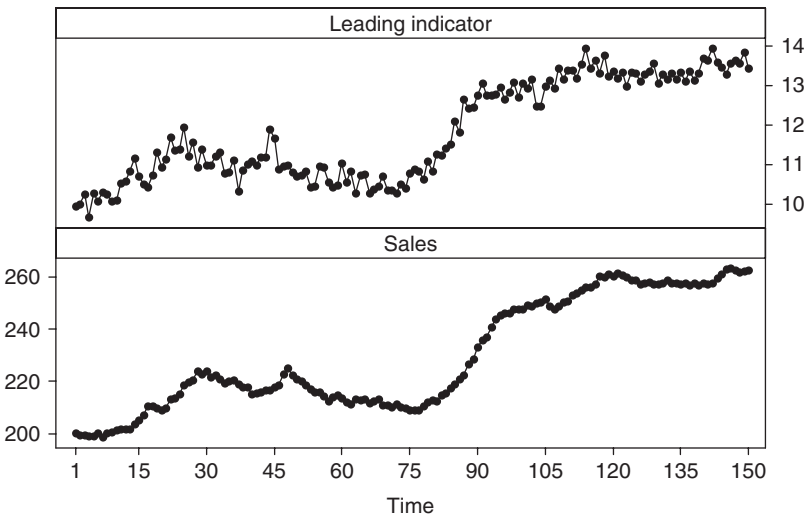

Figure 1.5 Time series plots of sales and a leading indicator.

the sales data is given at the bottom. They exhibit similar behavior; however, the important task is to find out whether there exists a lagged relationship between these two time series. If such a relationship exists, then from the current and past behavior of the leading indicator, it may be possible to determine how the sales will behave in the near future. This example will be studied in greater detail in Chapter 8.

Sometimes, the natural course of time series is interrupted because of some known causes such as public policy changes, strikes, new advertisement campaigns, and so on. In Chapter 8, the classic example of the market share fight between Colgate–Palmolive's "Colgate Dental Cream" and Proctor and Gamble's "Crest Toothpaste" will be discussed. Before the introduction of Crest by Proctor and Gamble into the US market, Colgate enjoyed a market leadership with a close to 50% market share. However, in 1960, the Council on Dental Therapeutics of the American Dental Association (ADA) endorsed Crest as an "important aid in any program of dental hygiene." Figure 1.6 shows the market shares of the two brands during the period before and after the endorsement. Now is it possible to deduce from this data that ADA's endorsement had any impact on the market shares? If so, was the effect permanent or temporary? In our analysis of these series in Chapter 8, some answers to these questions have been provided through an "intervention analysis."

This book also covers many engineering examples, most of which come from Box *et al*. (2008) (BJR hereafter). The time series plot of hourly temperature readings from a ceramic furnace is given in Figure 1.7. Even though the time interval considered consists of only 80 observations, the series looks stationary in the sense that both the mean and the variance do not seem to vary over time. The analysis of this series has been performed in Chapter 4.

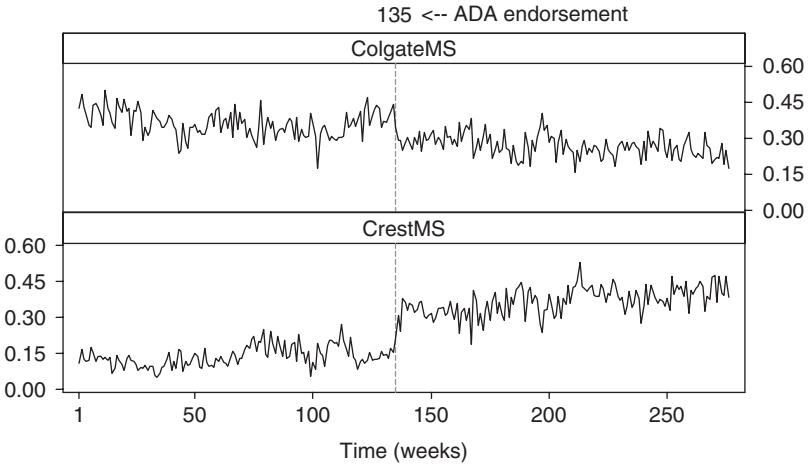

Figure 1.6 Time series plot of the weekly Colgate market share (ColgateMS) and Crest market share (CrestMS).

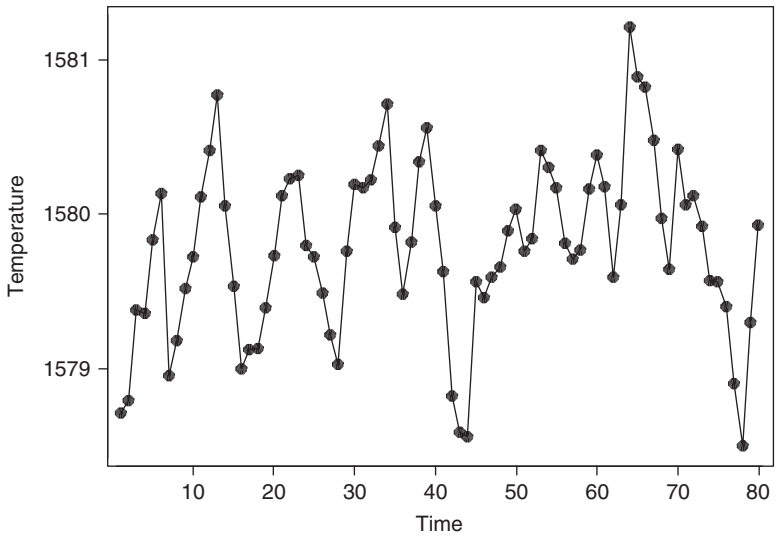

Figure 1.7 A time series plot of 80 consecutive hourly temperature observations from a ceramic furnace.

Figures 1.8 and 1.9 show the concentration and temperature readings, respectively, of a chemical process. The data come from series A and C of BJR. Both series exhibit nonstationary behavior in the sense that the mean seems to vary over time. This is to be expected from many engineering processes

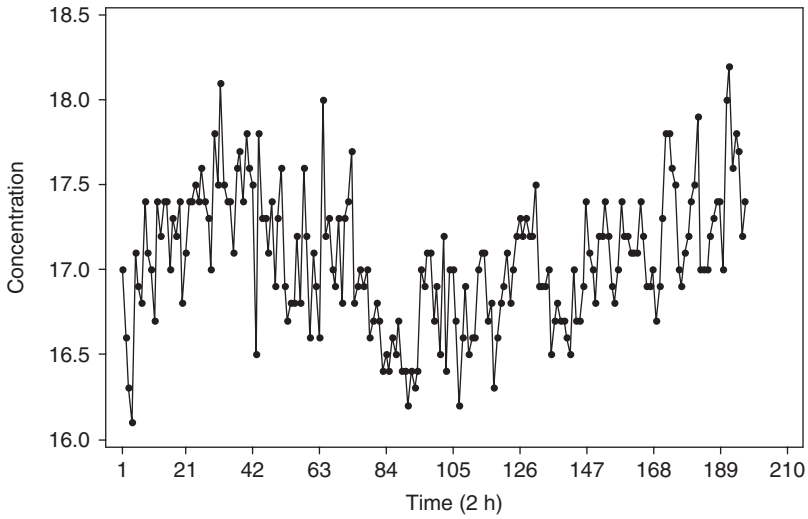

Figure 1.8 Time series plot of chemical process concentration readings sampled every 2 h (BJR series A).

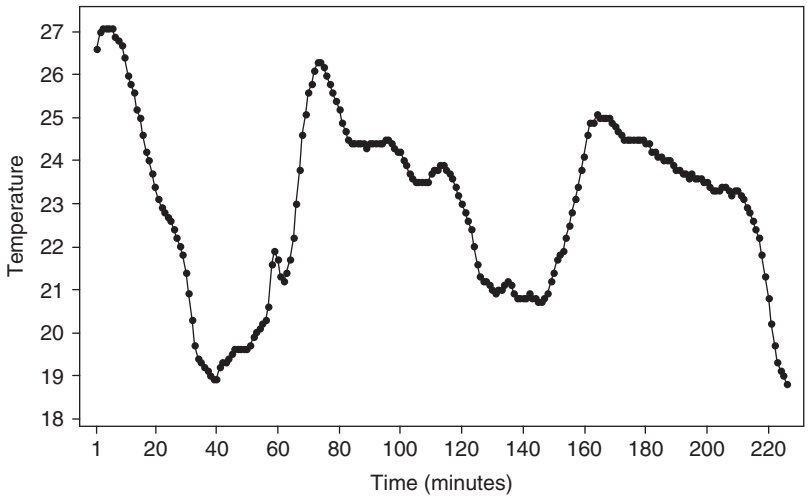

Figure 1.9 Time series plot of the temperature from a pilot plant observed every minute (BJR series C).

that are not tightly controlled. The analysis of both series has been provided in Chapter 4.

Data for another engineering example is given in the series J of BJR where the dynamic relationship between the input variable, methane gas rate, and the output, CO_2 concentration, in a pilot plant is discussed. In Figure 1.10, we can observe an apparent relationship between these two variables but a rigorous analysis is needed to fit a so-called transfer function–noise model to quantify this

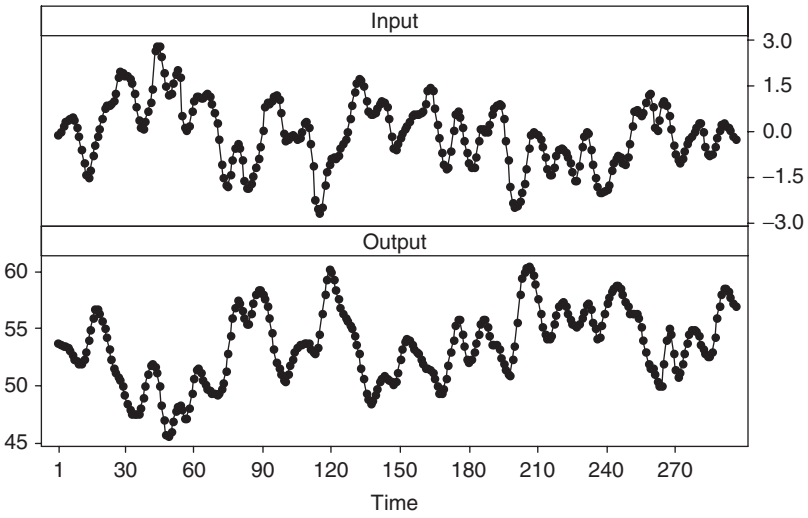

Figure 1.10 Time series plots of gas furnace data (BJR series J).

relationship. This is one of the examples used in Chapter 8 to illustrate some of the finer points in transfer function–noise models.

Time series data is of course not limited to economics, finance, business, and engineering. There are several other fields where the data is collected as a sequence in time and shows serial dependence. Consider the number of internet users over a 100-min period given in Figure 1.11. The data clearly does not follow a local mean but wanders around showing signs of "nonstationarity." This data is used in Chapter 6 to discuss how seemingly different models can fit a dataset equally well.

Figure 1.12 shows the annual sea level data for Copenhagen, Denmark, from 1889 to 2006. The data seems to have a stationary behavior with a subtle increase during the last couple of decades. What can city officials expect in the near future when it comes to sea levels rising? Can we make any generalizations regarding the sea levels all around the world based on this data? The data is available at www.psmsl.org. It is interesting to observe that the behavior we see in Figure 1.12 is only one of many different behaviors exhibited by similar datasets collected at various locations around the world. Note that in Figure 1.12, we observe missing data points, which is a surprisingly common problem with this type of data, and hence provides an excellent example to discuss the missing observations issue in Chapter 7.

There are also many examples in healthcare where time series data is collected and analyzed. In the fall of 2009, H1N1 flu pandemic generated a lot of fear throughout the world. The plot of the weekly number of reported cases in the United States is given in Figure 1.13. On the basis of this data, can we predict the number of flu cases in the autumn of 2010 and winter of 2011? What could

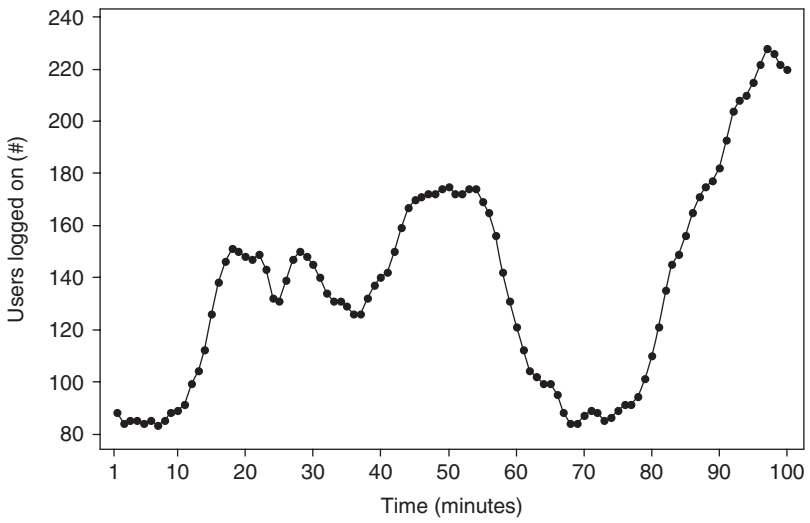

Figure 1.11 Time series plot of the number of internet server users over a 100-min period.

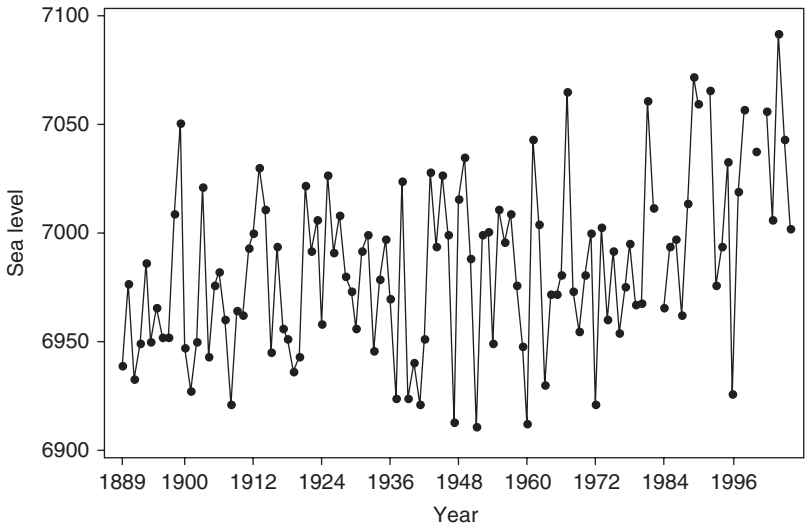

Figure 1.12 The annual sea levels in millimeters for Copenhagen, Denmark. *Source:* www.psmsl.org.

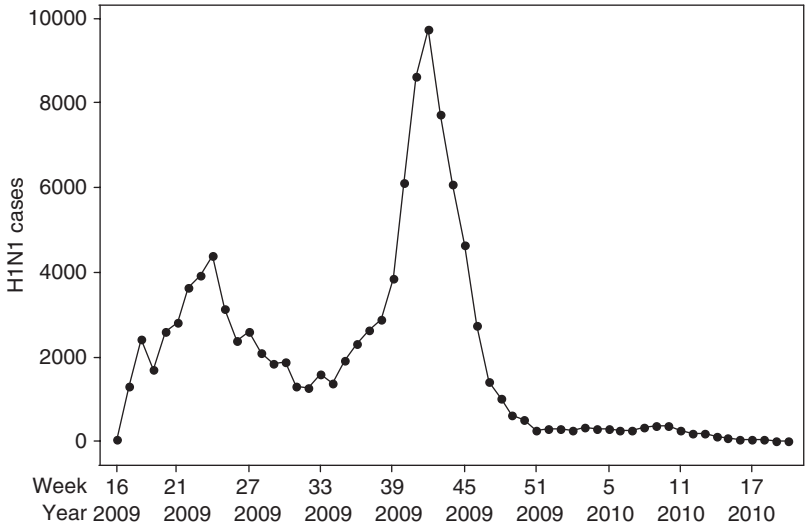

Figure 1.13 H1N1 flu cases in the United States from week 16 of 2009 to week 20 of 2010. *Source:* US Center for Disease Control CDC.

be the reason for a considerable decline in the number of cases at the end of 2009—a successful vaccination campaign, a successful "wash your hands" campaign, or people's improved immune system? Needless to say, an appropriate analysis of this data can greatly help to better prepare for the new flu season.

These examples can be extended to many other fields. The common thread is the data that is collected in time exhibiting a certain behavior, implying serial dependence. The tools and methodologies presented in this book will, in many cases, be proved very useful in identifying underlying patterns and dynamics in a process and allow the analyst to make sensible forecasts about its future behavior.

1.3 UNDERSTANDING AUTOCORRELATION

Modern time series modeling dates back to 1927 when the statistician G. U. Yule published an article where he used the dynamic movement of a pendulum as the inspiration to formulate an *autoregressive* model for the time dependency in an observed time series. We now demonstrate how Yule's pendulum analogue is an excellent vehicle for gaining intuition more generally about the dynamic behavior of time series models.

First, let us review the basic physics of the pendulum shown in Figure 1.14. If a pendulum in equilibrium with mass m under the influence of gravity is suddenly hit by a single impulse force, it will begin to swing back and forth. Yule describes this as a simple pendulum that is in equilibrium in the middle of the room, being pelted by peas thrown by some naughty boys in the room. This of course causes the harmonic motion that the pendulum displays subsequently. The frequency of this harmonic motion depends on the length of the pendulum,

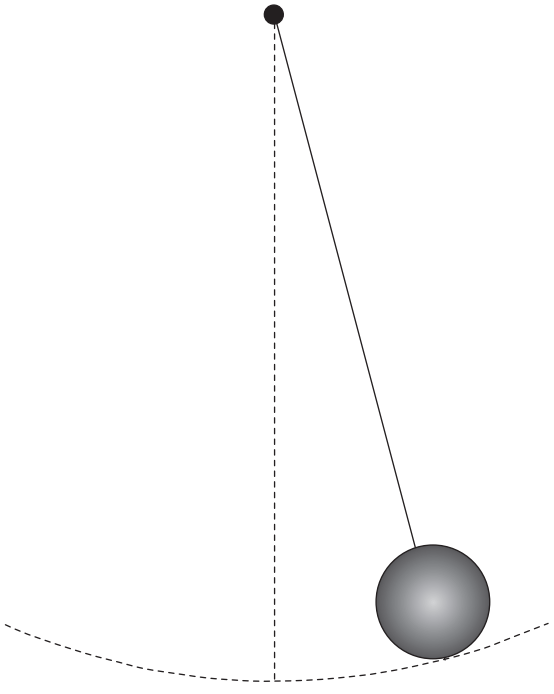

Figure 1.14 A simple pendulum.

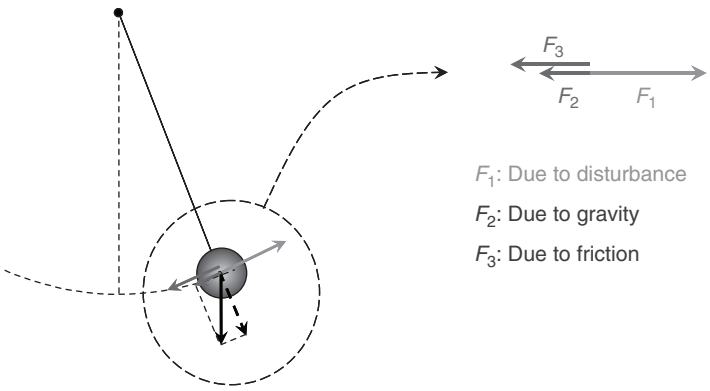

Figure 1.15 A simple pendulum in motion.

the amplitude of the mass of the bob, the impulse force, and the dissipative forces of friction and viscosity of the surrounding medium. The forces affecting a pendulum in motion are given in Figure 1.15.

After the initial impulse, the pendulum will gradually be slowed down by the dissipative forces until it eventually reaches the equilibrium again. How this happens provides an insight into the dynamic behavior of the pendulum—is it a short or long pendulum, is the bob light or heavy, is the friction small or large, and is the pendulum swinging in air or in a more viscous medium such as water?

An example for the displacement of the pendulum referenced to the equilibrium position at 0 is given in Figure 1.16. The harmonic movement $z(t)$ of a pendulum as a function of time t can be at least approximately described by a

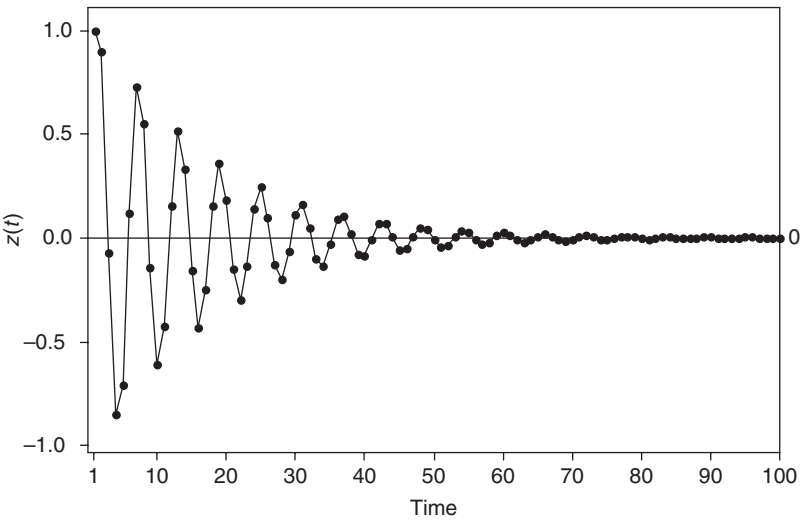

Figure 1.16 The displacement of a simple pendulum in motion.

second order linear differential equation with constant coefficients

$$m\frac{d^2z}{dt^2} + \gamma\frac{dz}{dt} + kz = a\delta(t) \tag{1.1}$$

where $\delta(t)$ is an impulse (delta) function that, like a pea shot, at time $t = 0$ forces the pendulum away from its equilibrium and a is the size of the impact by the pea. It is easy to imagine that the curve traced by this second order differential equation is a damped sinusoidal function of time although, if the friction or viscosity is sufficiently large, the (overdamped) pendulum may gradually come to rest following an exponential curve without ever crossing the centerline.

Differential equations are used to describe the dynamic process behavior in continuous time. But time series data is typically sampled (observed) at discrete times—for example, every hour or every minute. Yule therefore showed that if we replace the first- and second order differentials with discrete first- and second order differences, $\nabla z_t = z_t - z_{t-1}$ and $\nabla^2 z_t = \nabla(\nabla z_t) = z_t - 2z_{t-1} + z_{t-2}$, we can rewrite Equation (1.1) as a second order difference equation $\beta_2\nabla^2\tilde{z}_t + \beta_1\nabla\tilde{z}_t + \beta_0\tilde{z}_t = a_t$ where a_t mimics a random pea shot at time t and $\tilde{z}_t = z_t - \mu$ is the deviation from the pendulum's equilibrium position. After simple substitutions and rearrangements, this can be written as

$$\tilde{z}_t = \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + a_t \tag{1.2}$$

which is called a *second order autoregressive time series model* where the current observation $\tilde{z}_t$ is regressed on the two previous observations $\tilde{z}_{t-1}$ and $\tilde{z}_{t-2}$ and the error term is a_t. Therefore, if observed in discrete time, the oscillatory behavior of a pendulum can be described by Equation (1.2).

The model in Equation (1.2) is called an *autoregressive model* as the position of the pendulum at any given time t can be modeled using the position of the same pendulum at times $t - 1$ and $t - 2$. Borrowing the standard linear regression terminology, this model corresponds to the one where the position of the pendulum at any given time is (auto)-regressed onto itself at previous times. The reason that the model uses only two positions that are immediately preceding the current time is that the governing physics of the behavior of a simple pendulum dictates that it should follow second order dynamics. We should not expect all systems to follow the same second order dynamics. Nor do we expect to have a prior knowledge or even a guess of such dynamics for any given system. Therefore, empirical models where the current value is modeled using the previous values of appropriate lags are deemed appropriate for modeling time series data. The determination of the "appropriate" lags will be explored in the following chapters.

1.4 THE WOLD DECOMPOSITION

It is possible to provide another intuitive interpretation of Equation (1.2). The current position is given as a function of not only two previous positions but also of the current disturbance, a_t. This is, however, an incomplete account of what is going on here. If Equation (1.2) is valid for $\tilde{z}_t$, it should also be valid for $\tilde{z}_{t-1}$ and

$\tilde{z}_{t-2}$ for which a_{t-1} and a_{t-2} will be used respectively as the disturbance term in Equation (1.2). Therefore, the equation for $\tilde{z}_t$ does not only have a_t on the right-hand side but also a_{t-1} and a_{t-2} through the inclusion of the autoregressive terms $\tilde{z}_{t-1}$ and $\tilde{z}_{t-2}$. Using the same argument, we can further show that the equation for $\tilde{z}_t$ contains all previous disturbances. In fact, the powers of the coefficients in front of the autoregressive terms in Equation (1.2), namely ϕ_1 and ϕ_2, serve as the "weights" of these past disturbances. Therefore, certain coefficients can lead to an unstable infinite sum as these weights can increase exponentially as we move back in the past. For example, consider $+2$ and $+3$ for ϕ_1 and ϕ_2, respectively. This combination will give exponentially increasing weights for the past disturbances. Hence, only certain combinations of the coefficients will provide stable behavior in the weights and lead to a *stationary* time series. Indeed, stationary time series provide the foundation for discussing more general time series that exhibit trend and seasonality later in this book.

Stationary time series are characterized by having a distribution that is independent of time shifts. Most often, we will only require that the mean and variance of these processes are constant and that autocorrelation is only lag dependent. This is also called *weak stationarity*.

Now that we have introduced stationarity, we can also discuss one of the most fundamental results of modern time series analysis, the Wold decomposition theorem (see BJR). It essentially shows that any stationary time series process can be written as an infinite sum of weighted random shocks

$$\begin{aligned}\tilde{z}_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} \end{aligned} \tag{1.3}$$

where $\tilde{z}_t = z_t - \mu$ is the deviation from the mean, a_t's are uncorrelated random shocks with zero mean and constant variance, and $\{\psi_i\}$ satisfies $\sum_{i=0}^{+\infty} \psi_i^2 < \infty$. For most practical purposes, the Wold decomposition involving an infinite sum and an infinite number of parameters ψ_j is mostly of theoretical interest but not very useful in practice. However, we can often generate the ψ_j's from a few parameters. For example, if we let $\psi_j = \phi_1^j$, we can generate the entire infinite sequence of ψ_j's as the powers of a single parameter ϕ_1. It should be noted that although this imposes a strong restriction on the otherwise unrelated ψ_j's, it also allows us to represent infinitely many parameters with only one. Moreover, for most processes encountered in practice, most of the ψ_j weights will be small and without much consequence except for a relatively small number related to the most recent a_t's. Indeed, one of the essential ideas of the groundbreaking Box–Jenkins approach to time series analysis (see BJR) was their recognition that it was possible to approximate a wide variety of ψ weight patterns occurring in practice using models with only a few parameters. It is this idea of "parsimonious" models that led them to introduce the autoregressive moving average (ARMA) models that will be discussed in great detail in Chapter 3.

It should also be noted that while the models for stationary time series, such as the ARMA models, constitute the foundation of many methodologies we present in this book, the assumption that a time series is stationary is quite unrealistic in real life. For a system to exhibit a stationary behavior, it has to be tightly controlled and maintained in time. Otherwise, systems will tend to drift away from a stationary behavior following the second law of thermodynamics, which, as George E. P. Box, one of the pioneers in time series analysis, would playfully state, dictates that everything goes to hell in a hand basket. What is much more realistic is to claim that the changes to a process, or the first difference, form a stationary process. And if that is not realistic, we may try to see if the changes of the changes, the second difference, form a stationary process. This observation is the basis for the very versatile use of time series models. Thus, as we will see in later chapters, simple manipulations such as taking the first difference, $\nabla z_t = z_t - z_{t-1}$ or $\nabla^2 z_t = \nabla(z_t - z_{t-1}) = z_t - 2z_{t-1} + z_{t-2}$, can make those first- or second order differences exhibit stationary behavior even if z_t did not. This will be discussed in greater detail in Chapter 4.

1.5 THE IMPULSE RESPONSE FUNCTION

We have now seen that a stationary time series process can be represented as the dynamic response of a linear filter to a series of random shocks as illustrated in Figure 1.17. But what is the significance of the ψ_j's? The reason we are interested in the ψ_j weights is that they tell us something interesting about the dynamic behavior of a system. To illustrate this, let us return to the pendulum example. Suppose we, for a period of time, had observed a pendulum swinging back and forth, and found "coincidentally" that the parameters were $\hat{\phi}_1 = 0.9824$ and $\hat{\phi}_2 = -0.3722$. (Note that these estimates are from the example that will be discussed in Chapter 3.) Now, suppose the pendulum is brought to rest, but then at time $t = 0$ it is suddenly hit by a single small pea shot and then again left alone. The pendulum, of course, will start to swing but after some time it will eventually return to rest. But how much will it swing and for how long? If we knew that, we would have a feel for the type and size of pendulum we are dealing

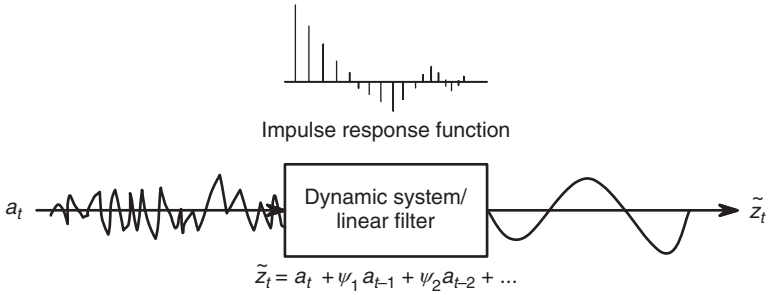

Figure 1.17 A time series model as a linear filter of random shock inputs.

with. In other words, we would be able to appreciate the dynamic behavior of the system under study, whether it is a pendulum, a ceramic furnace, the US economy, or something else. Fortunately, this question can directly be answered by studying the ψ_j's, also known as the *impulse response function*. Furthermore, the impulse response function can be computed easily with a spreadsheet program directly from the autoregressive model, $\tilde{z}_t = \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + a_t$.

Specifically, suppose we want to simulate that our pendulum is hit from the left with a single pea shot at time $t = 0$. Therefore, we let $a_0 = 1$ and $a_t = 0$ for $t > 0$. To get the computations started, suppose we start a few time units earlier, say $t = -2$. Since the pendulum is at rest, we set $z_{-1} = 0$ and $z_{-2} = 0$ and then recursively compute the responses as

$$\begin{aligned}
\tilde{z}_{-2} &= 0\\
\tilde{z}_{-1} &= 0\\
\tilde{z}_0 &= 0.9824\tilde{z}_{-1} - 0.3722\tilde{z}_{-2} + a_0 = 0.9824 \times 0 - 0.3722 \times 0 + 1 = 1\\
\tilde{z}_1 &= 0.9824\tilde{z}_0 - 0.3722\tilde{z}_{-1} + a_1 = 0.9824 \times 1 - 0.3722 \times 0 + 0 = 0.9824\\
\tilde{z}_2 &= 0.9824\tilde{z}_1 - 0.3722\tilde{z}_0 + a_2\\
&= 0.9824 \times 0.9824 - 0.3722 \times 1.0 + 0 = 0.59291\\
&\quad \text{and so on}
\end{aligned} \tag{1.4}$$

This type of recursive computation is easily set up in a spreadsheet. The response, $\tilde{z}_t, t = 1, 2, \ldots$, to the single impulse $a_0 = 1$ at $t = 0$ as it propagates through the system provides us with the ψ weights. The impulse response function is shown in Table 1.1 and plotted in Figure 1.18 where we see that the single pea shot causes the pendulum instantly to move to the right, then slowly returns back toward the centerline, crosses it at about $t = 4$, overshoots it a bit, again crosses the centerline about $t = 9$, and eventually comes to rest at about $t = 14$. In other words, our pendulum is relatively dampened as if it were moving in water or as if it were very long and had a heavy mass relative to the force of the small pea shot.

Now suppose we repeated the experiment with a much lighter and less damped pendulum with parameters $\phi_1 = 0.2$ and $\phi_2 = -0.8$.

The impulse response for this pendulum is shown in Figure 1.19. We see that it has a much more temperamental and oscillatory reaction to the pea shot and that the dynamic reaction stays much longer in the system.

1.6 SUPERPOSITION PRINCIPLE

The reaction of a linear filter model $\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots$ to a single pea shot has been discussed above. However, in general we will have a sequence of random shocks bombarding the system and not just a single shock. The reaction to each shock is given by the impulse response function. But for linear time series models, the reaction to a sequence of shocks can easily be generated by the superposition principle. That means, the individual responses can be added together

TABLE 1.1 The Impulse Response Function for the AR(2) for the Pendulum

Time (t)	a_t	$\tilde{z}_t = \psi_j$
−2	0	0.00000
−1	0	0.00000
0	1	1.00000
1	0	0.98240
2	0	0.59291
3	0	0.21683
4	0	−0.00767
5	0	−0.08824
6	0	−0.08383
7	0	−0.04951
8	0	−0.01744
9	0	0.00130
10	0	0.00776
11	0	0.00715
12	0	0.00413
13	0	0.00140
14	0	−0.00016
15	0	−0.00068

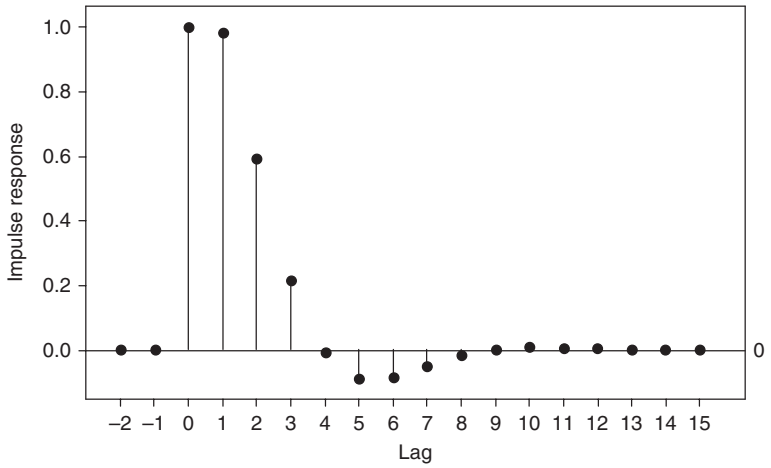

Figure 1.18 Impulse response function for the AR(2) for the pendulum.

to form the full response to a general sequence of inputs. Indeed, the impulse responses to each of the individual shocks are simply added up as they occur over time. For example, if the pendulum model $\tilde{z}_t = 0.9824\tilde{z}_{t-1} - 0.3722\tilde{z}_{t-2} + a_t$ was hit by a random sequence of 10 shocks as shown in Figure 1.20a starting

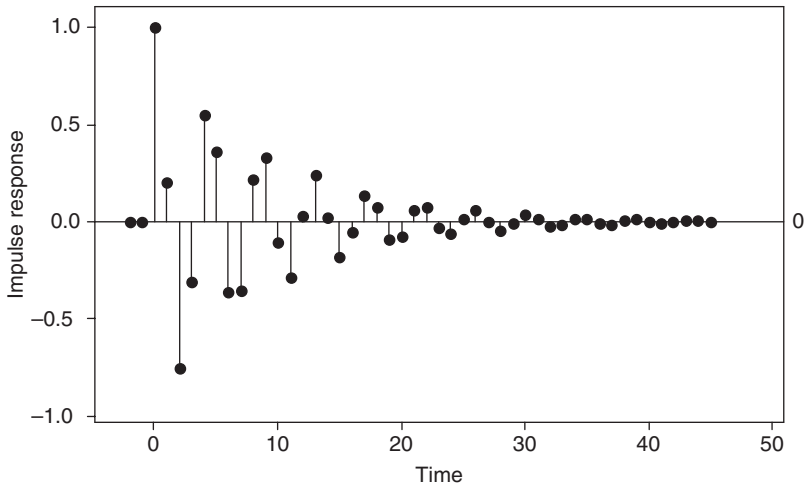

Figure 1.19 The impulse response for a pendulum with parameters $\phi_1 = 0.2$ and $\phi_2 = -0.8$.

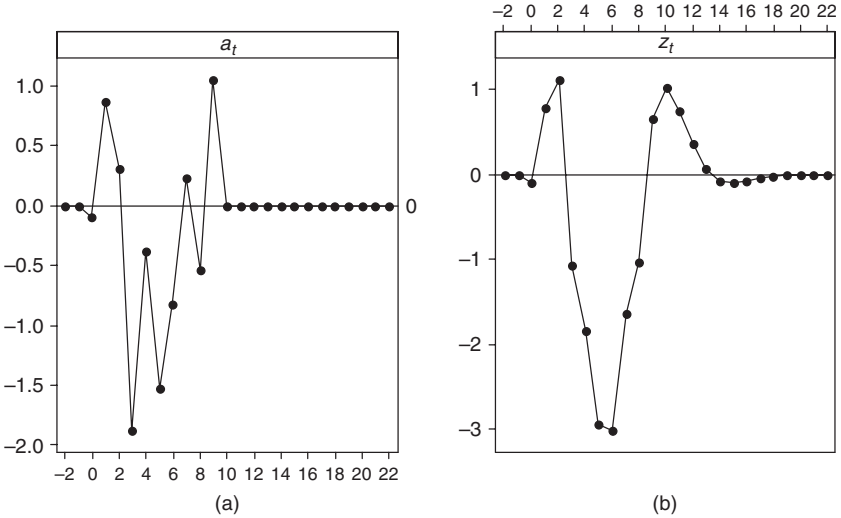

Figure 1.20 (a) Ten independent random white noise shocks $a_t, t = 1, \ldots, 10$ and (b) the superimposed responses of a linear filter generated by the AR(2) model $\tilde{z}_t = 0.9824\tilde{z}_{t-1} - 0.3722\tilde{z}_{t-2} + a_t$.

at time $t = 0$, then the pendulum's response over time would be as shown in Figure 1.20b.

The impulse response function helps us to visualize and gain an intuitive understanding of the dynamic reaction of a system. Specifically, we can consider any stationary time series model as a linear filter subject to a sequence of random shocks. How a process reacts to a single shock provides us with

important information about how the noise propagates through the system and what effect it has over time. Indeed, we can always intuitively think of any stationary time series model as a system that mimics the dynamic behavior of something like a pendulum subject to a sequence of small random pea shots. Further, if the process is nonstationary, what has been said above will apply to the first or possibly higher order difference of the data. In either case, the impulse response function is still a useful tool for visualizing the dynamic behavior of a system.

1.7 PARSIMONIOUS MODELS

In any modeling effort, we should always keep in mind that the model is only an approximation of the true behavior of the system in question. One of the cardinal sins of modeling is to fall in love with the model. As George Box famously stated, "All models are wrong. Some are useful." This is particularly true in time series modeling. There is quite a bit of personal judgment when it comes to determining the type of model we would like to use for a given data. Even though this interpretation adds extra excitement to the whole time series modeling process (we might admittedly be a bit biased when we say "excitement"), it also makes it subjective. When it comes to picking a model among many candidates, we should always keep in mind Occam's razor, which is attributed to philosopher and Franciscan friar William of Ockham (1285–1347/1349) who used it often in analyzing problems. In Latin it is "Pluralitas non est ponenda sine necessitate," which means "Plurality should not be posited without necessity" or "Entities are not to be multiplied beyond necessity." The principle was adapted by many scientists such as Nicole d'Oresme, a fourteenth century French physicist, and by Galileo in defending the simplest hypothesis of the heavens, the heliocentric system, or by Einstein who said "Everything should be made as simple as possible, but not simpler." In statistics, the application of this principle becomes obvious in modeling. Statistical models contain parameters that have to be estimated from the data. It is important to employ models with as few parameters as possible for adequate representation. Hence our principle should be, "When everything else is equal, choose the simplest model (. . . with the fewest parameters)." Why simpler models? Because they are easier to understand, easier to use, easier to interpret, and easier to explain. As opposed to simpler models, more complicated models with the prodigal use of parameters lead to poor estimates of the parameters. Models with large number of parameters will tend to overfit the data, meaning that locally they may provide very good fits; however, globally, that is, in forecasting, they tend to produce poor forecasts and larger forecast variances. Therefore, we strongly recommend the use of Occam's razor liberally in modeling efforts and always seek the simpler model when all else is the same.

EXERCISES

1.1 Discuss why we see serial dependence in data collected in time.

1.2 In the pendulum example given in Section 1.3, what are the factors that affect the serial dependence in the observations?

1.3 Find the Wold decomposition for the AR(2) model we obtain for the pendulum example.

1.4 The impulse response function in Table 1.1 is obtained when $a_0 = 1$. Repeat the same calculations with $a_0 = 1$ and $a_1 = 1$, and comment on your results.