CHAPTER 2

FUNDAMENTAL CONCEPTS

This chapter describes the fundamental concepts in the theory of time series models. In particular, we introduce the concepts of stochastic processes, mean and covariance functions, stationary processes, and autocorrelation functions.

2.1 Time Series and Stochastic Processes

The sequence of random variables $\{Y_t: t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$ is called a **stochastic process** and serves as a model for an observed time series. It is known that the complete probabilistic structure of such a process is determined by the set of distributions of all finite collections of the Y's. Fortunately, we will not have to deal explicitly with these multivariate distributions. Much of the information in these joint distributions can be described in terms of means, variances, and covariances. Consequently, we concentrate our efforts on these first and second moments. (If the joint distributions of the Y's are multivariate *normal* distributions, then the first and second moments completely determine all the joint distributions.)

2.2 Means, Variances, and Covariances

For a stochastic process $\{Y_t: t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$, the **mean function** is defined by

$$\mu_t = E(Y_t) \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots \tag{2.2.1}$$

That is, μ_t is just the expected value of the process at time t. In general, μ_t can be different at each time point t.

The **autocovariance function**, $\gamma_{t,s}$, is defined as

$$\gamma_{t,s} = Cov(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots \tag{2.2.2}$$

where $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$.

The **autocorrelation function**, $\rho_{t,s}$, is given by

$$\rho_{t,s} = Corr(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots \tag{2.2.3}$$

where

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \tag{2.2.4}$$

We review the basic properties of expectation, variance, covariance, and correlation in Appendix A on page 24.

Recall that both covariance and correlation are measures of the (linear) dependence between random variables but that the unitless correlation is somewhat easier to interpret. The following important properties follow from known results and our definitions:

$$\left.\begin{array}{ll} \gamma_{t,t} = Var(Y_t) & \rho_{t,t} = 1 \\ \gamma_{t,s} = \gamma_{s,t} & \rho_{t,s} = \rho_{s,t} \\ |\gamma_{t,s}| \le \sqrt{\gamma_{t,t}\gamma_{s,s}} & |\rho_{t,s}| \le 1 \end{array}\right\} \tag{2.2.5}$$

Values of $\rho_{t,s}$ near ± 1 indicate strong (linear) dependence, whereas values near zero indicate weak (linear) dependence. If $\rho_{t,s} = 0$, we say that Y_t and Y_s are *uncorrelated*.

To investigate the covariance properties of various time series models, the following result will be used repeatedly: If $c_1, c_2, \ldots, c_m$ and $d_1, d_2, \ldots, d_n$ are constants and $t_1, t_2, \ldots, t_m$ and $s_1, s_2, \ldots, s_n$ are time points, then

$$Cov\left[\sum_{i=1}^{m} c_i Y_{t_i}, \sum_{j=1}^{n} d_j Y_{s_j}\right] = \sum_{i=1}^{m} \sum_{j=1}^{n} c_i d_j Cov(Y_{t_i}, Y_{s_j}) \tag{2.2.6}$$

The proof of Equation (2.2.6), though tedious, is a straightforward application of the linear properties of expectation. As a special case, we obtain the well-known result

$$Var\left[\sum_{i=1}^{n} c_i Y_{t_i}\right] = \sum_{i=1}^{n} c_i^2 Var(Y_{t_i}) + 2\sum_{i=2}^{n} \sum_{j=1}^{i-1} c_i c_j Cov(Y_{t_i}, Y_{t_j}) \tag{2.2.7}$$

The Random Walk

Let $e_1, e_2, \ldots$ be a sequence of independent, identically distributed random variables each with zero mean and variance σ_e^2. The observed time series, $\{Y_t : t = 1, 2, \ldots\}$, is constructed as follows:

$$\left.\begin{array}{rl} Y_1 &= e_1 \\ Y_2 &= e_1 + e_2 \\ &\vdots \\ Y_t &= e_1 + e_2 + \cdots + e_t \end{array}\right\} \tag{2.2.8}$$

Alternatively, we can write

$$Y_t = Y_{t-1} + e_t \tag{2.2.9}$$

with "initial condition" $Y_1 = e_1$. If the e's are interpreted as the sizes of the "steps" taken (forward or backward) along a number line, then Y_t is the position of the "random walker" at time t. From Equation (2.2.8), we obtain the mean function

$$\begin{aligned}\mu_t &= E(Y_t) = E(e_1 + e_2 + \cdots + e_t) = E(e_1) + E(e_2) + \cdots + E(e_t) \\ &= 0 + 0 + \cdots + 0\end{aligned}$$

so that

$$\mu_t = 0 \quad \text{for all } t \tag{2.2.10}$$

We also have

$$\begin{aligned}Var(Y_t) &= Var(e_1 + e_2 + \cdots + e_t) = Var(e_1) + Var(e_2) + \cdots + Var(e_t) \\ &= \sigma_e^2 + \sigma_e^2 + \cdots + \sigma_e^2\end{aligned}$$

so that

$$Var(Y_t) = t\sigma_e^2 \tag{2.2.11}$$

Notice that the process variance increases linearly with time.

To investigate the covariance function, suppose that $1 \le t \le s$. Then we have

$$\gamma_{t,s} = Cov(Y_t, Y_s) = Cov(e_1 + e_2 + \cdots + e_t, e_1 + e_2 + \cdots + e_t + e_{t+1} + \cdots + e_s)$$

From Equation (2.2.6), we have

$$\gamma_{t,s} = \sum_{i=1}^{s} \sum_{j=1}^{t} Cov(e_i, e_j)$$

However, these covariances are zero unless $i = j$, in which case they equal $Var(e_i) = \sigma_e^2$. There are exactly t of these so that $\gamma_{t,s} = t\sigma_e^2$.

Since $\gamma_{t,s} = \gamma_{s,t}$, this specifies the autocovariance function for all time points t and s and we can write

$$\gamma_{t,s} = t\sigma_e^2 \qquad \text{for } 1 \le t \le s \tag{2.2.12}$$

The autocorrelation function for the random walk is now easily obtained as

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \sqrt{\frac{t}{s}} \qquad \text{for } 1 \le t \le s \tag{2.2.13}$$

The following numerical values help us understand the behavior of the random walk.

$$\rho_{1,2} = \sqrt{\frac{1}{2}} = 0.707 \qquad \rho_{8,9} = \sqrt{\frac{8}{9}} = 0.943$$

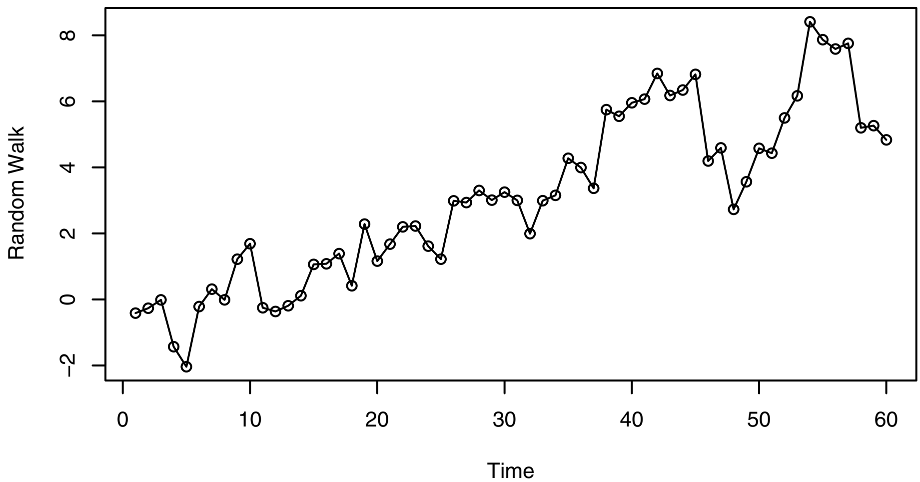
$$\rho_{24,25} = \sqrt{\frac{24}{25}} = 0.980 \qquad \rho_{1,25} = \sqrt{\frac{1}{25}} = 0.200$$

The values of Y at neighboring time points are more and more strongly and positively correlated as time goes by. On the other hand, the values of Y at distant time points are less and less correlated.

A simulated random walk is shown in Exhibit 2.1 where the e's were selected from a standard normal distribution. Note that even though the theoretical mean function is zero for all time points, the fact that the variance increases over time and that the correlation between process values nearby in time is nearly 1 indicate that we should expect long excursions of the process away from the mean level of zero.

The simple random walk process provides a good model (at least to a first approximation) for phenomena as diverse as the movement of common stock price, and the position of small particles suspended in a fluid—so-called Brownian motion.

Exhibit 2.1 Time Series Plot of a Random Walk




```
> win.graph(width=4.875, height=2.5,pointsize=8)
> data(rwalk) # rwalk contains a simulated random walk
> plot(rwalk,type='o',ylab='Random Walk')
```

A Moving Average

As a second example, suppose that $\{Y_t\}$ is constructed as

$$Y_t = \frac{e_t + e_{t-1}}{2} \tag{2.2.14}$$

where (as always throughout this book) the e's are assumed to be independent and identically distributed with zero mean and variance σ_e^2. Here

$$\begin{aligned}\mu_t &= E(Y_t) = E\left\{\frac{e_t + e_{t-1}}{2}\right\} = \frac{E(e_t) + E(e_{t-1})}{2} \\ &= 0\end{aligned}$$

and

$$Var(Y_t) = Var\left\{\frac{e_t + e_{t-1}}{2}\right\} = \frac{Var(e_t) + Var(e_{t-1})}{4}$$

$$= 0.5\sigma_e^2$$

Also

$$Cov(Y_t, Y_{t-1}) = Cov\left\{\frac{e_t + e_{t-1}}{2}, \frac{e_{t-1} + e_{t-2}}{2}\right\}$$

$$= \frac{Cov(e_t, e_{t-1}) + Cov(e_t, e_{t-2}) + Cov(e_{t-1}, e_{t-1})}{4}$$

$$+ \frac{Cov(e_{t-1}, e_{t-2})}{4}$$

$$= \frac{Cov(e_{t-1}, e_{t-1})}{4} \qquad \text{(as all the other covariances are zero)}$$

$$= 0.25\sigma_e^2$$

or

$$\gamma_{t,\,t-1} = 0.25\sigma_e^2 \qquad \text{for all } t \tag{2.2.15}$$

Furthermore,

$$Cov(Y_t, Y_{t-2}) = Cov\left\{\frac{e_t + e_{t-1}}{2}, \frac{e_{t-2} + e_{t-3}}{2}\right\}$$

$$= 0 \qquad \text{since the } e'\text{s are independent.}$$

Similarly, $Cov(Y_t, Y_{t-k}) = 0$ for $k > 1$, so we may write

$$\gamma_{t,\,s} = \begin{cases} 0.5\sigma_e^2 & \text{for } |t-s| = 0 \\ 0.25\sigma_e^2 & \text{for } |t-s| = 1 \\ 0 & \text{for } |t-s| > 1 \end{cases}$$

For the autocorrelation function, we have

$$\rho_{t,\,s} = \begin{cases} 1 & \text{for } |t-s| = 0 \\ 0.5 & \text{for } |t-s| = 1 \\ 0 & \text{for } |t-s| > 1 \end{cases} \tag{2.2.16}$$

since $0.25\sigma_e^2/0.5\sigma_e^2 = 0.5$.

Notice that $\rho_{2,1} = \rho_{3,2} = \rho_{4,3} = \rho_{9,8} = 0.5$. Values of Y precisely one time unit apart have exactly the same correlation no matter where they occur in time. Furthermore, $\rho_{3,1} = \rho_{4,2} = \rho_{t,\,t-2}$ and, more generally, $\rho_{t,\,t-k}$ is the same for all values of t. This leads us to the important concept of stationarity.

2.3 Stationarity

To make statistical inferences about the structure of a stochastic process on the basis of an observed record of that process, we must usually make some simplifying (and presumably reasonable) assumptions about that structure. The most important such assumption is that of **stationarity**. The basic idea of stationarity is that the probability laws that govern the behavior of the process do not change over time. In a sense, the process is in statistical equilibrium. Specifically, a process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2},\ldots, Y_{t_n}$ is the same as the joint distribution of $Y_{t_1-k}\ Y_{t_2-k},\ldots, Y_{t_n-k}$ for all choices of time points $t_1, t_2,\ldots, t_n$ and all choices of time lag k.

Thus, when $n = 1$ the (univariate) distribution of Y_t is the same as that of Y_{t-k} for all t and k; in other words, the Y's are (marginally) identically distributed. It then follows that $E(Y_t) = E(Y_{t-k})$ for all t and k so that the mean function is constant for all time. Additionally, $Var(Y_t) = Var(Y_{t-k})$ for all t and k so that the variance is also constant over time.

Setting $n = 2$ in the stationarity definition we see that the bivariate distribution of Y_t and Y_s must be the same as that of Y_{t-k} and Y_{s-k} from which it follows that $Cov(Y_t, Y_s) = Cov(Y_{t-k}, Y_{s-k})$ for all t, s, and k. Putting $k = s$ and then $k = t$, we obtain

$$\begin{aligned}\gamma_{t,s} &= Cov(Y_{t-s}, Y_0)\\ &= Cov(Y_0, Y_{s-t})\\ &= Cov(Y_0, Y_{|t-s|})\\ &= \gamma_{0,|t-s|}\end{aligned}$$

That is, the covariance between Y_t and Y_s depends on time only through the time difference $|t-s|$ and not otherwise on the actual times t and s. Thus, for a stationary process, we can simplify our notation and write

$$\gamma_k = Cov(Y_t, Y_{t-k}) \qquad \text{and} \qquad \rho_k = Corr(Y_t, Y_{t-k}) \tag{2.3.1}$$

Note also that

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

The general properties given in Equation (2.2.5) now become

$$\left.\begin{aligned}\gamma_0 &= Var(Y_t) & \rho_0 &= 1\\ \gamma_k &= \gamma_{-k} & \rho_k &= \rho_{-k}\\ |\gamma_k| &\le \gamma_0 & |\rho_k| &\le 1\end{aligned}\right\} \tag{2.3.2}$$

If a process is strictly stationary and has finite variance, then the covariance function must depend only on the time lag.

A definition that is similar to that of strict stationarity but is mathematically weaker is the following: A stochastic process $\{Y_t\}$ is said to be **weakly** (or **second-order**) **stationary** if

1. The mean function is constant over time, and
2. $\gamma_{t,\,t-k} = \gamma_{0,\,k}$ for all time t and lag k

In this book the term stationary when used alone will always refer to this weaker form of stationarity. However, if the joint distributions for the process are all multivariate normal distributions, it can be shown that the two definitions coincide. For stationary processes, we usually only consider $k \ge 0$.

White Noise

A very important example of a stationary process is the so-called **white noise** process, which is defined as a sequence of independent, identically distributed random variables $\{e_t\}$. Its importance stems not from the fact that it is an interesting model itself but from the fact that many useful processes can be constructed from white noise. The fact that $\{e_t\}$ is strictly stationary is easy to see since

$$
\begin{aligned}
& Pr(e_{t_1} \le x_1, e_{t_2} \le x_2, \ldots, e_{t_n} \le x_n) \\
&= Pr(e_{t_1} \le x_1) Pr(e_{t_2} \le x_2) \cdots Pr(e_{t_n} \le x_n) \qquad \text{(by independence)} \\
&= Pr(e_{t_1-k} \le x_1) Pr(e_{t_2-k} \le x_2) \cdots Pr(e_{t_n-k} \le x_n) \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(identical distributions)} \\
&= Pr(e_{t_1-k} \le x_1, e_{t_2-k} \le x_2, \ldots, e_{t_n-k} \le x_n) \qquad \text{(by independence)}
\end{aligned}
$$

as required. Also, $\mu_t = E(e_t)$ is constant and

$$
\gamma_k = \begin{cases} Var(e_t) & \text{for } k = 0 \\ 0 & \text{for } k \ne 0 \end{cases}
$$

Alternatively, we can write

$$
\rho_k = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \ne 0 \end{cases} \tag{2.3.3}
$$

The term white noise arises from the fact that a frequency analysis of the model shows that, in analogy with white light, all frequencies enter equally. We usually assume that the white noise process has mean zero and denote $Var(e_t)$ by σ_e^2.

The moving average example, on page 14, where $Y_t = (e_t + e_{t-1})/2$, is another example of a stationary process constructed from white noise. In our new notation, we have for the moving average process that

$$
\rho_k = \begin{cases} 1 & \text{for } k = 0 \\ 0.5 & \text{for } |k| = 1 \\ 0 & \text{for } |k| \ge 2 \end{cases}
$$

Random Cosine Wave

As a somewhat different example,[†] consider the process defined as follows:

$$Y_t = \cos\left[2\pi\left(\frac{t}{12} + \Phi\right)\right] \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots$$

where Φ is selected (once) from a uniform distribution on the interval from 0 to 1. A sample from such a process will appear highly deterministic since Y_t will repeat itself identically every 12 time units and look like a perfect (discrete time) cosine curve. However, its maximum will not occur at $t = 0$ but will be determined by the random phase Φ. The phase Φ can be interpreted as the fraction of a complete cycle completed by time $t = 0$. Still, the statistical properties of this process can be computed as follows:

$$\begin{aligned}
E(Y_t) &= E\left\{\cos\left[2\pi\left(\frac{t}{12} + \Phi\right)\right]\right\} \\
&= \int_0^1 \cos\left[2\pi\left(\frac{t}{12} + \phi\right)\right] d\phi \\
&= \frac{1}{2\pi}\sin\left[2\pi\left(\frac{t}{12} + \phi\right)\right]\Bigg|_{\phi = 0}^{1} \\
&= \frac{1}{2\pi}\left[\sin\left(2\pi\frac{t}{12} + 2\pi\right) - \sin\left(2\pi\frac{t}{12}\right)\right]
\end{aligned}$$

But this is zero since the sines must agree. So $\mu_t = 0$ for all t.

Also

$$\begin{aligned}
\gamma_{t,s} &= E\left\{\cos\left[2\pi\left(\frac{t}{12} + \Phi\right)\right]\cos\left[2\pi\left(\frac{s}{12} + \Phi\right)\right]\right\} \\
&= \int_0^1 \cos\left[2\pi\left(\frac{t}{12} + \phi\right)\right]\cos\left[2\pi\left(\frac{s}{12} + \phi\right)\right] d\phi \\
&= \frac{1}{2}\int_0^1 \left\{\cos\left[2\pi\left(\frac{t-s}{12}\right)\right] + \cos\left[2\pi\left(\frac{t+s}{12} + 2\phi\right)\right]\right\} d\phi \\
&= \frac{1}{2}\left\{\cos\left[2\pi\left(\frac{t-s}{12}\right)\right] + \frac{1}{4\pi}\sin\left[2\pi\left(\frac{t+s}{12} + 2\phi\right)\right]\Bigg|_{\phi = 0}^{1}\right\} \\
&= \frac{1}{2}\cos\left[2\pi\left(\frac{|t-s|}{12}\right)\right]
\end{aligned}$$

[†] This example contains optional material that is not needed in order to understand most of the remainder of this book. It will be used in Chapter 13, Introduction to Spectral Analysis.

So the process is stationary with autocorrelation function

$$\rho_k = \cos\left(2\pi\frac{k}{12}\right) \qquad \text{for } k = 0, \pm1, \pm2, \ldots \tag{2.3.4}$$

This example suggests that it will be difficult to assess whether or not stationarity is a reasonable assumption for a given time series on the basis of the time sequence plot of the observed data.

The random walk of page 12, where $Y_t = e_1 + e_2 + \cdots + e_t$, is also constructed from white noise but is *not* stationary. For example, the variance function, $Var(Y_t) = t\sigma_e^2$, is *not* constant; furthermore, the covariance function $\gamma_{t,s} = t\sigma_e^2$ for $0 \le t \le s$ does not depend only on time lag. However, suppose that instead of analyzing $\{Y_t\}$ directly, we consider the differences of successive Y-values, denoted ∇Y_t. Then $\nabla Y_t = Y_t - Y_{t-1} = e_t$, so the *differenced series*, $\{\nabla Y_t\}$, is stationary. This represents a simple example of a technique found to be extremely useful in many applications. Clearly, many real time series cannot be reasonably modeled by stationary processes since they are not in statistical equilibrium but are evolving over time. However, we can frequently transform nonstationary series into stationary series by simple techniques such as differencing. Such techniques will be vigorously pursued in the remaining chapters.

2.4 Summary

In this chapter we have introduced the basic concepts of stochastic processes that serve as models for time series. In particular, you should now be familiar with the important concepts of mean functions, autocovariance functions, and autocorrelation functions. We illustrated these concepts with the basic processes: the random walk, white noise, a simple moving average, and a random cosine wave. Finally, the fundamental concept of stationarity introduced here will be used throughout the book.

EXERCISES

2.1 Suppose $E(X) = 2$, $Var(X) = 9$, $E(Y) = 0$, $Var(Y) = 4$, and $Corr(X,Y) = 0.25$. Find:
(a) $Var(X + Y)$.
(b) $Cov(X, X + Y)$.
(c) $Corr(X + Y, X - Y)$.

2.2 If X and Y are dependent but $Var(X) = Var(Y)$, find $Cov(X + Y, X - Y)$.

2.3 Let X have a distribution with mean μ and variance σ^2, and let $Y_t = X$ for all t.
(a) Show that $\{Y_t\}$ is strictly and weakly stationary.
(b) Find the autocovariance function for $\{Y_t\}$.
(c) Sketch a "typical" time plot of Y_t.

2.4 Let $\{e_t\}$ be a zero mean white noise process. Suppose that the observed process is $Y_t = e_t + \theta e_{t-1}$, where θ is either 3 or 1/3.
(a) Find the autocorrelation function for $\{Y_t\}$ both when $\theta = 3$ and when $\theta = 1/3$.
(b) You should have discovered that the time series is stationary regardless of the value of θ and that the autocorrelation functions are the same for $\theta = 3$ and $\theta = 1/3$. For simplicity, suppose that the process mean is known to be zero and the variance of Y_t is known to be 1. You observe the series $\{Y_t\}$ for $t = 1, 2,\ldots, n$ and suppose that you can produce good estimates of the autocorrelations ρ_k. Do you think that you could determine which value of θ is correct (3 or 1/3) based on the estimate of ρ_k? Why or why not?

2.5 Suppose $Y_t = 5 + 2t + X_t$, where $\{X_t\}$ is a zero-mean stationary series with autocovariance function γ_k.
(a) Find the mean function for $\{Y_t\}$.
(b) Find the autocovariance function for $\{Y_t\}$.
(c) Is $\{Y_t\}$ stationary? Why or why not?

2.6 Let $\{X_t\}$ be a stationary time series, and define $Y_t = \begin{cases} X_t & \text{for } t \text{ odd} \\ X_t + 3 & \text{for } t \text{ even.} \end{cases}$
(a) Show that $Cov(Y_t, Y_{t-k})$ is free of t for all lags k.
(b) Is $\{Y_t\}$ stationary?

2.7 Suppose that $\{Y_t\}$ is stationary with autocovariance function γ_k.
(a) Show that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary by finding the mean and autocovariance function for $\{W_t\}$.
(b) Show that $U_t = \nabla^2 Y_t = \nabla[Y_t - Y_{t-1}] = Y_t - 2Y_{t-1} + Y_{t-2}$ is stationary. (You need not find the mean and autocovariance function for $\{U_t\}$.)

2.8 Suppose that $\{Y_t\}$ is stationary with autocovariance function γ_k. Show that for any fixed positive integer n and any constants $c_1, c_2,\ldots, c_n$, the process $\{W_t\}$ defined by $W_t = c_1 Y_t + c_2 Y_{t-1} + \cdots + c_n Y_{t-n+1}$ is stationary. (Note that Exercise 2.7 is a special case of this result.)

2.9 Suppose $Y_t = \beta_0 + \beta_1 t + X_t$, where $\{X_t\}$ is a zero-mean stationary series with autocovariance function γ_k and β_0 and β_1 are constants.
(a) Show that $\{Y_t\}$ is not stationary but that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary.
(b) In general, show that if $Y_t = \mu_t + X_t$, where $\{X_t\}$ is a zero-mean stationary series and μ_t is a polynomial in t of degree d, then $\nabla^m Y_t = \nabla(\nabla^{m-1} Y_t)$ is stationary for $m \geq d$ and nonstationary for $0 \leq m < d$.

2.10 Let $\{X_t\}$ be a zero-mean, unit-variance stationary process with autocorrelation function ρ_k. Suppose that μ_t is a nonconstant function and that σ_t is a positive-valued nonconstant function. The observed series is formed as $Y_t = \mu_t + \sigma_t X_t$.
(a) Find the mean and covariance function for the $\{Y_t\}$ process.
(b) Show that the autocorrelation function for the $\{Y_t\}$ process depends only on the time lag. Is the $\{Y_t\}$ process stationary?
(c) Is it possible to have a time series with a constant mean and with $Corr(Y_t, Y_{t-k})$ free of t but with $\{Y_t\}$ *not* stationary?

2.11 Suppose $Cov(X_t, X_{t-k}) = \gamma_k$ is free of t but that $E(X_t) = 3t$.
(a) Is $\{X_t\}$ stationary?
(b) Let $Y_t = 7 - 3t + X_t$. Is $\{Y_t\}$ stationary?

2.12 Suppose that $Y_t = e_t - e_{t-12}$. Show that $\{Y_t\}$ is stationary and that, for $k > 0$, its autocorrelation function is nonzero only for lag $k = 12$.

2.13 Let $Y_t = e_t - \theta(e_{t-1})^2$. For this exercise, assume that the white noise series is normally distributed.
(a) Find the autocorrelation function for $\{Y_t\}$.
(b) Is $\{Y_t\}$ stationary?

2.14 Evaluate the mean and covariance function for each of the following processes. In each case, determine whether or not the process is stationary.
(a) $Y_t = \theta_0 + t e_t$.
(b) $W_t = \nabla Y_t$, where Y_t is as given in part (a).
(c) $Y_t = e_t e_{t-1}$. (You may assume that $\{e_t\}$ is normal white noise.)

2.15 Suppose that X is a random variable with zero mean. Define a time series by $Y_t = (-1)^t X$.
(a) Find the mean function for $\{Y_t\}$.
(b) Find the covariance function for $\{Y_t\}$.
(c) Is $\{Y_t\}$ stationary?

2.16 Suppose $Y_t = A + X_t$, where $\{X_t\}$ is stationary and A is random but independent of $\{X_t\}$. Find the mean and covariance function for $\{Y_t\}$ in terms of the mean and autocovariance function for $\{X_t\}$ and the mean and variance of A.

2.17 Let $\{Y_t\}$ be stationary with autocovariance function γ_k. Let $\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$. Show that

$$Var(\bar{Y}) = \frac{\gamma_0}{n} + \frac{2}{n}\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\gamma_k$$

$$= \frac{1}{n}\sum_{k=-n+1}^{n-1}\left(1 - \frac{|k|}{n}\right)\gamma_k$$

2.18 Let $\{Y_t\}$ be stationary with autocovariance function γ_k. Define the sample variance as $S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \bar{Y})^2$.
(a) First show that $\sum_{t=1}^{n}(Y_t - \mu)^2 = \sum_{t=1}^{n}(Y_t - \bar{Y})^2 + n(\bar{Y} - \mu)^2$.
(b) Use part (a) to show that
(c) $E(S^2) = \frac{n}{n-1}\gamma_0 - \frac{n}{n-1}Var(\bar{Y}) = \gamma_0 - \frac{2}{n-1}\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\gamma_k$.
(Use the results of Exercise 2.17 for the last expression.)
(d) If $\{Y_t\}$ is a white noise process with variance γ_0, show that $E(S^2) = \gamma_0$.

2.19 Let $Y_1 = \theta_0 + e_1$, and then for $t > 1$ define Y_t recursively by $Y_t = \theta_0 + Y_{t-1} + e_t$. Here θ_0 is a constant. The process $\{Y_t\}$ is called a **random walk with drift**.
(a) Show that Y_t may be rewritten as $Y_t = t\theta_0 + e_t + e_{t-1} + \cdots + e_1$.
(b) Find the mean function for Y_t.
(c) Find the autocovariance function for Y_t.

2.20 Consider the standard random walk model where $Y_t = Y_{t-1} + e_t$ with $Y_1 = e_1$.
(a) Use the representation of Y_t above to show that $\mu_t = \mu_{t-1}$ for $t > 1$ with initial condition $\mu_1 = E(e_1) = 0$. Hence show that $\mu_t = 0$ for all t.
(b) Similarly, show that $Var(Y_t) = Var(Y_{t-1}) + \sigma_e^2$ for $t > 1$ with $Var(Y_1) = \sigma_e^2$ and hence $Var(Y_t) = t\sigma_e^2$.
(c) For $0 \le t \le s$, use $Y_s = Y_t + e_{t+1} + e_{t+2} + \cdots + e_s$ to show that $Cov(Y_t, Y_s) = Var(Y_t)$ and, hence, that $Cov(Y_t, Y_s) = \min(t, s)\sigma_e^2$.

2.21 For a random walk with random starting value, let $Y_t = Y_0 + e_t + e_{t-1} + \cdots + e_1$ for $t > 0$, where Y_0 has a distribution with mean μ_0 and variance σ_0^2. Suppose further that $Y_0, e_1, \ldots, e_t$ are independent.
(a) Show that $E(Y_t) = \mu_0$ for all t.
(b) Show that $Var(Y_t) = t\sigma_e^2 + \sigma_0^2$.
(c) Show that $Cov(Y_t, Y_s) = \min(t, s)\sigma_e^2 + \sigma_0^2$.
(d) Show that $Corr(Y_t, Y_s) = \sqrt{\dfrac{t\sigma_a^2 + \sigma_0^2}{s\sigma_a^2 + \sigma_0^2}}$ for $0 \le t \le s$.

2.22 Let $\{e_t\}$ be a zero-mean white noise process, and let c be a constant with $|c| < 1$. Define Y_t recursively by $Y_t = cY_{t-1} + e_t$ with $Y_1 = e_1$.
(a) Show that $E(Y_t) = 0$.
(b) Show that $Var(Y_t) = \sigma_e^2(1 + c^2 + c^4 + \cdots + c^{2t-2})$. Is $\{Y_t\}$ stationary?
(c) Show that

$$Corr(Y_t, Y_{t-1}) = c\sqrt{\frac{Var(Y_{t-1})}{Var(Y_t)}} \text{ and, in general,}$$

$$Corr(Y_t, Y_{t-k}) = c^k\sqrt{\frac{Var(Y_{t-k})}{Var(Y_t)}} \qquad \text{for } k > 0$$

Hint: Argue that Y_{t-1} is independent of e_t. Then use

$$Cov(Y_t, Y_{t-1}) = Cov(cY_{t-1} + e_t, Y_{t-1})$$

(d) For large t, argue that

$$Var(Y_t) \approx \frac{\sigma_e^2}{1 - c^2} \qquad \text{and} \qquad Corr(Y_t, Y_{t-k}) \approx c^k \qquad \text{for } k > 0$$

so that $\{Y_t\}$ could be called **asymptotically stationary**.
(e) Suppose now that we alter the initial condition and put $Y_1 = \dfrac{e_1}{\sqrt{1 - c^2}}$. Show that now $\{Y_t\}$ is stationary.

2.23 Two processes $\{Z_t\}$ and $\{Y_t\}$ are said to be **independent** if for any time points t_1, $t_2, \ldots, t_m$ and $s_1, s_2, \ldots, s_n$ the random variables $\{Z_{t_1}, Z_{t_2}, \ldots, Z_{t_m}\}$ are independent of the random variables $\{Y_{s_1}, Y_{s_2}, \ldots, Y_{s_n}\}$. Show that if $\{Z_t\}$ and $\{Y_t\}$ are independent stationary processes, then $W_t = Z_t + Y_t$ is stationary.

2.24 Let $\{X_t\}$ be a time series in which we are interested. However, because the measurement process itself is not perfect, we actually observe $Y_t = X_t + e_t$. We assume that $\{X_t\}$ and $\{e_t\}$ are independent processes. We call X_t the **signal** and e_t the **measurement noise** or **error process**.

If $\{X_t\}$ is stationary with autocorrelation function ρ_k, show that $\{Y_t\}$ is also stationary with

$$Corr(Y_t, Y_{t-k}) = \frac{\rho_k}{1 + \sigma_e^2/\sigma_X^2} \qquad \text{for } k \geq 1$$

We call σ_X^2/σ_e^2 the **signal-to-noise ratio**, or SNR. Note that the larger the SNR, the closer the autocorrelation function of the observed process $\{Y_t\}$ is to the autocorrelation function of the desired signal $\{X_t\}$.

2.25 Suppose $Y_t = \beta_0 + \sum_{i=1}^{k} [A_i \cos(2\pi f_i t) + B_i \sin(2\pi f_i t)]$, where $\beta_0, f_1, f_2, \ldots, f_k$ are constants and $A_1, A_2, \ldots, A_k, B_1, B_2, \ldots, B_k$ are independent random variables with zero means and variances $Var(A_i) = Var(B_i) = \sigma_i^2$. Show that $\{Y_t\}$ is stationary and find its covariance function.

2.26 Define the function $\Gamma_{t,s} = \frac{1}{2}E[(Y_t - Y_s)^2]$. In geostatistics, $\Gamma_{t,s}$ is called the *semivariogram.*

(a) Show that for a stationary process $\Gamma_{t,s} = \gamma_0 - \gamma_{|t-s|}$.

(b) A process is said to be *intrinsically stationary* if $\Gamma_{t,s}$ depends only on the time difference $|t - s|$. Show that the random walk process is intrinsically stationary.

2.27 For a fixed, positive integer r and constant ϕ, consider the time series defined by $Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^r e_{t-r}$.

(a) Show that this process is stationary for any value of ϕ.

(b) Find the autocorrelation function.

2.28 (Random cosine wave extended) Suppose that

$$Y_t = R\cos(2\pi(ft + \Phi)) \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots$$

where $0 < f < ½$ is a fixed frequency and R and Φ are uncorrelated random variables and with Φ uniformly distributed on the interval (0,1).

(a) Show that $E(Y_t) = 0$ for all t.

(b) Show that the process is stationary with $\gamma_k = \frac{1}{2}E(R^2)\cos(2\pi f k)$.

Hint: Use the calculations leading up to Equation (2.3.4), on page 19.

2.29 (Random cosine wave extended further) Suppose that

$$Y_t = \sum_{j=1}^{m} R_j \cos[2\pi(f_j t + \Phi_j)] \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots$$

where $0 < f_1 < f_2 < \cdots < f_m < ½$ are m fixed frequencies, and $R_1, \Phi_1, R_2, \Phi_2, \ldots, R_m, \Phi_m$ are uncorrelated random variables with each Φ_j uniformly distributed on the interval (0,1).

(a) Show that $E(Y_t) = 0$ for all t.

(b) Show that the process is stationary with $\gamma_k = \frac{1}{2} \sum_{j=1}^{m} E(R_j^2) \cos(2\pi f_j k)$. Hint: Do Exercise 2.28 first.

2.30 (Mathematical statistics required) Suppose that

$$Y_t = R \cos[2\pi(ft + \Phi)] \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots$$

where R and Φ are independent random variables and f is a fixed frequency. The phase Φ is assumed to be uniformly distributed on (0,1), and the amplitude R has a Rayleigh distribution with pdf $f(r) = re^{-r^2/2}$ for $r > 0$. Show that for each time point t, Y_t has a normal distribution. (Hint: Let $Y = R\cos[2\pi(ft + \Phi)]$ and $X = R\sin[2\pi(ft + \Phi)]$. Now find the joint distribution of X and Y. It can also be shown that all of the finite dimensional distributions are multivariate normal and hence the process is strictly stationary.)

Appendix A: Expectation, Variance, Covariance, and Correlation

In this appendix, we define expectation for continuous random variables. However, all of the properties described hold for all types of random variables, discrete, continuous, or otherwise. Let X have probability density function $f(x)$ and let the pair (X,Y) have joint probability density function $f(x,y)$.

The **expected value** of X is defined as $E(X) = \int_{-\infty}^{\infty} x f(x) dx$.

(If $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$; otherwise $E(X)$ is undefined.) $E(X)$ is also called the **expectation** of X or the **mean** of X and is often denoted μ or μ_X.

Properties of Expectation

If $h(x)$ is a function such that $\int_{-\infty}^{\infty} |h(x)| f(x) dx < \infty$, it may be shown that

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) dx$$

Similarly, if $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(x,y)| f(x, y) dx dy < \infty$, it may be shown that

$$E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy \tag{2.A.1}$$

As a corollary to Equation (2.A.1), we easily obtain the important result

$$E(aX + bY + c) = aE(X) + bE(Y) + c \tag{2.A.2}$$

We also have

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx dy \tag{2.A.3}$$

The **variance** of a random variable X is defined as

$$Var(X) = E\{[X - E(X)]^2\} \tag{2.A.4}$$

(provided $E(X^2)$ exists). The variance of X is often denoted by σ^2 or σ_X^2.

Properties of Variance

$$Var(X) \geq 0 \tag{2.A.5}$$

$$Var(a + bX) = b^2 Var(X) \tag{2.A.6}$$

If X and Y are independent, then

$$Var(X + Y) = Var(X) + Var(Y) \tag{2.A.7}$$

In general, it may be shown that

$$Var(X) = E(X^2) - [E(X)]^2 \tag{2.A.8}$$

The positive square root of the variance of X is called the **standard deviation** of X and is often denoted by σ or σ_X. The random variable $(X - \mu_X)/\sigma_X$ is called the **standardized version** of X. The mean and standard deviation of a standardized variable are always zero and one, respectively.

The **covariance** of X and Y is defined as $Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$.

Properties of Covariance

$$Cov(a + bX, c + dY) = bdCov(X, Y) \tag{2.A.9}$$

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y) \tag{2.A.10}$$

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z) \tag{2.A.11}$$

$$Cov(X, X) = Var(X) \tag{2.A.12}$$

$$Cov(X, Y) = Cov(Y, X) \tag{2.A.13}$$

If X and Y are independent,

$$Cov(X, Y) = 0 \tag{2.A.14}$$

The **correlation coefficient** of X and Y, denoted by $Corr(X, Y)$ or ρ, is defined as

$$\rho = Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

Alternatively, if X^* is a standardized X and Y^* is a standardized Y, then $\rho = E(X^*Y^*)$.

Properties of Correlation

$$-1 \le Corr(X, Y) \le 1 \tag{2.A.15}$$

$$Corr(a + bX, c + dY) = sign(bd)Corr(X, Y)$$

$$\text{where } sign(bd) = \begin{cases} 1 \text{ if } bd > 0 \\ 0 \text{ if } bd = 0 \\ -1 \text{ if } bd < 0 \end{cases} \tag{2.A.16}$$

$Corr(X, Y) = \pm 1$ if and only if there are constants a and b such that $Pr(Y = a + bX) = 1$.

http://www.springer.com/978-0-387-75958-6

Time Series Analysis
With Applications in R
Cryer, J.D.; Chan, K.-S.
2008, XIV, 491 p., Hardcover
ISBN: 978-0-387-75958-6