CHAPTER 15

PROPERTIES OF STOCHASTIC TIME SERIES

In the last chapter we discussed a number of simple extrapolation techniques. In this chapter we begin our treatment of the construction and use of time-series models. Such models provide a more sophisticated method of extrapolating time series, in that they are based on the notion that the series to be forecasted has been generated by a *stochastic* (or *random*) *process*, with a structure that can be characterized and described. In other words, a time-series model provides a description of the random nature of the (stochastic) process that generated the sample of observations under study. The description is given not in terms of a cause-and-effect relationship (as would be the case in a regression model) but in terms of how that randomness is embodied in the process.

This chapter begins with an introduction to the nature of stochastic time-series models and shows how those models characterize the stochastic structure of the underlying process that generated the particular series. The chapter then turns to the properties of stochastic time series, focusing on the concept of *stationarity*. This material is important for the discussion of model construction in the following chapters. We next present a statistical test (the Dickey-Fuller test) for stationarity. Finally, we discuss *co-integrated* time series—series which are nonstationary but can be combined to form a stationary series.

15.1 INTRODUCTION TO STOCHASTIC TIME-SERIES MODELS

The time-series models developed in this and the following chapters are all based on an important assumption—that the series to be forecasted has been generated by a *stochastic process*. In other words, we assume that each value $y_1, y_2,$

$\ldots, y_T$ in the series is drawn randomly from a probability distribution. In modeling such a process, we attempt to describe the characteristics of its randomness. This should help us to infer something about the probabilities associated with alternative future values of the series.

To be completely general, we could assume that the observed series $y_1, \ldots, y_T$ is drawn from a set of *jointly distributed random variables*. If we could somehow numerically specify the probability distribution function for our series, then we could actually determine the probability of one or another future outcome.

Unfortunately, the complete specification of the probability distribution function for a time series is usually impossible. However, it usually is possible to construct a simplified model of the time series which explains its randomness in a manner that is useful for forecasting purposes. For example, we might believe that the values of $y_1, \ldots, y_T$ are normally distributed and are correlated with each other according to a simple first-order autoregressive process. The actual distribution might be more complicated, but this simple model may be a reasonable approximation. Of course, the usefulness of such a model depends on how closely it captures the true probability distribution and thus the true random behavior of the series. Note that it *need not* (and usually will not) match the actual past behavior of the series *since the series and the model are stochastic.* It should simply capture the characteristics of the series' randomness.

15.1.1 Random Walks

Our first (and simplest) example of a stochastic time series is the *random walk* process.[1] In the simplest random walk process, each successive *change* in y_t is drawn *independently* from a probability distribution with 0 mean. Thus, y_t is determined by

$$y_t = y_{t-1} + \varepsilon_t \tag{15.1}$$

with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t\varepsilon_s) = 0$ for $t \neq s$. Such a process could be generated by successive flips of a coin, where a head receives a value of $+1$ and a tail receives a value of -1.

Suppose we wanted to make a forecast for such a random walk process. The forecast is given by

$$\hat{y}_{T+1} = E(y_{T+1}|y_T, \ldots, y_1) \tag{15.2}$$

But $y_{T+1} = y_T + \varepsilon_{T+1}$ is independent of $y_{T-1}, \ldots, y_1$. Thus, the forecast one period ahead is simply

$$\hat{y}_{T+1} = y_T + E(\varepsilon_{T+1}) = y_T \tag{15.3}$$

[1] The random walk process has often been used as a model for the movement of stock market prices. See, for example, E. F. Fama, "Random Walks in Stock Market Prices," *Financial Analysts Journal*, September-October 1965.

The forecast two periods ahead is

$$\hat{y}_{T+2} = E(y_{T+2}|y_T, \ldots, y_1) = E(y_{T+1} + \varepsilon_{T+2})$$
$$= E(y_T + \varepsilon_{T+1} + \varepsilon_{T+2}) = y_T \qquad (15.4)$$

Similarly, the forecast l periods ahead is also y_T.

Although the forecast $\hat{y}_{T+l}$ will be the same no matter how large l is, the variance of the forecast error will grow as l becomes larger. For the one-period forecast, the forecast error is given by

$$e_1 = y_{T+1} - \hat{y}_{T+1}$$
$$= y_T + \varepsilon_{T+1} - y_T = \varepsilon_{T+1} \qquad (15.5)$$

and its variance is just $E(\varepsilon^2_{T+1}) = \sigma^2_\varepsilon$. For the two-period forecast,

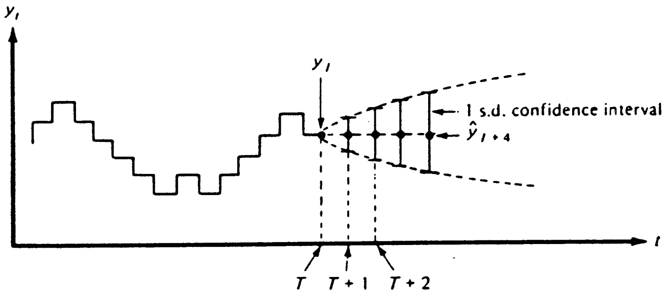
$$e_2 = y_{T+2} - \hat{y}_{T+2}$$
$$= y_T + \varepsilon_{T+1} + \varepsilon_{T+2} - y_T = \varepsilon_{T+1} + \varepsilon_{T+2} \qquad (15.6)$$

and its variance is

$$E[(\varepsilon_{T+1} + \varepsilon_{T+2})^2] = E(\varepsilon^2_{T+1}) + E(\varepsilon^2_{T+2}) + 2E(\varepsilon_{T+1}\varepsilon_{T+2}) \qquad (15.7)$$

Since ε_{T+1} and ε_{T+2} are independent, the third term in Eq. (15.7) is 0 and the error variance is $2\sigma^2_\varepsilon$. Similarly, for the l-period forecast, the error variance is $l\sigma^2_\varepsilon$. Thus, the *standard error of forecast* increases with the square root of l. We can thus obtain *confidence intervals* for our forecasts, and these intervals will become wider as the forecast horizon increases. This is illustrated in Fig. 15.1. Note that the forecasts are all equal to the last observation y_T, but the confidence intervals represented by 1 standard deviation in the forecast error increase as the square root of l.

FIGURE 15.1
Forecasting a random walk.



The fact that we can generate confidence intervals of this sort is an important advantage of stochastic time-series models. As we explained in Chapter 8, policy makers need to know the margin of error that has to be associated with a particular forecast, so confidence intervals can be as important as the forecasts themselves.

A simple extension of the random walk process discussed above is the random walk with drift. This process accounts for a trend (upward or downward) in the series y_t and thereby allows us to embody that trend in our forecast. In this process, y_t is determined by

$$y_t = y_{t-1} + d + \varepsilon_t \qquad (15.8)$$

so that on the average the process will tend to move upward (for $d > 0$). Now the one-period forecast is

$$\hat{y}_{T+1} = E(y_{T+1}|y_T, \ldots, y_1) = y_T + d \qquad (15.9)$$

and the l-period forecast is

$$\hat{y}_{T+l} = y_T + ld \qquad (15.10)$$

The standard error of forecast will be the same as before. For one period,

$$e_1 = y_{T+1} - \hat{y}_{T+1}$$
$$= y_T + d + \varepsilon_{T+1} - y_T - d = \varepsilon_{T+1} \qquad (15.11)$$

as before. The process, together with forecasts and forecast confidence intervals, is illustrated in Fig. 15.2. As can be seen in that figure, the forecasts increase linearly with l, and the standard error of forecast increases with the square root of l.

In the next chapter we examine a general class of stochastic time-series models. Later, we will see how that class of models can be used to make forecasts for a wide variety of time series. First, however, it is necessary to introduce some basic concepts about stochastic processes and their properties.

15.1.2 Stationary and Nonstationary Time Series

As we begin to develop models for time series, we want to know whether or not the underlying stochastic process that generated the series can be assumed to be *invariant with respect to time*. If the characteristics of the stochastic process change over time, i.e., if the process is *nonstationary*, it will often be difficult to represent the time series over past and future intervals of time by a simple algebraic model.[2] On the other hand, if the stochastic process is fixed in time, i.e., if it is

[2] The random walk with drift is one example of a nonstationary process for which a simple forecasting model can be constructed.

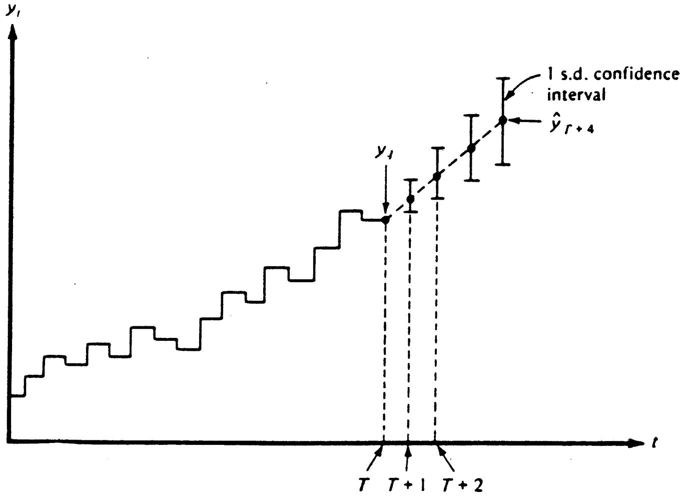


FIGURE 15.2
Forecasting a random walk with drift.

stationary, then one can model the process via an equation with fixed coefficients that can be estimated from past data. This is analogous to the single-equation regression model in which one economic variable is related to other economic variables, with coefficients that are estimated under the assumption that the structural relationship described by the equation is invariant over time (i.e., is stationary). If the structural relationship changed over time, we could not apply the techniques of Chapter 8 in using a regression model to forecast.

The models developed in detail in the next chapter of the book represent stochastic processes that are assumed to be in equilibrium about a constant mean level. The probability of a given fluctuation in the process from that mean level is assumed to be the same *at any point in time*. In other words, the stochastic properties of the stationary process are assumed to be *invariant with respect to time*.

One would suspect that many of the time series that one encounters in business and economics are not generated by stationary processes. The GNP, for example, has for the most part been growing steadily, and for this reason alone its stochastic properties in 1980 are different from those in 1933. Although it can be difficult to model nonstationary processes, we will see that nonstationary processes can often be transformed into stationary or approximately stationary processes.

15.1.3 Properties of Stationary Processes

We have said that any stochastic time series $y_1, \ldots, y_T$ can be thought of as having been generated by a set of jointly distributed random variables; i.e., the

set of data points $y_1, \ldots, y_T$ represents a particular outcome of the joint probability distribution function $p(y_1, \ldots, y_T)$.[3] Similarly, a *future* observation y_{T+1} can be thought of as being generated by a *conditional probability distribution function* $p(y_{T+1}|y_1, \ldots, y_T)$, that is, a probability distribution for y_{T+1} given the past observations $y_1, \ldots, y_T$. We define a *stationary* process, then, as one whose joint distribution and conditional distribution both are *invariant with respect to displacement in time*. In other words, if the series y_t is stationary, then

$$p(y_t, \ldots, y_{t+k}) = p(y_{t+m}, \ldots, y_{t+k+m}) \tag{15.12}$$

and

$$p(y_t) = p(y_{t+m}) \tag{15.13}$$

for any t, k, and m.

Note that if the series y_t is stationary, the *mean* of the series, defined as

$$\mu_y = E(y_t) \tag{15.14}$$

must also be stationary, so that $E(y_t) = E(y_{t+m})$, for any t and m. Furthermore, the *variance* of the series,

$$\sigma_y^2 = E[(y_t - \mu_y)^2] \tag{15.15}$$

must be stationary, so that $E[(y_t - \mu_y)^2] = E[y_{t+m} - \mu_y)^2]$, and finally, for any lag k, the *covariance* of the series,

$$\gamma_k = \text{Cov}\,(y_t, y_{t+k}) = E[(y_t - \mu_y)(y_{t+k} - \mu_y)] \tag{15.16}$$

must be stationary, so that $\text{Cov}\,(y_t, y_{t+k}) = \text{Cov}\,(y_{t+m}, y_{t+m+k})$.[4]

If a stochastic process is stationary, the probability distribution $p(y_t)$ is the same for all time t and its shape (or at least some of its properties) can be inferred by looking at a histogram of the observations $y_1, \ldots, y_T$ that make up the observed series. Also, an estimate of the mean μ_y of the process can be obtained from the *sample mean* of the series

$$\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t \tag{15.17}$$

[3] This outcome is called a *realization*. Thus $y_1, \ldots, y_T$ represent one particular realization of the stochastic process represented by the probability distribution $p(y_1, \ldots, y_T)$.

[4] It is possible for the mean, variance, and covariances of the series to be stationary but not the joint probability distribution. If the probability distributions are stationary, we term the series *strict-sense stationary*. If the mean, variance, and covariances are stationary, we term the series *wide-sense stationary*. Note that strict-sense stationarity implies wide-sense stationarity but that the converse is not true.

and an estimate of the variance σ_y^2 can be obtained from the *sample variance*

$$\hat{\sigma}_y^2 = \frac{1}{T}\sum_{t=1}^{T} (y_t - \bar{y})^2 \tag{15.18}$$

15.2 CHARACTERIZING TIME SERIES: THE AUTOCORRELATION FUNCTION

While it is usually impossible to obtain a complete description of a stochastic process (i.e., actually specify the underlying probability distributions), the autocorrelation function is extremely useful because it provides a partial description of the process for modeling purposes. The autocorrelation function tells us how much correlation there is (and by implication how much interdependency there is) between neighboring data points in the series y_t. We define the *autocorrelation with lag k* as

$$\rho_k = \frac{E[(y_t - \mu_y)(y_{t+k} - \mu_y)]}{\sqrt{E[(y_t - \mu_y)^2]E[(y_{t+k} - \mu_y)^2]}} = \frac{\text{Cov}\,(y_t, y_{t+k})}{\sigma_{y_t}\sigma_{y_{t+k}}} \tag{15.19}$$

For a stationary process the variance at time t in the denominator of Eq. (15.19) is the same as the variance at time $t + k$; thus the denominator is just the variance of the stochastic process, and

$$\rho_k = \frac{E[(y_t - \mu_y)(y_{t+k} - \mu_y)]}{\sigma_y^2} \tag{15.20}$$

Note that the numerator of Eq. (15.20) is the covariance between y_t and y_{t+k}, γ_k, so that

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{15.21}$$

and thus $\rho_0 = 1$ for *any* stochastic process.

Suppose that the stochastic process is simply

$$y_t = \varepsilon_t \tag{15.22}$$

where ε_t is an independently distributed random variable with zero mean. Then it is easy to see from Eq. (15.20) that the autocorrelation function for this process is given by $\rho_0 = 1$, $\rho_k = 0$ for $k > 0$. The process of Eq. (15.22) is called *white noise*, and there is no model that can provide a forecast any better than $\hat{y}_{T+l} = 0$

for all l. Thus if the autocorrelation function is zero (or close to zero) for all $k > 0$, there is little or no value in using a model to forecast the series.

Of course the autocorrelation function in Eq. (15.20) is purely theoretical, in that it describes a stochastic process for which we have only a limited number of observations. In practice, then, we must calculate an *estimate* of the autocorrelation function, called the *sample autocorrelation function:*

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2} \tag{15.23}$$

It is easy to see from their definitions that both the theoretical and estimated autocorrelation functions are symmetrical, i.e., that the correlation for a positive displacement is the same as that for a negative displacement, so that

$$\rho_k = \rho_{-k} \tag{15.24}$$

Then, when plotting an autocorrelation function (i.e., plotting ρ_k for different values of k), one need consider only positive values of k.

It is often useful to determine whether a particular value of the sample autocorrelation function $\hat{\rho}_k$ is close enough to zero to permit assuming that the *true* value of the autocorrelation function ρ_k is indeed equal to zero. It is also useful to test whether *all* the values of the autocorrelation function for $k > 0$ are equal to zero. (If they are, we know that we are dealing with white noise.) Fortunately, simple statistical tests exist that can be used to test the hypothesis that $\rho_k = 0$ for a particular k or to test the hypothesis that $\rho_k = 0$ for all $k > 0$.

To test whether a particular value of the autocorrelation function ρ_k is equal to zero we use a result obtained by Bartlett. He showed that if a time series has been generated by a *white noise* process, the sample autocorrelation coefficients (for $k > 0$) are approximately distributed according to a normal distribution with mean 0 and standard deviation $1/\sqrt{T}$ (where T is the number of observations in the series).[5] Thus, if a particular series consists of, say, 100 data points, we can attach a standard error of .1 to each autocorrelation coefficient. Therefore, if a particular coefficient was greater in magnitude than .2, we could be 95 percent sure that the true autocorrelation coefficient is not zero.

To test the *joint hypothesis* that *all* the autocorrelation coefficients are zero we use the Q statistic introduced by Box and Pierce. We will discuss this statistic in some detail in Chapter 17 in the context of performing diagnostic checks on

[5] See M. S. Bartlett, "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series," *Journal of the Royal Statistical Society*, ser. B8, vol. 27, 1946. Also see G. E. P. Box and G. M. Jenkins, *Time Series Analysis* (San Francisco: Holden-Day, 1970).

estimated time-series models, so here we only mention it in passing. Box and Pierce show that the statistic

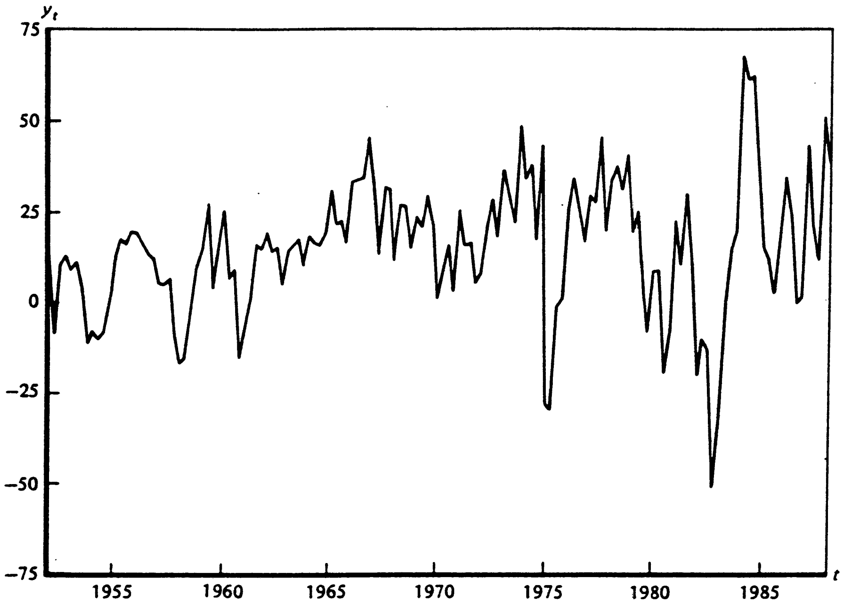
$$Q = T \sum_{k=1}^{K} \hat{\rho}_k^2 \tag{15.25}$$

is (approximately) distributed as chi square with K degrees of freedom. Thus if the calculated value of Q is greater than, say, the critical 5 percent level, we can be 95 percent sure that the *true* autocorrelation coefficients $\rho_1, \ldots, \rho_k$ are not all zero.

In practice people tend to use the critical 10 percent level as a cutoff for this test. For example, if Q turned out to be 18.5 for a total of $K = 15$ lags, we would observe that this is below the critical level of 22.31 and accept the hypothesis that the time series was generated by a white noise process.

Let us now turn to an example of an estimated autocorrelation function for a stationary economic time series. We have calculated $\hat{\rho}_k$ for quarterly data on real nonfarm inventory investment (measured in billions of 1982 dollars). The time series itself (covering the period 1952 through the first two quarters of 1988) is shown in Fig. 15.3, and the sample autocorrelation function is shown in Fig. 15.4. Note that the autocorrelation function falls off rather quickly as the lag k increases. This is typical of a stationary time series, such as inventory investment.

FIGURE 15.3
Nonfarm inventory investment (in 1982 constant dollars).



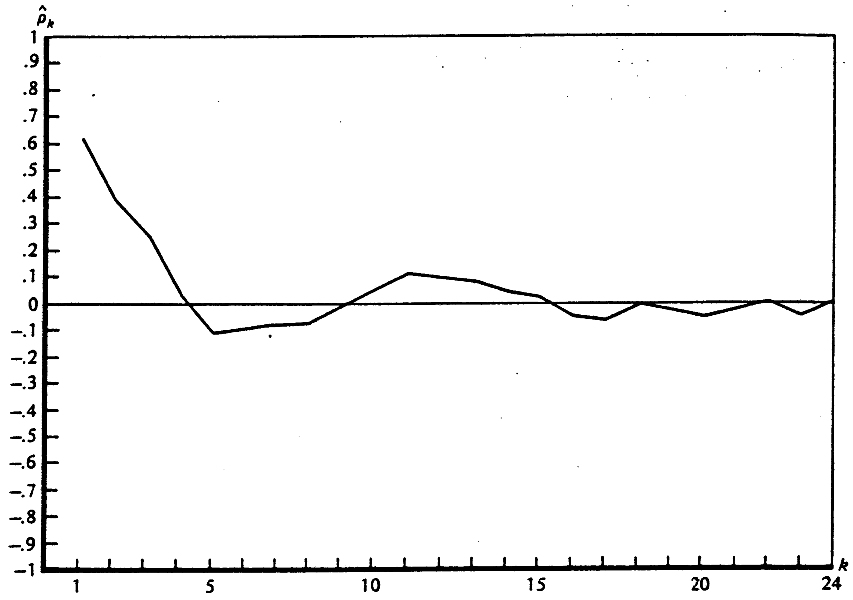


FIGURE 15.4
Nonfarm inventory investment: sample autocorrelation function.

In fact, as we will see, the autocorrelation function can be used to test whether or not a series is stationary. If $\hat{\rho}_k$ does not fall off quickly as k increases, this is an indication of nonstationarity. We will discuss more formal tests of nonstationarity ("unit root" tests) in Section 15.3.

If a time series is stationary, there exist certain analytical conditions which place bounds on the values that can be taken by the individual points of the autocorrelation function. However, the derivation of these conditions is somewhat complicated and will not be presented at this point. Furthermore, the conditions themselves are rather cumbersome and of limited usefulness in applied time-series modeling. Therefore, we have relegated them to Appendix 15.1. We turn our attention now to the properties of those time series which are nonstationary but which can be transformed into stationary series.

15.2.1 Homogeneous Nonstationary Processes

Probably very few of the time series one meets in practice are stationary. Fortunately, however, many of the nonstationary time series encountered (and this includes most of those that arise in economics and business) have the desirable property that if they are *differenced one or more times, the resulting series will be stationary*. Such a nonstationary series is termed *homogeneous*. The number of times that the original series must be differenced before a stationary series results is called the *order* of homogeneity. Thus, if y_t is first-order homogeneous nonstationary, the series

$$w_t = y_t - y_{t-1} = \Delta y_t \tag{15.26}$$

is stationary. If y_t happened to be second-order homogeneous, the series

$$w_t = \Delta^2 y_t = \Delta y_t - \Delta y_{t-1} \tag{15.27}$$

would be stationary.

As an example of a first-order homogeneous nonstationary process, consider the simple random walk process that we introduced earlier:

$$y_t = y_{t-1} + \varepsilon_t \tag{15.28}$$

Let us examine the variance of this process:

$$\begin{aligned}\gamma_0 &= E(y_t^2) = E[(y_{t-1} + \varepsilon_t)^2] = E(y_{t-1}^2) + \sigma_\varepsilon^2 \\ &= E(y_{t-2}^2) + 2\sigma_\varepsilon^2 \\ &\vdots\end{aligned} \tag{15.29}$$

or

$$\gamma_0 = E(y_{t-n}^2) + n\sigma_\varepsilon^2 \tag{15.30}$$

Observe from this recursive relation that the variance is infinite and hence undefined. The same is true for the covariances, since, for example,

$$\gamma_1 = E(y_t y_{t-1}) = E[y_{t-1}(y_{t-1} + \varepsilon_t)] = E(y_{t-1}^2) \tag{15.31}$$

Now let us look at the series that results from differencing the random walk process, i.e., the series

$$w_t = \Delta y_t = y_t - y_{t-1} = \varepsilon_t \tag{15.32}$$

Since the ε_t are assumed independent over time, w_t is clearly a stationary process. Thus, we see that the random walk process is first-order homogeneous. In fact, w_t is just a white noise process, and it has the autocorrelation function $\rho_0 = 1$, but $\rho_k = 0$ for $k > 0$.

15.2.2 Stationarity and the Autocorrelation Function

The GNP or a series of sales figures for a firm are both likely to be nonstationary. Each has been growing (on average) over time, so that the mean of each series is time-dependent. It is quite likely, however, that if the GNP or company sales figures are first-differenced one or more times, the resulting series will be stationary. Thus, if we want to build a time-series model to forecast the GNP, we can difference the series one or two times, construct a model for this new series, make our forecasts, and then *integrate* (i.e., undifference) the model and its forecasts to arrive back at GNP.

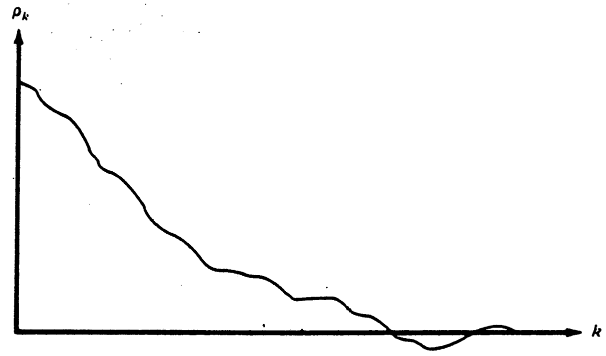
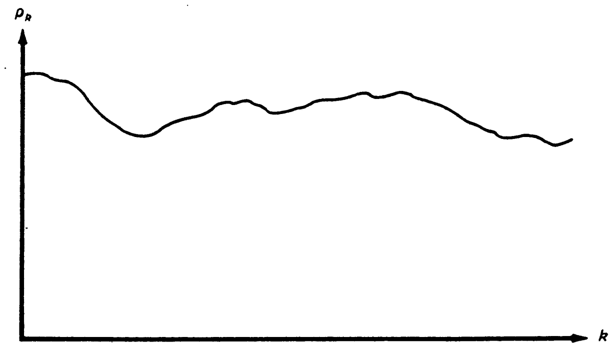


FIGURE 15.5
Stationary series.

How can we decide whether a series is stationary or determine the appropriate number of times a homogeneous nonstationary series should be differenced to arrive at a stationary series? We can begin by looking at a plot of the autocorrelation function (called a *correlogram*). Figures 15.5 and 15.6 show autocorrelation functions for stationary and nonstationary series. The autocorrelation function for a stationary series drops off as k, the number of lags, becomes large, but this is usually not the case for a nonstationary series. If we are differencing a nonstationary series, we can test each succeeding difference by looking at the autocorrelation function. If, for example, the second round of differencing

FIGURE 15.6
Nonstationary series.

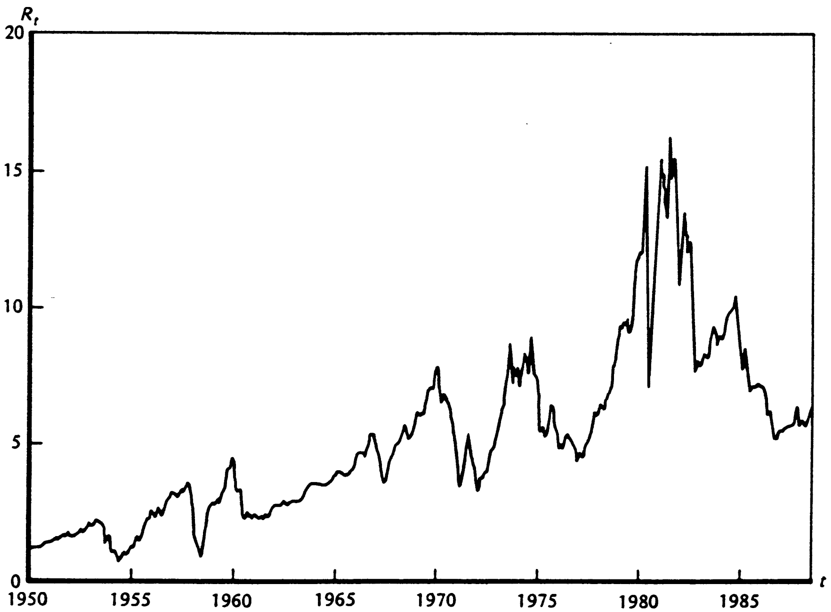


results in a series whose autocorrelation function drops off rapidly, we can determine that the original series is second-order homogeneous. If the resulting series is still nonstationary, the autocorrelation function will remain large even for long lags.

Example 15.1 Interest Rate Often in applied work it is not clear how many times a nonstationary series should be differenced to yield a stationary one, and one must make a judgment based on experience and intuition. As an example, we will examine the interest rate on 3-month government Treasury bills. This series, consisting of monthly data from the beginning of 1950 through June 1988, is shown in Fig. 15.7, and its autocorrelation function is shown in Fig. 15.8. The autocorrelation function does decline as the number of lags becomes large, but only very slowly. In addition, the series exhibits an upward trend (so that the mean is not constant over time). We would therefore suspect that this series has been generated by a homogeneous nonstationary process. To check, we difference the series and recalculate the sample autocorrelation function.

The differenced series is shown in Fig. 15.9. Note that the mean of the series is now about constant, although the variance becomes unusually high

FIGURE 15.7
Three-month Treasury bill rate.



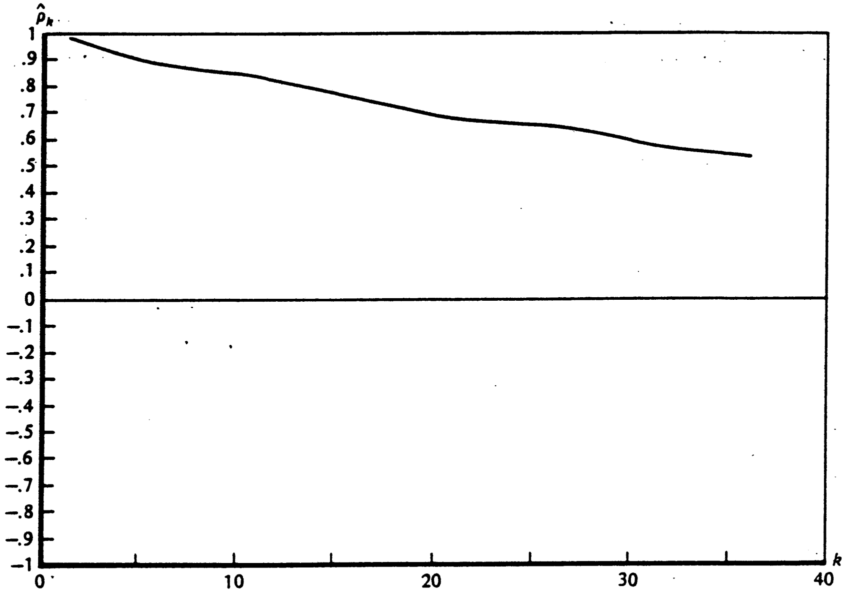
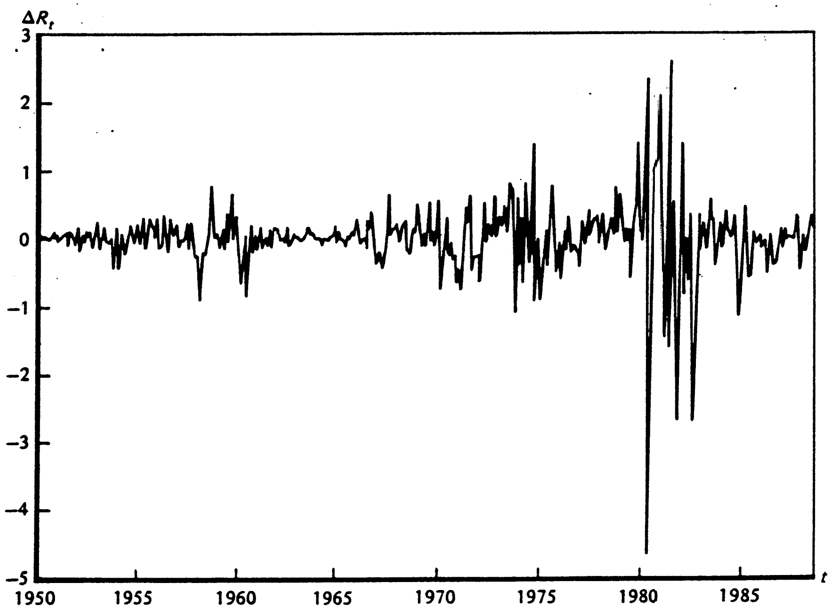


FIGURE 15.8
Three-month Treasury bill rate: sample autocorrelation function.

FIGURE 15.9
Three-month Treasury bill rate—first differences.



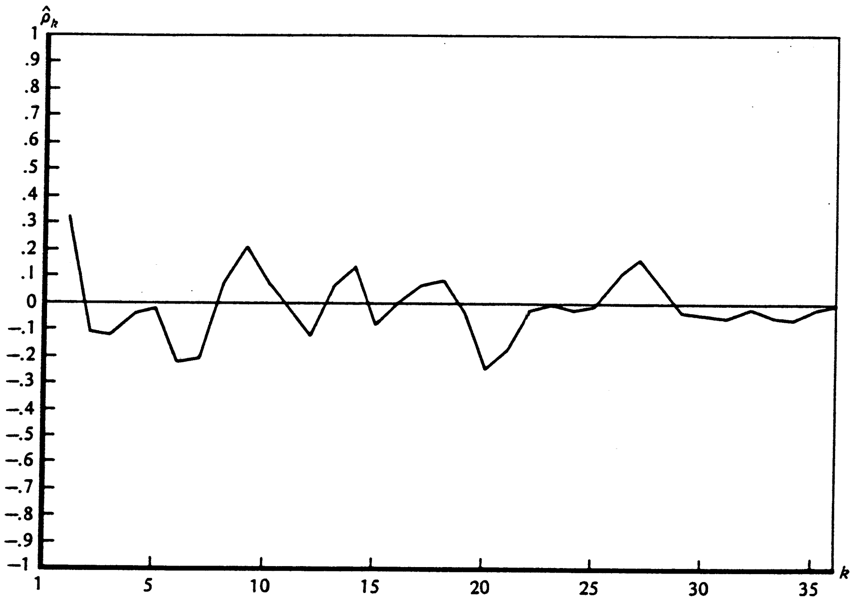


FIGURE 15.10
Interest rate—first differences: sample autocorrelation function.

during the early 1980s (a period when the Federal Reserve targeted the money supply, allowing interest rates to fluctuate). The sample autocorrelation function for the differenced series is shown in Fig. 15.10. It declines rapidly, consistent with a stationary series. We also tried differencing the series a second time. The twice-differenced series, $\Delta^2 R_t = \Delta R_t - \Delta R_{t-1}$, is shown in Fig. 15.11, and its sample autocorrelation function in Fig. 15.12. The results do not seem qualitatively different from the previous case. Our conclusion, then, would be that differencing once should be sufficient to ensure stationarity.

Example 15.2 Daily Hog Prices[6] As a second example, let us examine a time series for the daily market price of hogs. If a forecasting model could be developed for this series, one could conceivably make money by speculating on the futures market for hogs and using the model to outperform the market.

[6] This example is from a paper by R. Leuthold, A. MacCormick, A. Schmitz, and D. Watts, "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques," *Journal of the American Statistical Association*, March 1970, Applications Section, pp. 90–107.

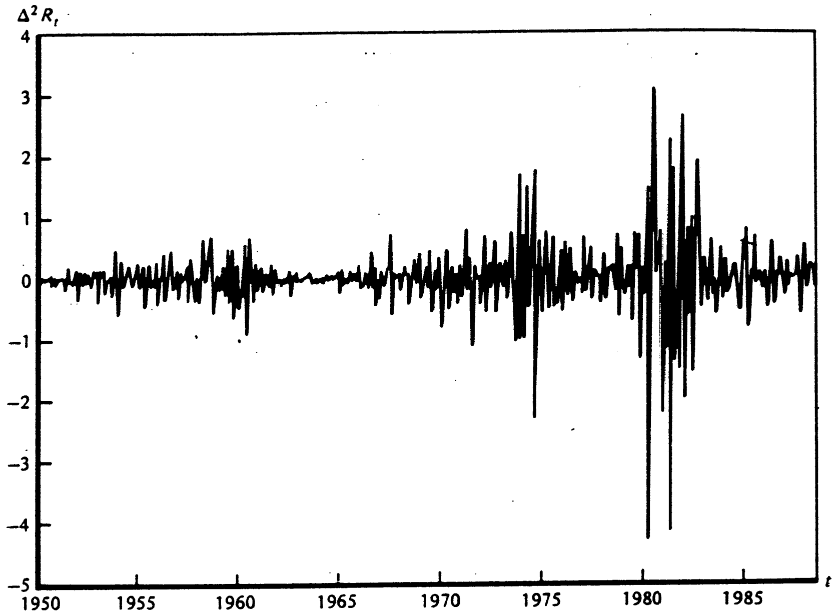
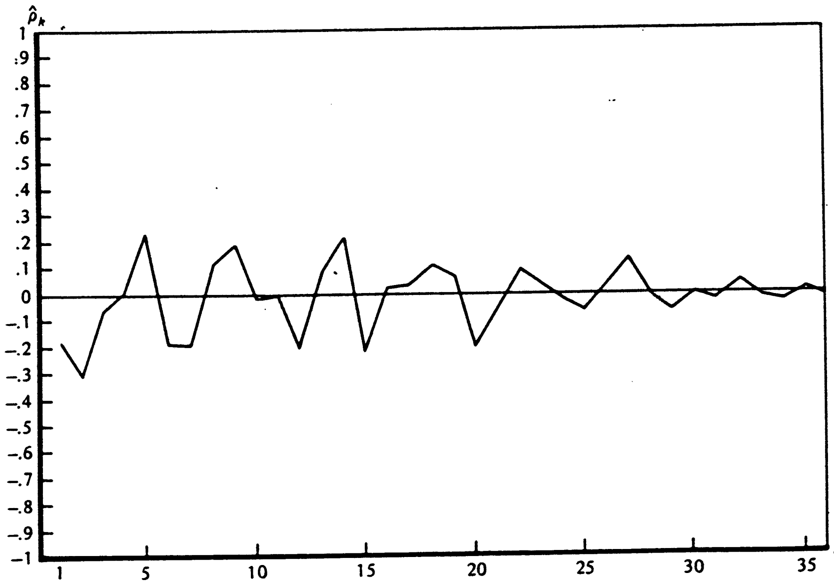


FIGURE 15.11
Three-month Treasury bill rate—second differences.

FIGURE 15.12
Interest rate—second differences: sample autocorrelation function.



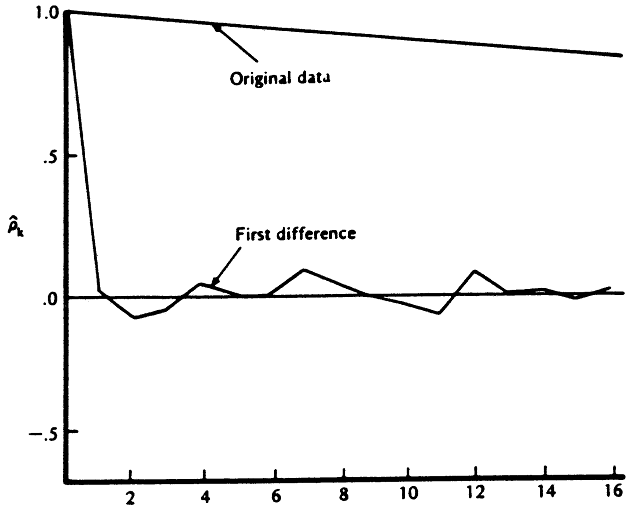


FIGURE 15.13
Sample autocorrelation functions of daily hog price data.

The series consists of 250 daily data points covering all the trading days in 1965. The price variable is the average price in dollars per hundredweight of all hogs sold in the eight regional markets in the United States on a particular day. The sample autocorrelation functions for the original price series and for the first difference of the series are shown in Fig. 15.13.

Observe that the original series is clearly nonstationary. The autocorrelation function barely declines, even after a 16-period lag. The series is, however, first-order homogeneous, since its first difference is clearly stationary.

In fact, not only is the first-differenced series stationary, but it appears to resemble white noise, since the sample autocorrelation function $\hat{\rho}_k$ is close to zero for all $k > 0$. To determine whether the differenced series is indeed white noise, let us calculate the Q statistic for the first 15 lags. The value of this statistic is 14.62, which, with 15 degrees of freedom, is insignificant at the 10 percent level. We can therefore conclude that the differenced series is white noise and that the original price series can best be modeled as a *random walk:*

$$P_t = P_{t-1} + \varepsilon_t \tag{15.33}$$

As is the case of most stock market prices, our best forecast of P_t is its most recent value, and (sadly) there is no model that can help us outperform the market.

15.2.3 Seasonality and the Autocorrelation Function

We have just seen that the autocorrelation function can reveal information about the stationarity of a time series. In the remaining chapters of this book we will see that other information about a time series can be obtained from its autocorrelation function. However, we continue here by examining the relationship between the autocorrelation function and the *seasonality* of a time series.

As discussed in the previous chapter, seasonality is just a cyclical behavior that occurs on a regular calendar basis. An example of a highly seasonal time series would be toy sales, which exhibit a strong peak every Christmas. Sales of ice cream and iced-tea mix show seasonal peaks each summer in response to increased demand brought about by warmer weather; Peruvian anchovy production shows seasonal troughs once every 7 years in response to decreased supply brought about by cyclical changes in the ocean currents.

Often seasonal peaks and troughs are easy to spot by direct observation of the time series. However, if the time series fluctuates considerably, seasonal peaks and troughs might not be distinguishable from the other fluctuations. Recognition of seasonality is important because it provides information about "regularity" in the series that can aid us in making a forecast. Fortunately, that recognition can be made easier with the help of the autocorrelation function.

If a monthly time series y_t exhibits annual seasonality, the data points in the series should show some degree of correlation with the corresponding data points which lead or lag by 12 months. In other words, we would expect to see some degree of correlation between y_t and y_{t-12}. Since y_t and y_{t-12} will be correlated, as will y_{t-12} and y_{t-24}, we should also see correlation between y_t and y_{t-24}. Similarly there will be correlation between y_t and y_{t-36}, y_t and y_{t-48}, etc. These correlations should manifest themselves in the sample autocorrelation function $\hat{\rho}_k$, which will exhibit peaks at k = 12, 24, 36, 48, etc. Thus we can identify seasonality by observing regular peaks in the autocorrelation function, even if seasonal peaks cannot be discerned in the time series itself.

Example 15.3 Hog Production As an example, look at the time series for the monthly production of hogs in the United States, shown in Fig. 15.14. It would take a somewhat experienced eye to easily discern seasonality in that series. The seasonality of the series, however, is readily apparent in its sample autocorrelation function, which is shown in Fig. 15.15. Note the peaks that occur at k = 12, 24, and 36, indicating annual cycles in the series.

A crude method of removing the annual cycles ("deseasonalizing" the data) would be to take a 12-month difference, obtaining a new series $z_t = y_t - y_{t-12}$. As can be seen in Fig. 15.16, the sample autocorrelation function for this 12-month differenced series does not exhibit strong seasonality. We will see in later chapters that z_t represents an extremely simple time-