Chapter 5

Random Processes

5.1 INTRODUCTION

In this chapter, we introduce the concept of a random (or stochastic) process. The theory of random processes was first developed in connection with the study of fluctuations and noise in physical systems. A random process is the mathematical model of an empirical process whose development is governed by probability laws. Random processes provides useful models for the studies of such diverse fields as statistical physics, communication and control, time series analysis, population growth, and management sciences.

5.2 RANDOM PROCESSES

1. Defintion:

A *random process* is a family of r.v.'s $\{X(t), t \in T\}$ defined on a given probability space, indexed by the parameter t, where t varies over an index set T.

Recall that a random variable is a function defined on the sample space S (Sec. 2.2). Thus, a random process $\{X(t), t \in T\}$ is really a function of two arguments $\{X(t, \zeta), t \in T, \zeta \in S\}$. For a fixed $t(=t_k)$, $X(t_k, \zeta) = X_k(\zeta)$ is a r.v. denoted by $X(t_k)$, as ζ varies over the sample space S. On the other hand, for a fixed sample point $\zeta_i \in S$, $X(t, \zeta_i) = X_i(t)$ is a single function of time t, called a *sample function* or a *realization* of the process. The totality of all sample functions is called an *ensemble*.

Of course if both ζ and t are fixed, $X(t_k, \zeta_i)$ is simply a real number. In the following we use the notation $X(t)$ to represent $X(t, \zeta)$.

B. Description of a Random Process:

In a random process $\{X(t), t \in T\}$, the index set T is called the *parameter set* of the random process. The values assumed by $X(t)$ are called *states*, and the set of all possible values forms the *state space E* of the random process. If the index set T of a random process is discrete, then the process is called a *discrete-parameter* (or *discrete-time*) process. A discrete-parameter process is also called a *random sequence* and is denoted by $\{X_n, n = 1, 2, \ldots\}$. If T is continuous, then we have a *continuous-parameter* (or *continuous-time*) process. If the state space E of a random process is discrete, then the process is called a *discrete-state* process, often referred to as a *chain*. In this case, the state space E is often assumed to be $\{0, 1, 2, \ldots\}$. If the state space E is continuous, then we have a *continuous-state* process.

A complex random process $X(t)$ is defined by

$$X(t) = X_1(t) + jX_2(t)$$

where $X_1(t)$ and $X_2(t)$ are (real) random processes and $j = \sqrt{-1}$. Throughout this book, all random processes are real random processes unless specified otherwise.

5.3 CHARACTERIZATION OF RANDOM PROCESSES

A. Probabilistic Descriptions:

Consider a random process $X(t)$. For a fixed time t_1, $X(t_1) = X_1$ is a r.v., and its cdf $F_X(x_1; t_1)$ is defined as

$$F_X(x_1; t_1) = P\{X(t_1) \le x_1\} \tag{5.1}$$

$F_X(x_1; t_1)$ is known as the *first-order distribution* of $X(t)$. Similarly, given t_1 and t_2, $X(t_1) = X_1$ and $X(t_2) = X_2$ represent two r.v.'s. Their joint distribution is known as the *second-order distribution* of $X(t)$ and is given by

$$F_X(x_1, x_2; t_1, t_2) = P\{X(t_1) \le x_1, X(t_2) \le x_2\} \tag{5.2}$$

In general, we define the *nth-order distribution* of $X(t)$ by

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = P\{X(t_1) \le x_1, \ldots, X(t_n) \le x_n\} \tag{5.3}$$

If $X(t)$ is a discrete-time process, then $X(t)$ is specified by a collection of pmf's:

$$p_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = P\{X(t_1) = x_1, \ldots, X(t_n) = x_n\} \tag{5.4}$$

If $X(t)$ is a continuous-time process, then $X(t)$ is specified by a collection of pdf's:

$$f_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = \frac{\partial^n F_X(x_1, \ldots, x_n; t_1, \ldots, t_n)}{\partial x_1 \cdots \partial x_n} \tag{5.5}$$

The complete characterization of $X(t)$ requires knowledge of all the distributions as $n \to \infty$. Fortunately, often much less is sufficient.

B. Mean, Correlation, and Covariance Functions:

As in the case of r.v.'s, random processes are often described by using statistical averages.

The *mean* of $X(t)$ is defined by

$$\mu_X(t) = E[X(t)] \tag{5.6}$$

where $X(t)$ is treated as a random variable for a fixed value of t. In general, $\mu_X(t)$ is a function of time, and it is often called the *ensemble average* of $X(t)$. A measure of dependence among the r.v.'s of $X(t)$ is provided by its *autocorrelation function*, defined by

$$R_X(t, s) = E[X(t)X(s)] \tag{5.7}$$

Note that

$$R_X(t, s) = R_X(s, t) \tag{5.8}$$

and

$$R_X(t, t) = E[X^2(t)] \tag{5.9}$$

The *autocovariance function* of $X(t)$ is defined by

$$\begin{aligned} K_X(t, s) = \operatorname{Cov}[X(t), X(s)] &= E\{[X(t) - \mu_X(t)][X(s) - \mu_X(s)]\} \\ &= R_X(t, s) - \mu_X(t)\mu_X(s) \end{aligned} \tag{5.10}$$

It is clear that if the mean of $X(t)$ is zero, then $K_X(t, s) = R_X(t, s)$. Note that the *variance* of $X(t)$ is given by

$$\sigma_X{}^2(t) = \operatorname{Var}[X(t)] = E\{[X(t) - \mu_X(t)]^2\} = K_X(t, t) \tag{5.11}$$

If $X(t)$ is a complex random process, then its autocorrelation function $R_X(t, s)$ and autocovariance function $K_X(t, s)$ are defined, respectively, by

$$R_X(t, s) = E[X(t)X^*(s)] \tag{5.12}$$

and

$$K_X(t, s) = E\{[X(t) - \mu_X(t)][X(s) - \mu_X(s)]^*\} \tag{5.13}$$

where * denotes the complex conjugate.

5.4 CLASSIFICATION OF RANDOM PROCESSES

If a random process $X(t)$ possesses some special probabilistic structure, we can specify less to characterize $X(t)$ completely. Some simple random processes are characterized completely by only the first- and second-order distributions.

A. Stationary Processes:

A random process $\{X(t), t \in T\}$ is said to be *stationary* or *strict-sense stationary* if, for all n and for every set of time instants $(t_i \in T, i = 1, 2, \ldots, n\}$,

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = F_X(x_1, \ldots, x_n; t_1 + \tau, \ldots, t_n + \tau) \tag{5.14}$$

for any τ. Hence, the distribution of a stationary process will be unaffected by a shift in the time origin, and $X(t)$ and $X(t + \tau)$ will have the same distributions for any τ. Thus, for the first-order distribution,

$$F_X(x; t) = F_X(x; t + \tau) = F_X(x) \tag{5.15}$$

and

$$f_X(x; t) = f_X(x) \tag{5.16}$$

Then

$$\mu_X(t) = E[X(t)] = \mu \tag{5.17}$$

$$\text{Var}[X(t)] = \sigma^2 \tag{5.18}$$

where μ and σ^2 are contants. Similarly, for the second-order distribution,

$$F_X(x_1, x_2; t_1, t_2) = F_X(x_1, x_2; t_2 - t_1) \tag{5.19}$$

and

$$f_X(x_1, x_2; t_1, t_2) = f_X(x_1, x_2; t_2 - t_1) \tag{5.20}$$

Nonstationary processes are characterized by distributions depending on the points $t_1, t_2, \ldots, t_n$.

B. Wide-Sense Stationary Processes:

If stationary condition (*5.14*) of a random process $X(t)$ does not hold for all n but holds for $n \le k$, then we say that the process $X(t)$ is *stationary to order k*. If $X(t)$ is stationary to order 2, then $X(t)$ is said to be *wide-sense stationary* (WSS) or *weak stationary*. If $X(t)$ is a WSS random process, then we have

1. $E[X(t)] = \mu$ (constant) (*5.21*)
2. $R_X(t, s) = E[X(t)X(s)] = R_X(|s - t|)$ (*5.22*)

Note that a strict-sense stationary process is also a WSS process, but, in general, the converse is not true.

C. Independent Processes:

In a random process $X(t)$, if $X(t_i)$ for $i = 1, 2, \ldots, n$ are independent r.v.'s, so that for $n = 2, 3, \ldots$,

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = \prod_{i=1}^{n} F_X(x_i; t_i) \tag{5.23}$$

then we call $X(t)$ an *independent random process*. Thus, a first-order distribution is sufficient to characterize an independent random process $X(t)$.

D. Processes with Stationary Independent Increments:

A random process $\{X(t), t \ge 0)$ is said to have *independent increments* if whenever $0 < t_1 < t_2 < \cdots < t_n$,

$$X(0),\ X(t_1) - X(0),\ X(t_2) - X(t_1),\ \ldots,\ X(t_n) - X(t_{n-1})$$

are independent. If $\{X(t), t \geq 0)$ has independent increments and $X(t) - X(s)$ has the same distribution as $X(t + h) - X(s + h)$ for all $s, t, h \geq 0, s < t$, then the process $X(t)$ is said to have *stationary independent increments.*

Let $\{X(t), t \geq 0\}$ be a random process with stationary independent increments and assume that $X(0) = 0$. Then (Probs. 5.21 and 5.22)

$$E[X(t)] = \mu_1 t \tag{5.24}$$

where $\mu_1 = E[X(1)]$ and

$$\text{Var}[X(t)] = \sigma_1{}^2 t \tag{5.25}$$

where $\sigma_1{}^2 = \text{Var}[X(1)]$.

From Eq. (5.24), we see that processes with stationary independent increments are nonstationary. Examples of processes with stationary independent increments are Poisson processes and Wiener processes, which are discussed in later sections.

E. Markov Processes:

A random process $\{X(t), t \in T\}$ is said to be a *Markov process* if

$$P\{X(t_{n+1}) \leq x_{n+1} \mid X(t_1) = x_1, X(t_2) = x_2, \ldots, X(t_n) = x_n\} = P\{X(t_{n+1}) \leq x_{n+1} \mid X(t_n) = x_n\} \tag{5.26}$$

whenever $t_1 < t_2 < \cdots < t_n < t_{n+1}$.

A discrete-state Markov process is called a *Markov chain.* For a discrete-parameter Markov chain $\{X_n, n \geq 0\}$ (see Sec. 5.5), we have for every n

$$P(X_{n+1} = j \mid X_0 = i_0, X_i = i_1, \ldots, X_n = i) = P(X_{n+1} = j \mid X_n = i) \tag{5.27}$$

Equation (5.26) or Eq. (5.27) is referred to as the *Markov property* (which is also known as the *memoryless property*). This property of a Markov process states that the future state of the process depends only on the present state and not on the past history. Clearly, any process with independent increments is a Markov process.

Using the Markov property, the nth-order distribution of a Markov process $X(t)$ can be expressed as (Prob. 5.25)

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = F_X(x_1; t_1) \prod_{k=2}^{n} P\{X(t_k) \leq x_k\} \mid X(t_{k-1}) = x_{k-1}\} \tag{5.28}$$

Thus, all finite-order distributions of a Markov process can be expressed in terms of the second-order distributions.

F. Normal Processes:

A random process $\{X(t), t \in T\}$ is said to be a *normal* (or *gaussian*) process if for any integer n and any subset $\{t_1, \ldots, t_n\}$ of T, the n r.v.'s $X(t_1), \ldots, X(t_n)$ are jointly normally distributed in the sense that their joint characteristic function is given by

$$\begin{aligned}\Psi_{X(t_1)\cdots X(t_n)}(\omega_1, \ldots, \omega_n) &= E\{\exp j[\omega_1 X(t_1) + \cdots + \omega_n X(t_n)]\} \\ &= \exp\left\{j \sum_{i=1}^{n} \omega_i E[X(t_i)] - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \omega_i \omega_k \text{Cov}[X(t_i), X(t_k)]\right\}\end{aligned} \tag{5.29}$$

where $\omega_1, \ldots, \omega_n$ are any real numbers (see Probs. 5.59 and 5.60). Equation (5.29) shows that a normal process is completely characterized by the second-order distributions. Thus, if a normal process is wide-sense stationary, then it is also strictly stationary.

G. Ergodic Processes:

Consider a random process $\{X(t), -\infty < t < \infty\}$ with a typical sample function $x(t)$. The time average of $x(t)$ is defined as

$$\langle x(t)\rangle = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)\, dt \tag{5.30}$$

Similarly, the time autocorrelation function $\bar{R}_X(\tau)$ of $x(t)$ is defined as

$$\bar{R}_X(\tau) = \langle x(t)x(t+\tau)\rangle = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)x(t+\tau)\, dt \tag{5.31}$$

A random process is said to be *ergodic* if it has the property that the time averages of sample functions of the process are equal to the corresponding statistical or ensemble averages. The subject of *ergodicity* is extremely complicated. However, in most physical applications, it is assumed that stationary processes are ergodic.

5.5 DISCRETE-PARAMETER MARKOV CHAINS

In this section we treat a discrete-parameter Markov chain $\{X_n, n \ge 0\}$ with a discrete state space $E = \{0, 1, 2, \ldots\}$, where this set may be finite or infinite. If $X_n = i$, then the Markov chain is said to be in state i at time n (or the nth step). A discrete-parameter Markov chain $\{X_n, n \ge 0\}$ is characterized by [Eq. (*5.27*)]

$$P(X_{n+1} = j \mid X_0 = i_0, X_i = i_1, \ldots, X_n = i) = P(X_{n+1} = j \mid X_n = i) \tag{5.32}$$

where $P\{x_{n+1} = j \mid X_n = i\}$ are known as one-step transition probabilities. If $P\{x_{n+1} = j \mid X_n = i\}$ is independent of n, then the Markov chain is said to possess *stationary transition probabilities* and the process is referred to as a *homogeneous* Markov chain. Otherwise the process is known as a *nonhomogeneous* Markov chain. Note that the concepts of a Markov chain's having stationary transition probabilities and being a stationary random process should not be confused. The Markov process, in general, is not stationary. We shall consider only homogeneous Markov chains in this section.

A. Transition Probability Matrix:

Let $\{X_n, n \ge 0\}$ be a homogeneous Markov chain with a discrete infinite state space $E = \{0, 1, 2, \ldots\}$. Then

$$p_{ij} = P\{X_{n+1} = j \mid X_n = i\} \qquad i \ge 0, j \ge 0 \tag{5.33}$$

regardless of the value of n. A *transition probability matrix of* $\{X_n, n \ge 0\}$ is defined by

$$P = [p_{ij}] = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

where the elements satisfy

$$p_{ij} \ge 0 \qquad \sum_{j=0}^{\infty} p_{ij} = 1 \qquad i = 0, 1, 2, \ldots \tag{5.34}$$

In the case where the state space E is finite and equal to $\{1, 2, \ldots, m\}$, P is $m \times m$ dimensional; that is,

$$P = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

where $$p_{ij} \geq 0 \qquad \sum_{j=1}^{m} p_{ij} = 1 \qquad i = 1, 2, \ldots, m \tag{5.35}$$

A square matrix whose elements satisfy Eq. (5.34) or (5.35) is called a *Markov* matrix or *stochastic* matrix.

B. Higher-Order Transition Probabilities—Chapman-Kolmogorov Equation:

Tractability of Markov chain models is based on the fact that the probability distribution of $\{X_n, n \geq 0\}$ can be computed by matrix manipulations.

Let $P = [p_{ij}]$ be the transition probability matrix of a Markov chain $\{X_n, n \geq 0\}$. Matrix powers of P are defined by

$$P^2 = PP$$

with the (i, j)th element given by

$$p_{ij}^{(2)} = \sum_k p_{ik} p_{kj}$$

Note that when the state space E is infinite, the series above converges, since by Eq. (5.34),

$$\sum_k p_{ik} p_{kj} \leq \sum_k p_{ik} = 1$$

Similarly, $P^3 = PP^2$ has the (i, j)th element

$$p_{ij}^{(3)} = \sum_k p_{ik} p_{kj}^{(2)}$$

and in general, $P^{n+1} = PP^n$ has the (i, j)th element

$$p_{ij}^{(n+1)} = \sum_k p_{ik} p_{kj}^{(n)} \tag{5.36}$$

Finally, we define $P^0 = I$, where I is the identity matrix.

The n-step transition probabilities for the homogeneous Markov chain $\{X_n, n \geq 0\}$ are defined by

$$P(X_n = j \mid X_0 = i)$$

Then we can show that (Prob. 5.70)

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i) \tag{5.37}$$

We compute $p_{ij}^{(n)}$ by taking matrix powers.

The matrix identity

$$P^{n+m} = P^n P^m \qquad n, m \geq 0$$

when written in terms of elements

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{kj}^{(m)} \tag{5.38}$$

is known as the *Chapman-Kolmogorov equation.* It expresses the fact that a transition from i to j in $n+m$ steps can be achieved by moving from i to an intermediate k in n steps (with probability $p_{ik}^{(n)}$), and then proceeding to j from k in m steps (with probability $p_{kj}^{(m)}$). Furthermore, the events "go from i to k in n steps" and "go from k to j in m steps" are independent. Hence the probability of the transition from i to j in $n+m$ steps via i, k, j is $p_{ik}^{(n)}p_{kj}^{(m)}$. Finally, the probability of the transition from i to j is obtained by summing over the intermediate state k.

C. The Probability Distribution of $\{X_n, n \geq 0\}$:

Let $p_i(n) = P(X_n = i)$ and

$$\mathbf{p}(n) = [p_0(n) \quad p_1(n) \quad p_2(n) \quad \cdots]$$

where

$$\sum_k p_k(n) = 1$$

Then $p_i(0) = P(X_0 = i)$ are the *initial-state* probabilities,

$$\mathbf{p}(0) = [p_0(0) \quad p_1(0) \quad p_2(0) \quad \cdots]$$

is called the *initial-state probability vector*, and $\mathbf{p}(n)$ is called the *state probability vector after n transitions* or the *probability distribution* of X_n. Now it can be shown that (Prob. 5.29)

$$\mathbf{p}(n) = \mathbf{p}(0)P^n \tag{5.39}$$

which indicates that the probability distribution of a homogeneous Markov chain is completely determined by the one-step transition probability matrix P and the initial-state probability vector $\mathbf{p}(0)$.

D. Classification of States:

1. *Accessible States:*

State j is said to be *accessible* from state i if for some $n \geq 0$, $p_{ij}^{(n)} > 0$, and we write $i \to j$. Two states i and j accessible to each other are said to *communicate*, and we write $i \leftrightarrow j$. If all states communicate with each other, then we say that the Markov chain is *irreducible.*

2. *Recurrent States:*

Let T_j be the time (or the number of steps) of the first visit to state j after time zero, unless state j is never visited, in which case we set $T_j = \infty$. Then T_j is a discrete r.v. taking values in $\{1, 2, \ldots, \infty\}$. Let

$$f_{ij}^{(m)} = P(T_j = m \mid X_0 = i) = P(X_m = j, X_k \neq j, k = 1, 2, \ldots, m-1 \mid X_0 = i) \tag{5.40}$$

and $f_{ij}^{(0)} = 0$ since $T_j \geq 1$. Then

$$f_{ij}^{(1)} = P(T_j = 1 \mid X_0 = i) = P(X_1 = j \mid X_0 = i) = p_{ij} \tag{5.41}$$

and

$$f_{ij}^{(m)} = \sum_{k \neq j} p_{ik} f_{kj}^{(m-1)} \qquad m = 2, 3, \ldots \tag{5.42}$$

The probability of visiting j in finite time, starting from i, is given by

$$f_{ij} \equiv \sum_{n=0}^{\infty} f_{ij}^{(n)} = P(T_j < \infty \mid X_0 = i) \tag{5.43}$$

Now state j is said to be *recurrent* if

$$f_{jj} = P(T_j < \infty \mid X_0 = j) = 1 \tag{5.44}$$

That is, starting from j, the probability of eventual return to j is one. A recurrent state j is said to be *positive recurrent* if

$$E(T_j \mid X_0 = j) < \infty \tag{5.45}$$

and state j is said to be *null recurrent* if

$$E(T_j \mid X_0 = j) = \infty \tag{5.46}$$

Note that

$$E(T_j \mid X_0 = j) = \sum_{n=0}^{\infty} n f_{jj}^{(n)} \tag{5.47}$$

3. *Transient States:*

State j is said to be *transient* (or *nonrecurrent*) if

$$f_{jj} = P(T_j < \infty \mid X_0 = j) < 1 \tag{5.48}$$

In this case there is positive probability of never returning to state j.

4. *Periodic and Aperiodic States:*

We define the period of state j to be

$$d(j) = \gcd\{n \geq 1: p_{jj}^{(n)} > 0\}$$

where gcd stands for greatest common divisor.

If $d(j) > 1$, then state j is called *periodic* with period $d(j)$. If $d(j) = 1$, then state j is called *aperiodic*. Note that whenever $p_{jj} > 0$, j is aperiodic.

5. *Absorbing States:*

State j is said to be an *absorbing state* if $p_{jj} = 1$; that is, once state j is reached, it is never left.

E. Absorption Probabilities:

Consider a Markov chain $X(n) = \{X_n, n \geq 0\}$ with finite state space $E = \{1, 2, \ldots, N\}$ and transition probability matrix P. Let $A = \{1, \ldots, m\}$ be the set of absorbing states and $B = \{m+1, \ldots, N\}$ be a set of nonabsorbing states. Then the transition probability matrix P can be expressed as

$$P = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdots & 1 & 0 & \cdots & 0 \\ p_{m+1,1} & \cdot & \cdots & p_{m+1,m} & p_{m+1,m+1} & \cdots & p_{m+1,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{N,1} & \cdot & \cdots & p_{N,m} & p_{N,m+1} & \cdots & p_{N,N} \end{bmatrix} = \begin{bmatrix} I & O \\ R & Q \end{bmatrix} \tag{5.49a}$$

where I is an $m \times m$ identity matrix, O is an $m \times (N - m)$ zero matrix, and

$$R = \begin{bmatrix} p_{m+1,1} & \cdots & p_{m+1,m} \\ \vdots & \vdots & \vdots \\ p_{N,1} & \cdots & p_{N,m} \end{bmatrix} \qquad Q = \begin{bmatrix} p_{m+1,m+1} & \cdots & p_{m+1,N} \\ \vdots & \vdots & \vdots \\ p_{N,m+1} & \cdots & p_{N,N} \end{bmatrix} \tag{5.49b}$$

Note that the elements of R are the one-step transition probabilities from nonabsorbing to absorbing states, and the elements of Q are the one-step transition probabilities among the nonabsorbing states.

Let $U = [u_{kj}]$, where

$$u_{kj} = P\{X_n = j(\in A) \mid X_0 = k(\in B)\}$$

It is seen that U is an $(N - m) \times m$ matrix and its elements are the absorption probabilities for the various absorbing states. Then it can be shown that (Prob. 5.40)

$$U = (I - Q)^{-1}R = \Phi R \tag{5.50}$$

The matrix $\Phi = (I - Q)^{-1}$ is known as the *fundamental matrix* of the Markov chain $X(n)$. Let T_k denote the total time units (or steps) to absorption from state k. Let

$$\mathbf{T} = [T_{m+1} \quad T_{m+2} \quad \cdots \quad T_N]$$

Then it can be shown that (Prob. 5.74)

$$E(T_k) = \sum_{i=m+1}^{N} \phi_{ki} \qquad k = m + 1, \ldots, N \tag{5.51}$$

where ϕ_{ki} is the (k, i)th element of the fundamental matrix Φ.

F. Stationary Distributions:

Let P be the transition probability matrix of a homogeneous Markov chain $\{X_n, n \geq 0\}$. If there exists a probability vector $\hat{\mathbf{p}}$ such that

$$\hat{\mathbf{p}}P = \hat{\mathbf{p}} \tag{5.52}$$

then $\hat{\mathbf{p}}$ is called a *stationary distribution* for the Markov chain. Equation (*5.52*) indicates that a stationary distribution $\hat{\mathbf{p}}$ is a (left) *eigenvector* of P with *eigenvalue* 1. Note that any nonzero multiple of $\hat{\mathbf{p}}$ is also an eigenvector of P. But the stationary distribution $\hat{\mathbf{p}}$ is fixed by being a probability vector; that is, its components sum to unity.

G. Limiting Distributions:

A Markov chain is called *regular* if there is a finite positive integer m such that after m time-steps, every state has a nonzero chance of being occupied, no matter what the initial state. Let $A > O$ denote that every element a_{ij} of A satisfies the condition $a_{ij} > 0$. Then, for a regular Markov chain with transition probability matrix P, there exists an $m > 0$ such that $P^m > O$. For a regular homogeneous Markov chain we have the following theorem:

THEOREM 5.5.1

Let $\{X_n, n \geq 0\}$ be a regular homogeneous finite-state Markov chain with transition matrix P. Then

$$\lim_{n \to \infty} P^n = \hat{P} \tag{5.53}$$

where $\hat{P}$ is a matrix whose rows are identical and equal to the stationary distribution $\hat{\mathbf{p}}$ for the Markov chain defined by Eq. (*5.52*).

5.6 POISSON PROCESSES

A. Definitions:

Let t represent a time variable. Suppose an experiment begins at $t = 0$. Events of a particular kind occur randomly, the first at T_1, the second at T_2, and so on. The r.v. T_i denotes the time at which the ith event occurs, and the values t_i of T_i $(i = 1, 2, \ldots)$ are called *points of occurrence* (Fig. 5-1).

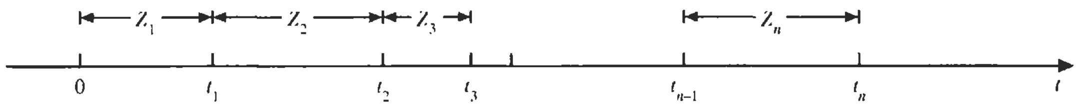

Fig. 5-1

Let
$$Z_n = T_n - T_{n-1} \tag{5.54}$$
and $T_0 = 0$. Then Z_n denotes the time between the $(n-1)$st and the nth events (Fig. 5-1). The sequence of ordered r.v.'s $\{Z_n, n \geq 1\}$ is sometimes called an *interarrival process.* If all r.v.'s Z_n are independent and identically distributed, then $\{Z_n, n \geq 1\}$ is called a *renewal process* or a *recurrent process.* From Eq. (*5.54*), we see that

$$T_n = Z_1 + Z_2 + \cdots + Z_n$$

where T_n denotes the time from the beginning until the occurrence of the nth event. Thus, $\{T_n, n \geq 0\}$ is sometimes called an *arrival process.*

B. Counting Processes:

A random process $\{X(t), t \geq 0\}$ is said to be a *counting process* if $X(t)$ represents the total number of "events" that have occurred in the interval $(0, t)$. From its definition, we see that for a counting process, $X(t)$ must satisfy the following conditions:

1. $X(t) \geq 0$ and $X(0) = 0$.
2. $X(t)$ is integer valued.
3. $X(s) \leq X(t)$ if $s < t$.
4. $X(t) - X(s)$ equals the number of events that have occurred on the interval (s, t).

A typical sample function (or realization) of $X(t)$ is shown in Fig. 5-2.

A counting process $X(t)$ is said to possess independent increments if the numbers of events which occur in disjoint time intervals are independent. A counting process $X(t)$ is said to possess stationary increments if the number of events in the interval $(s + h, t + h)$—that is, $X(t + h) - X(s + h)$—has the same distribution as the number of events in the interval (s, t)—that is, $X(t) - X(s)$—for all $s < t$ and $h > 0$.

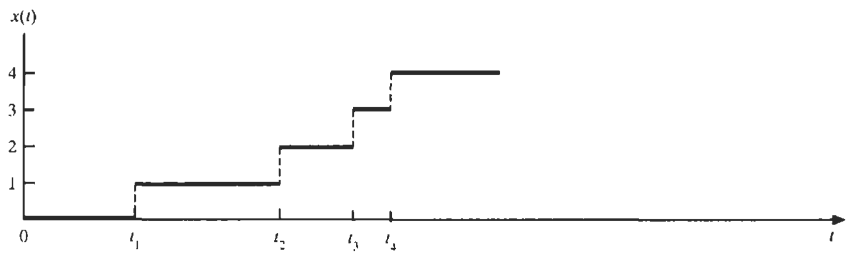

Fig. 5-2 A sample function of a counting process.

C. Poisson Processes:

One of the most important types of counting processes is the *Poisson process* (or *Poisson counting process*), which is defined as follows:

DEFINITION 5.6.1

A counting process $X(t)$ is said to be a Poisson process with *rate* (or *intensity*) $\lambda(>0)$ if

1. $X(0) = 0$.
2. $X(t)$ has independent increments.
3. The number of events in any interval of length t is Poisson distributed with mean λt; that is, for all $s, t > 0$,

$$P[X(t+s) - X(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \qquad n = 0, 1, 2, \ldots \tag{5.55}$$

It follows from condition 3 of Def. 5.6.1 that a Poisson process has stationary increments and that

$$E[X(t)] = \lambda t \tag{5.56}$$

Then by Eq. (*2.43*) (Sec. 2.7C), we have

$$\mathrm{Var}[X(t)] = \lambda t \tag{5.57}$$

Thus, the expected number of events in the unit interval (0, 1), or any other interval of unit length, is just λ (hence the name of the rate or intensity).

An alternative definition of a Poisson process is given as follows:

DEFINITION 5.6.2

A counting process $X(t)$ is said to be a Poisson process with rate (or intensity) $\lambda(>0)$ if

1. $X(0) = 0$.
2. $X(t)$ has independent and stationary increments.
3. $P[X(t+\Delta t) - X(t) = 1] = \lambda\,\Delta t + o(\Delta t)$
4. $P[X(t+\Delta t) - X(t) \geq 2] = o(\Delta t)$

where $o(\Delta t)$ is a function of Δt which goes to zero faster than does Δt; that is,

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0 \tag{5.58}$$

Note: Since addition or multiplication by a scalar does not change the property of approaching zero, even when divided by Δt, $o(\Delta t)$ satisfies useful identities such as $o(\Delta t) + o(\Delta t) = o(\Delta t)$ and $ao(\Delta t) = o(\Delta t)$ for all constant a.

It can be shown that Def. 5.6.1 and Def. 5.6.2 are equivalent (Prob. 5.49). Note that from conditions 3 and 4 of Def. 5.6.2, we have (Prob. 5.50)

$$P[X(t+\Delta t) - X(t) = 0] = 1 - \lambda\,\Delta t + o(\Delta t) \tag{5.59}$$

Equation (*5.59*) states that the probability that no event occurs in any short interval approaches unity as the duration of the interval approaches zero. It can be shown that in the Poisson process, the intervals between successive events are independent and identically distributed exponential r.v.'s (Prob. 5.53). Thus, we also identify the Poisson process as a renewal process with exponentially distributed intervals.

The autocorrelation function $R_X(t, s)$ and the autocovariance function $K_X(t, s)$ of a Poisson process $X(t)$ with rate λ are given by (Prob. 5.52)

$$R_X(t, s) = \lambda \min(t, s) + \lambda^2 ts \tag{5.60}$$

$$K_X(t, s) = \lambda \min(t, s) \tag{5.61}$$

5.7 WIENER PROCESSES

Another example of random processes with independent stationary increments is a *Wiener process.*

DEFINITION 5.7.1

A random process $\{X(t), t \geq 0\}$ is called a Wiener process if

1. $X(t)$ has stationary independent increments.
2. The increment $X(t) - X(s)$ $(t > s)$ is normally distributed.
3. $E[X(t)] = 0$.
4. $X(0) = 0$.

The Wiener process is also known as the *Brownian motion process,* since it originates as a model for Brownian motion, the motion of particles suspended in a fluid. From Def. 5.7.1, we can verify that a Wiener process is a normal process (Prob. 5.61) and

$$E[X(t)] = 0 \tag{5.62}$$

$$\mathrm{Var}[X(t)] = \sigma^2 t \tag{5.63}$$

where σ^2 is a parameter of the Wiener process which must be determined from observations. When $\sigma^2 = 1$, $X(t)$ is called a *standard* Wiener (or standard Brownian motion) process.

The autocorrelation function $R_X(t, s)$ and the autocovariance function $K_X(t, s)$ of a Wiener process $X(t)$ are given by (see Prob. 5.23)

$$R_X(t, s) = K_X(t, s) = \sigma^2 \min(t, s) \qquad s, t \geq 0 \tag{5.64}$$

DEFINITION 5.7.2

A random process $\{X(t), t \geq 0\}$ is called a *Wiener process with drift coefficient* μ if

1. $X(t)$ has stationary independent increments.
2. $X(t)$ is normally distributed with mean μt.
3. $X(0) = 0$.

From condition 2, the pdf of a standard Wiener process with drift coefficient μ is given by

$$f_{X(t)}(x) = \frac{1}{\sqrt{2\pi t}} e^{-(x-\mu t)^2/(2t)} \tag{5.65}$$

Solved Problems

RANDOM PROCESSES

5.1. Let $X_1, X_2, \ldots$ be independent Bernoulli r.v.'s (Sec. 2.7A) with $P(X_n = 1) = p$ and $P(X_n = 0) = q = 1 - p$ for all n. The collection of r.v.'s $\{X_n, n \geq 1\}$ is a random process, and it is called a *Bernoulli process.*

(*a*) Describe the Bernoulli process.

(*b*) Construct a typical sample sequence of the Bernoulli process.

(*a*) The Bernoulli process $\{X_n, n \geq 1\}$ is a discrete-parameter, discrete-state process. The state space is $E = \{0, 1\}$, and the index set is $T = \{1, 2, \ldots\}$.

(b) A sample sequence of the Bernoulli process can be obtained by tossing a coin consecutively. If a head appears, we assign 1, and if a tail appears, we assign 0. Thus, for instance,

n	1	2	3	4	5	6	7	8	9	10	...
Coin tossing	H	T	T	H	H	H	T	H	H	T	...
x_n	1	0	0	1	1	1	0	1	1	0	...

The sample sequence $\{x_n\}$ obtained above is plotted in Fig. 5-3.

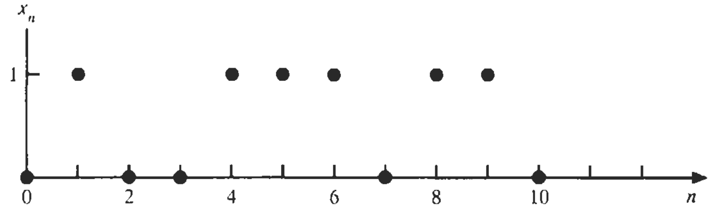

Fig. 5-3 A sample function of a Bernoulli process.

5.2. Let $Z_1, Z_2, \ldots$ be independent identically distributed r.v.'s with $P(Z_n = 1) = p$ and $P(Z_n = -1) = q = 1 - p$ for all n. Let

$$X_n = \sum_{i=1}^{n} Z_i \qquad n = 1, 2, \ldots \tag{5.66}$$

and $X_0 = 0$. The collection of r.v.'s $\{X_n, n \geq 0\}$ is a random process, and it is called the *simple random walk* $X(n)$ in one dimension.

(a) Describe the simple random walk $X(n)$.

(b) Construct a typical sample sequence (or realization) of $X(n)$.

(a) The simple random walk $X(n)$ is a discrete-parameter (or time), discrete-state random process. The state space is $E = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and the index parameter set is $T = \{0, 1, 2, \ldots\}$.

(b) A sample sequence $x(n)$ of a simple random walk $X(n)$ can be produced by tossing a coin every second and letting $x(n)$ increase by unity if a head appears and decrease by unity if a tail appears. Thus, for instance,

n	0	1	2	3	4	5	6	7	8	9	10	...
Coin tossing		H	T	T	H	H	H	T	H	H	T	...
$x(n)$	0	1	0	−1	0	1	2	1	2	3	2	...

The sample sequence $x(n)$ obtained above is plotted in Fig. 5-4. The simple random walk $X(n)$ specified in this problem is said to be *unrestricted* because there are no bounds on the possible values of X_n.

The simple random walk process is often used in the following primitive gambling model: Toss a coin. If a head appears, you win one dollar; if a tail appears, you lose one dollar (see Prob. 5.38).

5.3. Let $\{X_n, n \geq 0\}$ be a simple random walk of Prob. 5.2. Now let the random process $X(t)$ be defined by

$$X(t) = X_n \qquad n \leq t < n + 1$$

(a) Describe $X(t)$.

(b) Construct a typical sample function of $X(t)$.

(a) The random process $X(t)$ is a continuous-parameter (or time), discrete-state random process. The state space is $E = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and the index parameter set is $T = \{t, t \geq 0\}$.

(b) A sample function $x(t)$ of $X(t)$ corresponding to Fig. 5-4 is shown in Fig. 5-5.

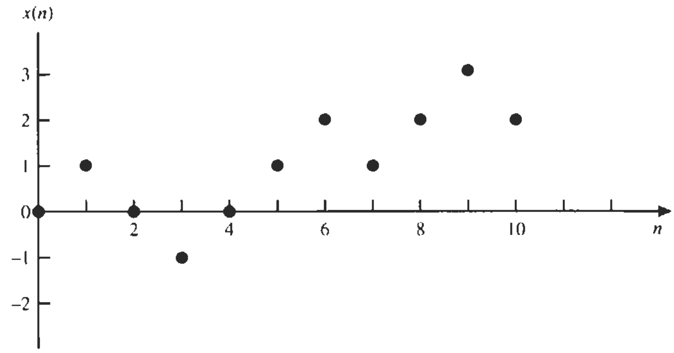

Fig. 5-4 A sample function of a random walk.

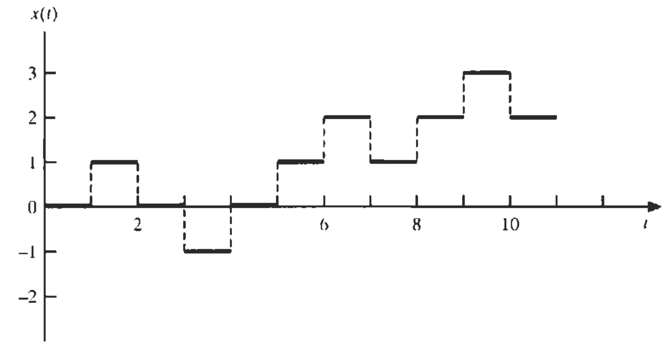

Fig. 5-5

5.4. Consider a random process $X(t)$ defined by

$$X(t) = Y \cos \omega t \qquad t \geq 0$$

where ω is a constant and Y is a uniform r.v. over (0, 1).

(*a*) Describe $X(t)$.

(*b*) Sketch a few typical sample functions of $X(t)$.

(*a*) The random process $X(t)$ is a continuous-parameter (or time), continuous-state random process. The state space is $E = \{x: -1 < x < 1\}$ and the index parameter set is $T = \{t: t \geq 0\}$.

(*b*) Three sample functions of $X(t)$ are sketched in Fig. 5-6.

5.5. Consider patients coming to a doctor's office at random points in time. Let X_n denote the time (in hours) that the nth patient has to wait in the office before being admitted to see the doctor.

(*a*) Describe the random process $X(n) = \{X_n, n \geq 1\}$.

(*b*) Construct a typical sample function of $X(n)$.

(*a*) The random process $X(n)$ is a discrete-parameter, continuous-state random process. The state space is $E = \{x: x \geq 0\}$, and the index parameter set is $T = \{1, 2, \ldots\}$.

(*b*) A sample function $x(n)$ of $X(n)$ is shown in Fig. 5-7.

CHARACTERIZATION OF RANDOM PROCESSES

5.6. Consider the Bernoulli process of Prob. 5.1. Determine the probability of occurrence of the sample sequence obtained in part (*b*) of Prob. 5.1.

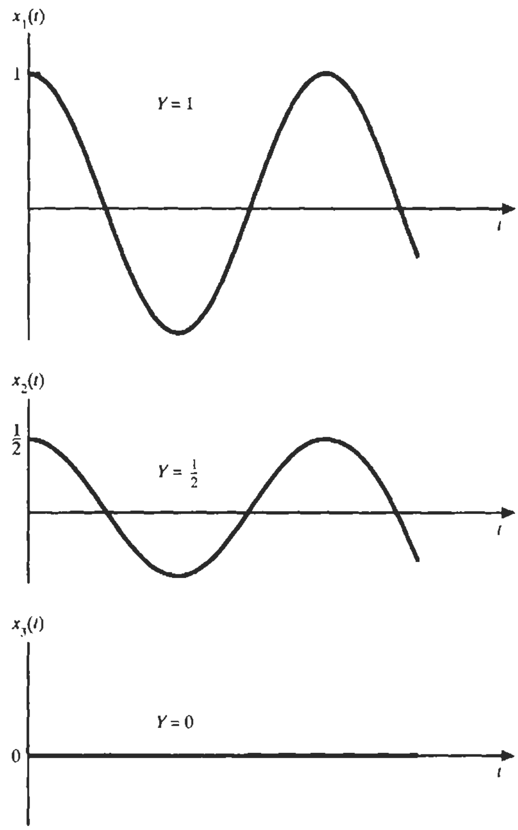

Fig. 5-6

Since X_n's are independent, we have

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2) \cdots P(X_n = x_n) \tag{5.67}$$

Thus, for the sample sequence of Fig. 5-3,

$$P(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 1, X_5 = 1, X_6 = 1, X_7 = 0, X_8 = 1, X_9 = 1, X_{10} = 0) = p^6q^4$$

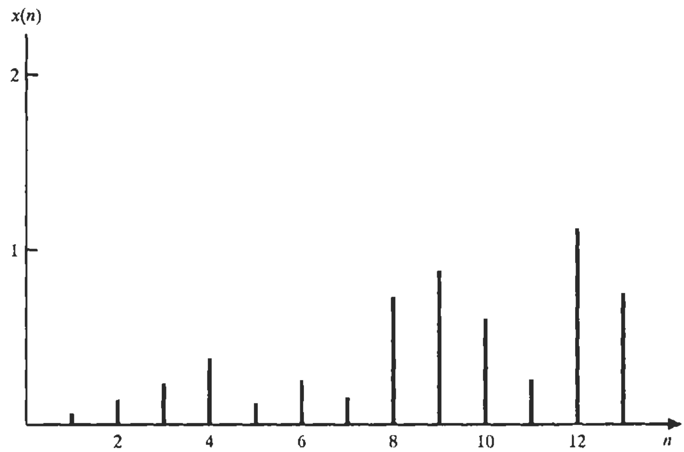

Fig. 5-7

5.7. Consider the random process $X(t)$ of Prob. 5.4. Determine the pdf's of $X(t)$ at $t = 0, \pi/4\omega, \pi/2\omega, \pi/\omega$.

For $t = 0$, $X(0) = Y \cos 0 = Y$. Thus,

$$f_{X(0)}(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

For $t = \pi/4\omega$, $X(\pi/4\omega) = Y \cos \pi/4 = 1/\sqrt{2}\, Y$. Thus,

$$f_{X(\pi/4\omega)}(x) = \begin{cases} \sqrt{2} & 0 < x < 1/\sqrt{2} \\ 0 & \text{otherwise} \end{cases}$$

For $t = \pi/2\omega$, $X(\pi/2\omega) = Y \cos \pi/2 = 0$; that is, $X(\pi/2\omega) = 0$ irrespective of the value of Y. Thus, the pmf of $X(\omega/2\omega)$ is

$$p_{X(\pi/2\omega)}(x) = P(X = 0) = 1$$

For $t = \pi/\omega$, $X(\pi/\omega) = Y \cos \pi = -Y$. Thus,

$$f_{X(\pi/\omega)}(x) = \begin{cases} 1 & -1 < x < 0 \\ 0 & \text{otherwise} \end{cases}$$

5.8. Derive the first-order probability distribution of the simple random walk $X(n)$ of Prob. 5.2.

The first-order probability distribution of the simple random walk $X(n)$ is given by

$$p_n(k) = P(X_n = k)$$

where k is an integer. Note that $P(X_0 = 0) = 1$. We note that $p_n(k) = 0$ if $n < |k|$ because the simple random walk cannot get to level k in less than $|k|$ steps. Thus, $n \geq |k|$.

Let N_n^+ and N_n^- be the r.v.'s denoting the numbers of $+1$s and -1s, respectively, in the first n steps. Then

$$n = N_n^+ + N_n^- \tag{5.68}$$

$$X_n = N_n^+ - N_n^- \tag{5.69}$$

Adding Eqs. (5.68) and (5.69), we get

$$N_n^+ = \tfrac{1}{2}(n + X_n) \tag{5.70}$$

Thus, $X_n = k$ if and only if $N_n^+ = \frac{1}{2}(n + k)$. From Eq. (5.70), we note that $2N_n^+ = n + X_n$ must be even. Thus, X_n must be even if n is even, and X_n must be odd if n is odd. We note that N_n^+ is a binomial r.v. with parameters (n, p). Thus, by Eq. (2.36), we obtain

$$p_n(k) = \binom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2} \qquad q = 1 - p \tag{5.71}$$

where $n \geq |k|$, and n and k are either both even or both odd.

5.9. Consider the simple random walk $X(n)$ of Prob. 5.2.

(*a*) Find the probability that $X(n) = -2$ after four steps.

(*b*) Verify the result of part (*a*) by enumerating all possible sample sequences that lead to the value $X(n) = -2$ after four steps.

(*a*) Setting $k = -2$ and $n = 4$ in Eq. (5.71), we obtain

$$P(X_4 = -2) = p_4(-2) = \binom{4}{1} pq^3 = 4pq^3 \qquad q = 1 - p$$

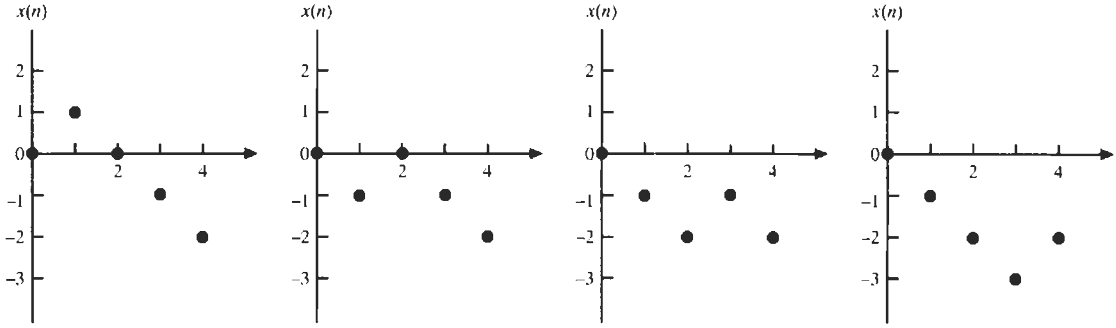

Fig. 5-8

(b) All possible sample functions that lead to the value $X_4 = -2$ after 4 steps are shown in Fig. 5-8. For each sample sequence, $P(X_4 = -2) = pq^3$. There are only four sample functions that lead to the value $X_4 = -2$ after four steps. Thus $P(X_4 = -2) = 4pq^3$.

5.10 Find the mean and variance of the simple random walk $X(n)$ of Prob. 5.2.

From Eq. (5.66), we have

$$X_n = X_{n-1} + Z_n \qquad n = 1, 2, \ldots \tag{5.72}$$

and $X_0 = 0$ and Z_n $(n = 1, 2, \ldots)$ are independent and identically distributed (iid) r.v.'s with

$$P(Z_n = +1) = p \qquad P(Z_n = -1) = q = 1 - p$$

From Eq. (5.72), we observe that

$$\begin{aligned} X_1 &= X_0 + Z_1 = Z_1 \\ X_2 &= X_1 + Z_2 = Z_1 + Z_2 \\ &\vdots \\ X_n &= Z_1 + Z_2 + \cdots + Z_n \end{aligned} \tag{5.73}$$

Then, because the Z_n are iid r.v.'s and $X_0 = 0$, by Eqs. (4.108) and (4.112), we have

$$E(X_n) = E\left(\sum_{k=1}^{n} Z_k\right) = nE(Z_k)$$

$$\text{Var}(X_n) = \text{Var}\left(\sum_{k=1}^{n} Z_k\right) = n\,\text{Var}(Z_k)$$

Now
$$E(Z_k) = (1)p + (-1)q = p - q \tag{5.74}$$

$$E(Z_k^2) = (1)^2 p + (-1)^2 q = p + q = 1 \tag{5.75}$$

Thus
$$\text{Var}(Z_k) = E(Z_k^2) - [E(Z_k)]^2 = 1 - (p - q)^2 = 4pq \tag{5.76}$$

Hence,
$$E(X_n) = n(p - q) \qquad q = 1 - p \tag{5.77}$$

$$\text{Var}(X_n) = 4npq \qquad q = 1 - p \tag{5.78}$$

Note that if $p = q = \frac{1}{2}$, then

$$E(X_n) = 0 \tag{5.79}$$

$$\text{Var}(X_n) = n \tag{5.80}$$

5.11. Find the autocorrelation function $R_X(n, m)$ of the simple random walk $X(n)$ of Prob. 5.2.

From Eq. (5.73), we can express X_n as

$$X_n = \sum_{i=0}^{n} Z_i \qquad n = 1, 2, \ldots \tag{5.81}$$

where $Z_0 = X_0 = 0$ and Z_i ($i \geq 1$) are iid r.v.'s with

$$P(Z_i = +1) = p \qquad P(Z_i = -1) = q = 1 - p$$

By Eq. (5.7),

$$R_X(n, m) = E[X(n)X(m)] = E(X_n X_m)$$

Then by Eq. (5.81),

$$R_X(n, m) = \sum_{i=0}^{n} \sum_{k=0}^{m} E(Z_i Z_k) = \sum_{i=0}^{\min(n, m)} E(Z_i^2) + \sum_{i=0}^{n} \sum_{\substack{k=0 \\ i \neq k}}^{m} E(Z_i)E(Z_k) \tag{5.82}$$

Using Eqs. (5.74) and (5.75), we obtain

$$R_X(n, m) = \min(n, m) + [nm - \min(n, m)](p - q)^2 \tag{5.83}$$

or

$$R_X(n, m) = \begin{cases} m + (nm - m)(p - q)^2 & m < n \\ n + (nm - n)(p - q)^2 & n < m \end{cases} \tag{5.84}$$

Note that if $p = q = \frac{1}{2}$, then

$$R_X(n, m) = \min(n, m) \qquad n, m > 0 \tag{5.85}$$

5.12. Consider the random process $X(t)$ of Prob. 5.4; that is,

$$X(t) = Y \cos \omega t \qquad t \geq 0$$

where ω is a constant and Y is a uniform r.v. over (0, 1).

(a) Find $E[X(t)]$.

(b) Find the autocorrelation function $R_X(t, s)$ of $X(t)$.

(c) Find the autocovariance function $K_X(t, s)$ of $X(t)$.

(a) From Eqs. (2.46) and (2.91), we have $E(Y) = \frac{1}{2}$ and $E(Y^2) = \frac{1}{3}$. Thus

$$E[X(t)] = E(Y \cos \omega t) = E(Y) \cos \omega t = \tfrac{1}{2} \cos \omega t \tag{5.86}$$

(b) By Eq. (5.7), we have

$$\begin{aligned} R_X(t, s) &= E[X(t)X(s)] = E(Y^2 \cos \omega t \cos \omega s) \\ &= E(Y^2) \cos \omega t \cos \omega s = \tfrac{1}{3} \cos \omega t \cos \omega s \end{aligned} \tag{5.87}$$

(c) By Eq. (5.10), we have

$$\begin{aligned} K_X(t, s) &= R_X(t, s) - E[X(t)]E[X(s)] \\ &= \tfrac{1}{3} \cos \omega t \cos \omega s - \tfrac{1}{4} \cos \omega t \cos \omega s \\ &= \tfrac{1}{12} \cos \omega t \cos \omega s \end{aligned} \tag{5.88}$$

5.13. Consider a discrete-parameter random process $X(n) = \{X_n, n \geq 1\}$ where the X_n's are iid r.v.'s with common cdf $F_X(x)$, mean μ, and variance σ^2.

(a) Find the joint cdf of $X(n)$.

(b) Find the mean of $X(n)$.

(c) Find the autocorrelation function $R_X(n, m)$ of $X(n)$.

(d) Find the autocovariance function $K_X(n, m)$ of $X(n)$.

(a) Since the X_n's are iid r.v.'s with common cdf $F_X(x)$, the joint cdf of $X(n)$ is given by

$$F_X(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_X(x_i) = [F_X(x)]^n \tag{5.89}$$

(b) The mean of $X(n)$ is

$$\mu_X(n) = E(X_n) = \mu \qquad \text{for all } n \tag{5.90}$$

(c) If $n \neq m$, by Eqs. (5.7) and (5.90),

$$R_X(n, m) = E(X_n X_m) = E(X_n)E(X_m) = \mu^2$$

If $n = m$, then by Eq. (2.31),

$$E(X_n^2) = \text{Var}(X_n) + [E(X_n)]^2 = \sigma^2 + \mu^2$$

Hence,
$$R_X(n, m) = \begin{cases} \mu^2 & n \neq m \\ \sigma^2 + \mu^2 & n = m \end{cases} \tag{5.91}$$

(d) By Eq. (5.10),

$$K_X(n, m) = R_X(n, m) - \mu_X(n)\mu_X(m) = \begin{cases} 0 & n \neq m \\ \sigma^2 & n = m \end{cases} \tag{5.92}$$

CLASSIFICATION OF RANDOM PROCESSES

5.14. Show that a random process which is stationary to order n is also stationary to all orders lower than n.

Assume that Eq. (5.14) holds for some particular n; that is,

$$P\{X(t_1) \le x_1, \ldots, X(t_n) \le x_n\} = P\{X(t_1 + \tau) \le x_1, \ldots, X(t_n + \tau) \le x_n\}$$

for any τ. Letting $x_n \to \infty$, we have [see Eq. (3.63)]

$$P\{X(t_1) \le x_1, \ldots, X(t_{n-1}) \le x_{n-1}\} = P\{X(t_1 + \tau) \le x_1, \ldots, X(t_{n-1} + \tau) \le x_{n-1}\}$$

and the process is stationary to order $n - 1$. Continuing the same procedure, we see that the process is stationary to all orders lower than n.

5.15. Show that if $\{X(t), t \in T\}$ is a strict-sense stationary random process, then it is also WSS.

Since $X(t)$ is strict-sense stationary, the first- and second-order distributions are invariant through time translation for all $\tau \in T$. Then we have

$$\mu_X(t) = E[X(t)] = E[X(t + \tau)] = \mu_X(t + \tau)$$

and hence the mean function $\mu_X(t)$ must be constant; that is,

$$E[X(t)] = \mu \text{ (constant)}$$

Similarly, we have

$$E[X(s)X(t)] = E[X(s + \tau)X(t + \tau)]$$

so that the autocorrelation function would depend on the time points s and t only through the difference $|t - s|$. Thus, $X(t)$ is WSS.

5.16. Let $\{X_n, n \ge 0\}$ be a sequence of iid r.v.'s with mean 0 and variance 1. Show that $\{X_n, n \ge 0\}$ is a WSS process.

By Eq. (5.90),

$$E(X_n) = 0 \text{ (constant)} \qquad \text{for all } n$$

and by Eq. (5.91),

$$R_X(n, n+k) = E(X_n X_{n+k}) = \begin{cases} E(X_n)E(X_{n+k}) = 0 & k \neq 0 \\ E(X_n^2) = \text{Var}(X_n) = 1 & k = 0 \end{cases}$$

which depends only on k. Thus, $\{X_n\}$ is a WSS process.

5.17. Show that if a random process $X(t)$ is WSS, then it must also be covariance stationary.

If $X(t)$ is WSS, then

$$E[X(t)] = \mu \text{ (constant)} \qquad \text{for all } t$$
$$R_X(t, t+\tau)] = R_X(\tau) \qquad \text{for all } t$$

Now

$$K_X(t, t+\tau) = \text{Cov}[X(t)X(t+\tau)] = R_X(t, t+\tau) - E[X(t)]E[X(t+\tau)] = R_X(\tau) - \mu^2$$

which indicates that $K_X(t, t+\tau)$ depends only on τ; thus, $X(t)$ is covariance stationary.

5.18. Consider a random process $X(t)$ defined by

$$X(t) = U \cos \omega t + V \sin \omega t \qquad -\infty < t < \infty \tag{5.93}$$

where ω is constant and U and V are r.v.'s.

(*a*) Show that the condition

$$E(U) = E(V) = 0 \tag{5.94}$$

is necessary for $X(t)$ to be stationary.

(*b*) Show that $X(t)$ is WSS if and only if U and V are uncorrelated with equal variance; that is,

$$E(UV) = 0 \qquad E(U^2) = E(V^2) = \sigma^2 \tag{5.95}$$

(*a*) Now

$$\mu_X(t) = E[X(t)] = E(U) \cos \omega t + E(V) \sin \omega t$$

must be independent of t for $X(t)$ to be stationary. This is possible only if $\mu_X(t) = 0$, that is, $E(U) = E(V) = 0$.

(*b*) If $X(t)$ is WSS, then

$$E[X^2(0)] = E\left[X^2\left(\frac{\pi}{2\omega}\right)\right] = R_{XX}(0) = \sigma_X^2$$

But $X(0) = U$ and $X(\pi/2\omega) = V$; thus

$$E(U^2) = E(V^2) = \sigma_X^2 = \sigma^2$$

Using the above result, we obtain

$$\begin{aligned} R_X(t, t+\tau) &= E[X(t)X(t+\tau)] \\ &= E\{(U \cos \omega t + V \sin \omega t)[U \cos \omega(t+\tau) + V \sin \omega(t+\tau)]\} \\ &= \sigma^2 \cos \omega\tau + E(UV) \sin(2\omega t + \omega\tau) \end{aligned} \tag{5.96}$$

which will be a function of τ only if $E(UV) = 0$. Conversely, if $E(UV) = 0$ and $E(U^2) = E(V^2) = \sigma^2$, then from the result of part (*a*) and Eq. (5.96), we have

$$\mu_X(t) = 0$$
$$R_X(t, t+\tau) = \sigma^2 \cos \omega\tau = R_X(\tau)$$

Hence, $X(t)$ is WSS.

5.19. Consider a random process $X(t)$ defined by

$$X(t) = U \cos t + V \sin t \qquad -\infty < t < \infty$$

where U and V are independent r.v.'s, each of which assumes the values -2 and 1 with the probabilities $\frac{1}{3}$ and $\frac{2}{3}$, respectively. Show that $X(t)$ is WSS but not strict-sense stationary.

We have

$$E(U) = E(V) = \tfrac{1}{3}(-2) + \tfrac{2}{3}(1) = 0$$

$$E(U^2) = E(V^2) = \tfrac{1}{3}(-2)^2 + \tfrac{2}{3}(1)^2 = 2$$

Since U and V are independent,

$$E(UV) = E(U)E(V) = 0$$

Thus, by the results of Prob. 5.18, $X(t)$ is WSS. To see if $X(t)$ is strict-sense stationary, we consider $E[X^3(t)]$.

$$\begin{aligned} E[X^3(t)] &= E[(U \cos t + V \sin t)^3] \\ &= E(U^3) \cos^3 t + 3E(U^2V) \cos^2 t \sin t + 3E(UV^2) \cos t \sin^2 t + E(V^3) \sin^3 t \end{aligned}$$

Now
$$E(U^3) = E(V^3) = \tfrac{1}{3}(-2)^3 + \tfrac{2}{3}(1)^3 = -2$$
$$E(U^2V) = E(U^2)E(V) = 0 \qquad E(UV^2) = E(U)E(V^2) = 0$$

Thus
$$E[X^3(t)] = -2(\cos^3 t + \sin^3 t)$$

which is a function of t. From Eq. (*5.16*), we see that all the moments of a strict-sense stationary process must be independent of time. Thus $X(t)$ is not strict-sense stationary.

5.20. Consider a random process $X(t)$ defined by

$$X(t) = A \cos(\omega t + \Theta) \qquad -\infty < t < \infty$$

where A and ω are constants and Θ is a uniform r.v. over $(-\pi, \pi)$. Show that $X(t)$ is WSS.

From Eq. (*2.44*), we have

$$f_\Theta(\theta) = \begin{cases} \dfrac{1}{2\pi} & -\pi < \theta < \pi \\ 0 & \text{otherwise} \end{cases}$$

Then
$$\mu_X(t) = \frac{A}{2\pi} \int_{-\pi}^{\pi} \cos(\omega t + \theta)\, d\theta = 0 \tag{5.97}$$

Setting $s = t + \tau$ in Eq. (5.7), we have

$$\begin{aligned} R_{XX}(t, t+\tau) &= \frac{A^2}{2\pi} \int_{-\pi}^{\pi} \cos(\omega t + \theta) \cos[\omega(t+\tau) + \theta]\, d\theta \\ &= \frac{A^2}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} [\cos \omega\tau + \cos(2\omega t + 2\theta + \omega\tau)]\, d\theta \\ &= \frac{A^2}{2} \cos \omega\tau \end{aligned} \tag{5.98}$$

Since the mean of $X(t)$ is a constant and the autocorrelation of $X(t)$ is a function of time difference only, we conclude that $X(t)$ is WSS.

5.21. Let $\{X(t), t \ge 0\}$ be a random process with stationary independent increments, and assume that $X(0) = 0$. Show that

$$E[X(t)] = \mu_1 t \tag{5.99}$$

where $\mu_1 = E[X(1)]$.

Let $$f(t) = E[X(t)] = E[X(t) - X(0)]$$

Then, for any t and s and using Eq. (*4.108*) and the property of the stationary independent increments, we have

$$\begin{aligned} f(t+s) &= E[X(t+s) - X(0)] \\ &= E[X(t+s) - X(s) + X(s) - X(0)] \\ &= E[X(t+s) - X(s)] + E[X(s) - X(0)] \\ &= E[X(t) - X(0)] + E[X(s) - X(0)] \\ &= f(t) + f(s) \end{aligned} \tag{5.100}$$

The only solution to the above functional equation is $f(t) = ct$, where c is a constant. Since $c = f(1) = E[X(1)]$, we obtain

$$E[X(t)] = \mu_1 t \qquad \mu_1 = E[X(1)]$$

5.22. Let $\{X(t), t \geq 0\}$ be a random process with stationary independent increments, and assume that $X(0) = 0$. Show that

(*a*) $$\text{Var}[X(t)] = \sigma_1{}^2 t \tag{5.101}$$

(*b*) $$\text{Var}[X(t) - X(s)] = \sigma_1{}^2(t - s) \qquad t > s \tag{5.102}$$

where $\sigma_1{}^2 = \text{Var}[X(1)]$.

(*a*) Let $$g(t) = \text{Var}[X(t)] = \text{Var}[X(t) - X(0)]$$

Then, for any t and s and using Eq. (*4.112*) and the property of the stationary independent increments, we get

$$\begin{aligned} g(t+s) &= \text{Var}[X(t+s) - X(0)] \\ &= \text{Var}[X(t+s) - X(s) + X(s) - X(0)] \\ &= \text{Var}[X(t+s) - X(s)] + \text{Var}[X(s) - X(0)] \\ &= \text{Var}[X(t) - X(0)] + \text{Var}[X(s) - X(0)] \\ &= g(t) + g(s) \end{aligned}$$

which is the same functional equation as Eq. (*5.100*). Thus, $g(t) = kt$, where k is a constant. Since $k = g(1) = \text{Var}[X(1)]$, we obtain

$$\text{Var}[X(t)] = \sigma_1{}^2 t \qquad \sigma_1{}^2 = \text{Var}[X(1)]$$

(*b*) Let $t > s$. Then

$$\begin{aligned} \text{Var}[X(t)] &= \text{Var}[X(t) - X(s) + X(s) - X(0)] \\ &= \text{Var}[X(t) - X(s)] + \text{Var}[X(s) - X(0)] \\ &= \text{Var}[X(t) - X(s)] + \text{Var}[X(s)] \end{aligned}$$

Thus, using Eq. (*5.101*), we obtain

$$\text{Var}[X(t) - X(s)] = \text{Var}[X(t)] - \text{Var}[X(s)] = \sigma_1{}^2(t - s)$$

5.23. Let $\{X(t), t \geq 0\}$ be a random process with stationary independent increments, and assume that $X(0) = 0$. Show that

$$\text{Cov}[X(t), X(s)] = K_X(t, s) = \sigma_1{}^2 \min(t, s) \tag{5.103}$$

where $\sigma_1{}^2 = \text{Var}[X(1)]$.

By definition (2.28),

$$\begin{aligned}\operatorname{Var}[X(t)-X(s)] &= E(\{X(t)-X(s)-E[X(t)-X(s)]\}^2)\\ &= E[(\{X(t)-E[X(t)]\}-\{X(s)-E[X(s)]\})^2]\\ &= E(\{X(t)-E[X(t)]\}^2-2\{X(t)-E[X(t)]\}\{X(s)-E[X(s)]\}+\{X(s)-E[X(s)]\}^2)\\ &= \operatorname{Var}[X(t)]-2\operatorname{Cov}[X(t),X(s)]+\operatorname{Var}[X(s)]\end{aligned}$$

Thus,

$$\operatorname{Cov}[X(t),X(s)] = \tfrac{1}{2}\{\operatorname{Var}[X(t)]+\operatorname{Var}[X(s)]-\operatorname{Var}[X(t)-X(s)]\}$$

Using Eqs. (5.101) and (5.102), we obtain

$$K_X(t,s) = \begin{cases}\tfrac{1}{2}\sigma_1{}^2[t+s-(t-s)] = \sigma_1{}^2 s & t > s\\ \tfrac{1}{2}\sigma_1{}^2[t+s-(s-t)] = \sigma_1{}^2 t & s > t\end{cases}$$

or

$$K_X(t,s) = \sigma_1{}^2 \min(t,s)$$

where $\sigma_1{}^2 = \operatorname{Var}[X(1)]$.

5.24. (a) Show that a simple random walk $X(n)$ of Prob. 5.2 is a Markov chain.
(b) Find its one-step transition probabilities.

(a) From Eq. (5.73) (Prob. 5.10), $X(n) = \{X_n, n \ge 0\}$ can be expressed as

$$X_0 = 0 \qquad X_n = \sum_{i=1}^{n} Z_i \qquad n \ge 1$$

where Z_n ($n = 1, 2, \ldots$) are iid r.v.'s with

$$P(Z_n = k) = a_k \qquad (k = 1, -1) \qquad \text{and} \qquad a_1 = p \qquad a_{-1} = q = 1 - p$$

Then $X(n) = \{X_n, n \ge 0\}$ is a Markov chain, since

$$\begin{aligned}P(X_{n+1} = i_{n+1} \mid X_0 = 0, X_1 = i_1, \ldots, X_n = i_n) &\\ = P(Z_{n+1} + i_n = i_{n+1} \mid X_0 = 0, X_1 = i_1, \ldots, X_n = i_n) &\\ = P(Z_{n+1} = i_{n+1} - i_n) = a_{i_{n+1}-i_n} = P(X_{n+1} = i_{n+1} \mid X_n = i_n)&\end{aligned}$$

since Z_{n+1} is independent of $X_0, X_1, \ldots, X_n$.

(b) The one-step transition probabilities are given by

$$p_{jk} = P(X_n = k \mid X_{n-1} = j) = \begin{cases} p & k = j+1\\ q = 1-p & k = j-1\\ 0 & \text{otherwise}\end{cases}$$

which do not depend on n. Thus, a simple random walk $X(n)$ is a homogeneous Markov chain.

5.25. Show that for a Markov process $X(t)$, the second-order distribution is sufficient to characterize $X(t)$.

Let $X(t)$ be a Markov process with the nth-order distribution

$$F_X(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = P\{X(t_1) \le x_1, X(t_2) \le x_2, \ldots, X(t_n) \le x_n\}$$

Then, using the Markov property (5.26), we have

$$\begin{aligned}F_X(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) &= P\{X(t_n) \le x_n \mid X(t_1) \le x_1, X(t_2) \le x_2, \ldots, X(t_{n-1}) \le x_{n-1}\}\\ &\quad \times P\{X(t_1) \le x_1, X(t_2) \le x_2, \ldots, X(t_{n-1}) \le x_{n-1}\}\\ &= P\{X(t_n) \le x_n \mid X(t_{n-1}) \le x_{n-1}\} F_X(x_1, \ldots, x_{n-1}; t_1, \ldots, t_{n-1})\end{aligned}$$

Applying the above relation repeatedly for lower-order distribution, we can write

$$F_X(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = F_X(x_1, t_1) \prod_{k=2}^{n} P\{X(t_k) \le x_k \mid X(t_{k-1}) \le x_{k-1}\} \tag{5.104}$$

Hence, all finite-order distributions of a Markov process can be completely determined by the second-order distribution.

5.26. Show that if a normal process is WSS, then it is also strict-sense stationary.

By Eq. (5.29), a normal random process $X(t)$ is completely characterized by the specification of the mean $E[X(t)]$ and the covariance function $K_X(t, s)$ of the process. Suppose that $X(t)$ is WSS. Then, by Eqs. (5.21) and (5.22), Eq. (5.29) becomes

$$\Psi_{X(t_1)\cdots X(t_n)}(\omega_1, \ldots, \omega_n) = \exp\left\{j \sum_{i=1}^{n} \mu\omega_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} K_X(t_i - t_k)\omega_i \omega_k\right\} \tag{5.105}$$

Now we translate all of the time instants $t_1, t_2, \ldots, t_n$ by the same amount τ. The joint characteristic function of the new r.v.'s $X(t_i + \tau)$, $i = 1, 2, \ldots, n$, is then

$$\begin{aligned}\Psi_{X(t_1+\tau)\cdots X(t_n+\tau)}(\omega_1, \ldots, \omega_n) &= \exp\left\{j \sum_{i=1}^{n} \mu\omega_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} K_X[t_i + \tau - (t_k + \tau)]\omega_i \omega_k\right\} \\ &= \exp\left\{j \sum_{i=1}^{n} \mu\omega_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} K_X(t_i - t_k)\omega_i \omega_k\right\} \\ &= \Psi_{X(t_1)\cdots X(t_n)}(\omega_1, \ldots, \omega_n)\end{aligned} \tag{5.106}$$

which indicates that the joint characteristic function (and hence the corresponding joint pdf) is unaffected by a shift in the time origin. Since this result holds for any n and any set of time instants ($t_i \in T$, $i = 1, 2, \ldots, n$), it follows that if a normal process is WSS, then it is also strict-sense stationary.

5.27. Let $\{X(t), -\infty < t < \infty\}$ be a zero-mean, stationary, normal process with the autocorrelation function

$$R_X(\tau) = \begin{cases} 1 - \dfrac{|\tau|}{T} & -T \le \tau \le T \\ 0 & \text{otherwise} \end{cases} \tag{5.107}$$

Let $\{X(t_i), i = 1, 2, \ldots, n\}$ be a sequence of n samples of the process taken at the time instants

$$t_i = i\frac{T}{2} \qquad i = 1, 2, \ldots, n$$

Find the mean and the variance of the sample mean

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} X(t_i) \tag{5.108}$$

Since $X(t)$ is zero-mean and stationary, we have

$$E[X(t_i)] = 0$$

and
$$R_X(t_i, t_k) = E[X(t_i)X(t_k)] = R_X(t_k - t_i) = R_X\left[(k - i)\frac{T}{2}\right]$$

Thus
$$E(\hat{\mu}_n) = E\left[\frac{1}{n} \sum_{i=1}^{n} X(t_i)\right] = \frac{1}{n} \sum_{i=1}^{n} E[X(t_i)] = 0 \tag{5.109}$$

and
$$\begin{aligned}\operatorname{Var}(\hat{\mu}_n) &= E\{[\hat{\mu}_n - E(\hat{\mu}_n)]^2\} = E(\hat{\mu}_n{}^2) \\ &= E\left\{\left[\frac{1}{n} \sum_{i=1}^{n} X(t_i)\right]\left[\frac{1}{n} \sum_{k=1}^{n} X(t_k)\right]\right\} \\ &= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} E[X(t_i)X(t_k)] = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} R_X\left[(k - i)\frac{T}{2}\right]\end{aligned}$$

By Eq. (5.107),

$$R_X[(k-i)T/2] = \begin{cases} 1 & k = i \\ \frac{1}{2} & |k-i| = 1 \\ 0 & |k-i| > 2 \end{cases}$$

Thus
$$\operatorname{Var}(\hat{\mu}_n) = \frac{1}{n^2}[n(1) + 2(n-1)(\tfrac{1}{2}) + 0] = \frac{1}{n^2}(2n-1) \tag{5.110}$$

DISCRETE-PARAMETER MARKOV CHAINS

5.28. Show that if P is a Markov matrix, then P^n is also a Markov matrix for any positive integer n.

Let
$$P = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{12} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

Then by the property of a Markov matrix [Eq. (5.35)], we can write

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{12} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

or
$$P\mathbf{a} = \mathbf{a} \tag{5.111}$$

where
$$\mathbf{a}^T = [1 \quad 1 \quad \cdots \quad 1]$$

Premultiplying both sides of Eq. (5.111) by P, we obtain

$$P^2\mathbf{a} = P\mathbf{a} = \mathbf{a}$$

which indicates that P^2 is also a Markov matrix. Repeated premultiplication by P yields

$$P^n\mathbf{a} = \mathbf{a}$$

which shows that P^n is also a Markov matrix.

5.29. Verify Eq. (5.39); that is,

$$\mathbf{p}(n) = \mathbf{p}(0)P^n$$

We verify Eq. (5.39) by induction. If the state of X_0 is i, state X_1 will be j only if a transition is made from i to j. The events $\{X_0 = i, i = 1, 2, \ldots\}$ are mutually exclusive, and one of them must occur. Hence, by the law of total probability [Eq. (1.44)],

$$P(X_1 = j) = \sum_i P(X_0 = i)P(X_1 = j \mid X_0 = i)$$

or
$$p_j(1) = \sum_i p_i(0)p_{ij} \qquad j = 1, 2, \ldots \tag{5.112}$$

In terms of vectors and matrices, Eq. (5.112) can be expressed as

$$\mathbf{p}(1) = \mathbf{p}(0)P \tag{5.113}$$

Thus, Eq. (5.39) is true for $n = 1$. Assume now that Eq. (5.39) is true for $n = k$; that is,

$$\mathbf{p}(k) = \mathbf{p}(0)P^k$$

Again, by the law of total probability,

$$P(X_{k+1} = j) = \sum_i P(X_k = i)P(X_{k+1} = j \mid X_k = i)$$

or

$$p_j(k+1) = \sum_i p_i(k)p_{ij} \qquad j = 1, 2, \ldots \tag{5.114}$$

In terms of vectors and matrices, Eq. (5.114) can be expressed as

$$\mathbf{p}(k+1) = \mathbf{p}(k)P = \mathbf{p}(0)P^kP = \mathbf{p}(0)P^{k+1} \tag{5.115}$$

which indicates that Eq. (5.39) is true for $k + 1$. Hence, we conclude that Eq. (5.39) is true for all $n \geq 1$.

5.30. Consider a two-state Markov chain with the transition probability matrix

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \qquad 0 < a < 1, 0 < b < 1 \tag{5.116}$$

(*a*) Show that the *n*-step transition probability matrix P^n is given by

$$P^n = \frac{1}{a+b}\left\{\begin{bmatrix} b & a \\ b & a \end{bmatrix} + (1-a-b)^n\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}\right\} \tag{5.117}$$

(*b*) Find P^n when $n \to \infty$.

(*a*) From matrix analysis, the characteristic equation of P is

$$\begin{aligned} c(\lambda) = |\lambda I - P| &= \begin{vmatrix} \lambda - (1-a) & -a \\ -b & \lambda - (1-b) \end{vmatrix} \\ &= (\lambda - 1)(\lambda - 1 + a + b) = 0 \end{aligned}$$

Thus, the eigenvalues of P are $\lambda_1 = 1$ and $\lambda_2 = 1 - a - b$. Then, using the spectral decomposition method, P^n can be expressed as

$$P^n = \lambda_1{}^n E_1 + \lambda_2{}^n E_2 \tag{5.118}$$

where E_1 and E_2 are constituent matrices of P, given by

$$E_1 = \frac{1}{\lambda_1 - \lambda_2}[P - \lambda_2 I] \qquad E_2 = \frac{1}{\lambda_2 - \lambda_1}[P - \lambda_1 I] \tag{5.119}$$

Substituting $\lambda_1 = 1$ and $\lambda_2 = 1 - a - b$ in the above expressions, we obtain

$$E_1 = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} \qquad E_2 = \frac{1}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}$$

Thus, by Eq. (5.118), we obtain

$$\begin{aligned} P^n &= E_1 + (1-a-b)^n E_2 \\ &= \frac{1}{a+b}\left\{\begin{bmatrix} b & a \\ b & a \end{bmatrix} + (1-a-b)^n\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}\right\} \end{aligned} \tag{5.120}$$

(*b*) If $0 < a < 1, 0 < b < 1$, then $0 < 1 - a < 1$ and $|1 - a - b| < 1$. So $\lim_{n \to \infty}(1 - a - b)^n = 0$ and

$$\lim_{n \to \infty} P^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} \tag{5.121}$$

Note that a limiting matrix exists and has the same rows (see Prob. 5.47).

5.31. An example of a two-state Markov chain is provided by a communication network consisting of the sequence (or cascade) of stages of binary communication channels shown in Fig. 5-9. Here X_n

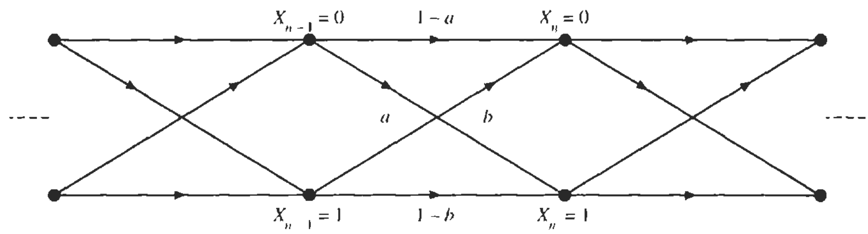

Fig. 5-9 Binary communication network.

denotes the digit leaving the nth stage of the channel and X_0 denotes the digit entering the first stage. The transition probability matrix of this communication network is often called the *channel matrix* and is given by Eq. (*5.116*); that is,

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \qquad 0 < a < 1, 0 < b < 1$$

Assume that $a = 0.1$ and $b = 0.2$, and the initial distribution is $P(X_0 = 0) = P(X_0 = 1) = 0.5$.

(*a*) Find the distribution of X_n.

(*b*) Find the distribution of X_n when $n \to \infty$.

(*a*) The channel matrix of the communication network is

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

and the initial distribution is

$$\mathbf{p}(0) = [0.5 \quad 0.5]$$

By Eq. (*5.39*), the distribution of X_n is given by

$$\mathbf{p}(n) = \mathbf{p}(0)P^n = [0.5 \quad 0.5]\begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^n$$

Letting $a = 0.1$ and $b = 0.2$ in Eq. (*5.117*), we get

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^n = \frac{1}{0.3}\begin{bmatrix} 0.2 & 0.1 \\ 0.2 & 0.1 \end{bmatrix} + \frac{(0.7)^n}{0.3}\begin{bmatrix} 0.1 & -0.1 \\ -0.2 & 0.2 \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{2+(0.7)^n}{3} & \dfrac{1-(0.7)^n}{3} \\ \dfrac{2-2(0.7)^n}{3} & \dfrac{1+2(0.7)^n}{3} \end{bmatrix}$$

Thus, the distribution of X_n is

$$\mathbf{p}(n) = [0.5 \quad 0.5]\begin{bmatrix} \dfrac{2+(0.7)^n}{3} & \dfrac{1-(0.7)^n}{3} \\ \dfrac{2-2(0.7)^n}{3} & \dfrac{1+2(0.7)^n}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{2}{3} - \dfrac{(0.7)^n}{6} & \dfrac{1}{3} + \dfrac{(0.7)^n}{6} \end{bmatrix}$$

that is,

$$P(X_n = 0) = \frac{2}{3} - \frac{(0.7)^n}{6} \qquad \text{and} \qquad P(X_n = 1) = \frac{1}{3} + \frac{(0.7)^n}{6}$$

(*b*) Since $\lim_{n \to \infty}(0.7)^n = 0$, the distribution of X_n when $n \to \infty$ is

$$P(X_\infty = 0) = \tfrac{2}{3} \qquad \text{and} \qquad P(X_\infty = 1) = \tfrac{1}{3}$$

5.32. Verify the transitivity property of the Markov chain; that is, if $i \to j$ and $j \to k$, then $i \to k$.

By definition, the relations $i \to j$ and $j \to k$ imply that there exist integers n and m such that $p_{ij}^{(n)} > 0$ and $p_{jk}^{(m)} > 0$. Then, by the Chapman-Kolmogorov equation (*5.38*), we have

$$p_{ik}^{(n+m)} = \sum_r p_{ir}^{(n)} p_{rk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0 \tag{5.122}$$

Therefore $i \to k$.

5.33. Verify Eq. (*5.42*).

If the Markov chain $\{X_n\}$ goes from state i to state j in m steps, the first step must take the chain from i to some state k, where $k \neq j$. Now after that first step to k, we have $m - 1$ steps left, and the chain must get to state j, from state k, on the last of those steps. That is, the first visit to state j must occur on the $(m - 1)$st step, starting now in state k. Thus we must have

$$f_{ij}^{(m)} = \sum_{k \neq j} p_{ik} f_{kj}^{(m-1)} \qquad m = 2, 3, \ldots$$

5.34. Show that in a finite-state Markov chain, not all states can be transient.

Suppose that the states are $0, 1, \ldots, m$, and suppose that they are all transient. Then by definition, after a finite amount of time (say T_0), state 0 will never be visited; after a finite amount of time (say T_1), state 1 will never be visited; and so on. Thus, after a finite time $T = \max\{T_0, T_1, \ldots, T_m\}$, no state will be visited. But as the process must be in some state after time T, we have a contradiction. Thus, we conclude that not all states can be transient and at least one of the states must be recurrent.

5.35. A state transition diagram of a finite-state Markov chain is a line diagram with a vertex corresponding to each state and a directed line between two vertices i and j if $p_{ij} > 0$. In such a diagram, if one can move from i and j by a path following the arrows, then $i \to j$. The diagram is useful to determine whether a finite-state Markov chain is irreducible or not, or to check for periodicities. Draw the state transition diagrams and classify the states of the Markov chains with the following transition probability matrices:

(*a*) $P = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$ (*b*) $P = \begin{bmatrix} 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

(*c*) $P = \begin{bmatrix} 0.3 & 0.4 & 0 & 0 & 0.3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

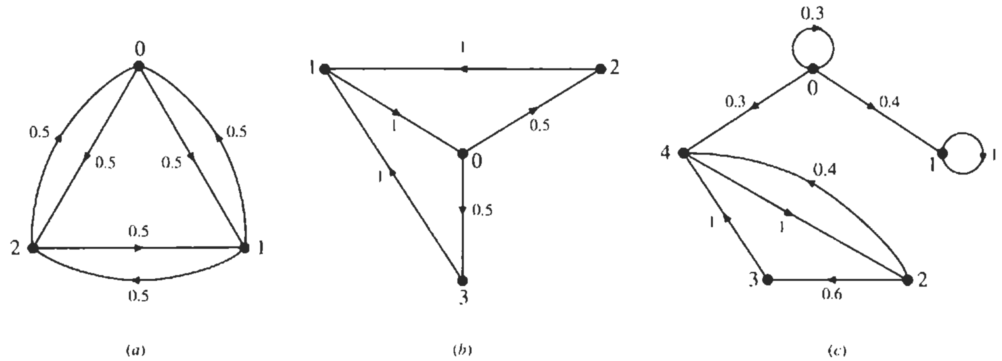

Fig. 5-10 State transition diagram.

(*a*) The state transition diagram of the Markov chain with P of part (*a*) is shown in Fig. 5-10(*a*). From Fig. 5-10(*a*), it is seen that the Markov chain is irreducible and aperiodic. For instance, one can get back to state 0 in two steps by going from 0 to 1 to 0. However, one can also get back to state 0 in three steps by going from 0 to 1 to 2 to 0. Hence 0 is aperiodic. Similarly, we can see that states 1 and 2 are also aperiodic.

(*b*) The state transition diagram of the Markov chain with P of part (*b*) is shown in Fig. 5-10(*b*). From Fig. 5-10(*b*), it is seen that the Markov chain is irreducible and periodic with period 3.

(*c*) The state transition diagram of the Markov chain with P of part (*c*) is shown in Fig. 5-10(*c*). From Fig. 5-10(*c*), it is seen that the Markov chain is not irreducible, since states 0 and 4 do not communicate, and state 1 is absorbing.

5.36. Consider a Markov chain with state space $\{0, 1\}$ and transition probability matrix

$$P = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

(*a*) Show that state 0 is recurrent.

(*b*) Show that state 1 is transient.

(*a*) By Eqs. (*5.41*) and (*5.42*), we have

$$f_{00}^{(1)} = p_{00} = 1 \qquad f_{10}^{(1)} = p_{10} = \tfrac{1}{2}$$
$$f_{00}^{(2)} = p_{01}\, f_{10}^{(1)} = (0)\tfrac{1}{2} = 0$$
$$f_{00}^{(n)} = 0 \qquad n \geq 2$$

Then, by Eqs. (*5.43*),

$$f_{00} = P(T_0 < \infty \mid X_0 = 0) = \sum_{n=0}^{\infty} f_{00}^{(n)} = 1 + 0 + 0 + \cdots = 1$$

Thus, by definition (*5.44*), state 0 is recurrent.

(*b*) Similarly, we have

$$f_{11}^{(1)} = p_{11} = \tfrac{1}{2} \qquad f_{01}^{(1)} = p_{01} = 0$$
$$f_{11}^{(2)} = p_{10}\, f_{01}^{(1)} = (\tfrac{1}{2})0 = 0$$
$$f_{11}^{(n)} = 0 \qquad n \geq 2$$

and
$$f_{11} = P(T_1 < \infty \mid X_0 = 1) = \sum_{n=0}^{\infty} f_{11}^{(n)} = \tfrac{1}{2} + 0 + 0 + \cdots = \tfrac{1}{2} < 1$$

Thus, by definition (*5.48*), state 1 is transient.

5.37. Consider a Markov chain with state space $\{0, 1, 2\}$ and transition probability matrix

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Show that state 0 is periodic with period 2.

The characteristic equation of P is given by

$$c(\lambda) = |\lambda I - P| = \begin{vmatrix} \lambda & -\frac{1}{2} & -\frac{1}{2} \\ -1 & \lambda & 0 \\ -1 & 0 & \lambda \end{vmatrix} = \lambda^3 - \lambda = 0$$

Thus, by the Cayley-Hamilton theorem (in matrix analysis), we have $P^3 = P$. Thus, for $n \ge 1$,

$$P^{(2n)} = P^2 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$P^{(2n+1)} = P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Therefore $$d(0) = \gcd\{n \ge 1 : p_{00}^{(n)} > 0\} = \gcd\{2, 5, 6, \ldots\} = 2$$

Thus, state 0 is periodic with period 2.

Note that the state transition diagram corresponding to the given P is shown in Fig. 5-11. From Fig. 5-11, it is clear that state 0 is periodic with period 2.

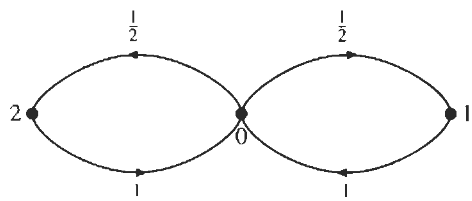

Fig. 5-11

5.38. Let two gamblers, A and B, initially have k dollars and m dollars, respectively. Suppose that at each round of their game, A wins one dollar from B with probability p and loses one dollar to B with probability $q = 1 - p$. Assume that A and B play until one of them has no money left. (This is known as the *Gambler's Ruin* problem.) Let X_n be A's capital after round n, where $n = 0, 1, 2, \ldots$ and $X_0 = k$.

(*a*) Show that $X(n) = \{X_n, n \ge 0\}$ is a Markov chain with absorbing states.

(*b*) Find its transition probability matrix P.

(*a*) The total capital of the two players at all times is

$$k + m = N$$

Let Z_n $(n \geq 1)$ be independent r.v.'s with $P(Z_n = 1) = p$ and $P(Z_n = -1) = q = 1 - p$ for all n. Then

$$X_n = X_{n-1} + Z_n \qquad n = 1, 2, \ldots$$

and $X_0 = k$. The game ends when $X_n = 0$ or $X_n = N$. Thus, by Probs. 5.2 and 5.24, $X(n) = \{X_n, n \geq 0\}$ is a Markov chain with state space $E = \{0, 1, 2, \ldots, N\}$, where states 0 and N are absorbing states. The Markov chain $X(n)$ is also known as a simple random walk with *absorbing barriers*.

(*b*) Since

$$\begin{aligned}
p_{i,i+1} &= P(X_{n+1} = i + 1 \mid X_n = i) = p \\
p_{i,i-1} &= P(X_{n+1} = i - 1 \mid X_n = i) = q \\
p_{i,i} &= P(X_{n+1} = i \mid X_n = i) = 0 \qquad i \neq 0, N \\
p_{0,0} &= P(X_{n+1} = 0 \mid X_n = 0) = 1 \\
p_{N,N} &= P(X_{n+1} = N \mid X_n = N) = 1
\end{aligned}$$

the transition probability matrix P is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ q & 0 & p & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & q & 0 & p & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \ddots & & & \vdots \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{5.123}$$

For example, when $p = q = \frac{1}{2}$ and $N = 4$,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.39. Consider a homogeneous Markov chain $X(n) = \{X_n, n \geq 0\}$ with a finite state space $E = \{0, 1, \ldots, N\}$, of which $A = \{0, 1, \ldots, m\}$, $m \geq 1$, is a set of absorbing states and $B = \{m + 1, \ldots, N\}$ is a set of nonabsorbing states. It is assumed that at least one of the absorbing states in A is accessible from any nonabsorbing states in B. Show that absorption of $X(n)$ in one or another of the absorbing states is certain.

If $X_0 \in A$, then there is nothing to prove, since $X(n)$ is already absorbed. Let $X_0 \in B$. By assumption, there is at least one state in A which is accessible from any state in B. Now assume that state $k \in A$ is accessible from $j \in B$. Let n_{jk} $(< \infty)$ be the smallest number n such that $p_{jk}^{(n)} > 0$. For a given state j, let n_j be the largest of n_{jk} as k varies and n' be the largest of n_j as j varies. After n' steps, no matter what the initial state of $X(n)$, there is a probability $p > 0$ that $X(n)$ is in an absorbing state. Therefore

$$P\{X_{n'} \in B\} = 1 - p$$

and $0 < 1 - p < 1$. It follows by homogeneity and the Markov property that

$$P\{X_{k(n')} \in B\} = (1 - p)^k \qquad k = 1, 2, \ldots$$

Now since $\lim_{k\to\infty}(1-p)^k = 0$, we have

$$\lim_{n\to\infty} P\{X_n \in B\} = 0 \qquad \text{or} \qquad \lim_{n\to\infty} P\{X_n \in \bar{B} = A\} = 1$$

which shows that absorption of $X(n)$ in one or another of the absorption states is certain.

5.40. Verify Eq. (5.50).

Let $X(n) = \{X_n, n \geq 0\}$ be a homogeneous Markov chain with a finite state space $E = \{0, 1, \ldots, N\}$, of which $A = \{0, 1, \ldots, m\}$, $m \geq 1$, is a set of absorbing states and $B = \{m+1, \ldots, N\}$ is a set of nonabsorbing states. Let state $k \in B$ at the first step go to $i \in E$ with probability p_{ki}. Then

$$\begin{aligned} u_{kj} &= P\{X_n = j(\in A) \mid X_0 = k(\in B)\} \\ &= \sum_{i=1}^{N} p_{ki} P\{X_n = j(\in A) \mid X_0 = i\} \end{aligned} \tag{5.124}$$

Now
$$P\{X_n = j(\in A), X_0 = i\} = \begin{cases} 1 & i = j \\ 0 & i \in A, i \neq j \\ u_{ij} & i \in B, i = m+1, \ldots, N \end{cases}$$

Then Eq. (5.124) becomes

$$u_{kj} = p_{kj} + \sum_{i=m+1}^{N} p_{ki} u_{ij} \qquad k = m+1, \ldots, N; j = 1, \ldots, m \tag{5.125}$$

But p_{kj}, $k = m+1, \ldots, N; j = 1, \ldots, m$, are the elements of R, whereas p_{ki}, $k = m+1, \ldots, N; i = m+1, \ldots, N$ are the elements of Q [see Eq. (5.49a)]. Hence, in matrix notation, Eq. (5.125) can be expressed as

$$U = R + QU \qquad \text{or} \qquad (I - Q)U = R \tag{5.126}$$

Premultiplying both sides of the second equation of Eq. (5.126) with $(I - Q)^{-1}$, we obtain

$$U = (I - Q)^{-1}R = \Phi R$$

5.41. Consider a simple random walk $X(n)$ with absorbing barriers at state 0 and state $N = 3$ (see Prob. 5.38).

(a) Find the transition probability matrix P.

(b) Find the probabilities of absorption into states 0 and 3.

(a) The transition probability matrix P is [Eq. (5.123)]

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

(b) Rearranging the transition probability matrix P as [Eq. (5.49a)],

$$P = \begin{matrix} & \begin{matrix} 0 & 3 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 3 \\ 1 \\ 2 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ q & 0 & 0 & p \\ 0 & p & q & 0 \end{bmatrix} \end{matrix}$$

and by Eq. (5.49b), the matrices Q and R are given by

$$R = \begin{bmatrix} p_{10} & p_{13} \\ p_{20} & p_{23} \end{bmatrix} = \begin{bmatrix} q & 0 \\ 0 & p \end{bmatrix} \qquad Q = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0 & p \\ q & 0 \end{bmatrix}$$

Then

$$I - Q = \begin{bmatrix} 1 & -p \\ -q & 1 \end{bmatrix}$$

and

$$\Phi = (I - Q)^{-1} = \frac{1}{1 - pq} \begin{bmatrix} 1 & p \\ q & 1 \end{bmatrix} \tag{5.127}$$

By Eq. (5.50),

$$U = \begin{bmatrix} u_{10} & u_{13} \\ u_{20} & u_{23} \end{bmatrix} = \Phi R = \frac{1}{1 - pq} \begin{bmatrix} 1 & p \\ q & 1 \end{bmatrix} \begin{bmatrix} q & 0 \\ 0 & p \end{bmatrix} = \frac{1}{1 - pq} \begin{bmatrix} q & p^2 \\ q^2 & p \end{bmatrix} \tag{5.128}$$

Thus, the probabilities of absorption into state 0 from states 1 and 2 are given, respectively, by

$$u_{10} = \frac{q}{1 - pq} \qquad \text{and} \qquad u_{20} = \frac{q^2}{1 - pq}$$

and the probabilities of absorption into state 3 from states 1 and 2 are given, respectively, by

$$u_{13} = \frac{p^2}{1 - pq} \qquad \text{and} \qquad u_{23} = \frac{p}{1 - pq}$$

Note that

$$u_{10} + u_{13} = \frac{q + p^2}{1 - pq} = \frac{1 - p + p^2}{1 - p(1 - p)} = 1$$

$$u_{20} + u_{23} = \frac{q^2 + p}{1 - pq} = \frac{q^2 + (1 - q)}{1 - (1 - q)q} = 1$$

which confirm the proposition of Prob. 5.39.

5.42. Consider the simple random walk $X(n)$ with absorbing barriers at 0 and 3 (Prob. 5.41). Find the expected time (or steps) to absorption when $X_0 = 1$ and when $X_0 = 2$.

The fundamental matrix Φ of $X(n)$ is [Eq. (5.127)]

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} = \frac{1}{1 - pq} \begin{bmatrix} 1 & p \\ q & 1 \end{bmatrix}$$

Let T_i be the time to absorption when $X_0 = i$. Then by Eq. (5.51), we get

$$E(T_1) = \frac{1}{1 - pq}(1 + p) \qquad E(T_2) = \frac{1}{1 - pq}(q + 1) \tag{5.129}$$

5.43. Consider the gambler's game described in Prob. 5.38. What is the probability of A's losing all his money?

Let $P(k)$, $k = 0, 1, 2, \ldots, N$, denote the probability that A loses all his money when his initial capital is k dollars. Equivalently, $P(k)$ is the probability of absorption at state 0 when $X_0 = k$ in the simple random walk $X(n)$ with absorbing barriers at states 0 and N. Now if $0 < k < N$, then

$$P(k) = pP(k+1) + qP(k-1) \qquad k = 1, 2, \ldots, N-1 \tag{5.130}$$

where $pP(k+1)$ is the probability that A wins the first round and subsequently loses all his money and $qP(k-1)$ is the probability that A loses the first round and subsequently loses all his money. Rewriting Eq. (5.130), we have

$$P(k+1) - \frac{1}{p} P(k) + \frac{q}{p} P(k-1) = 0 \qquad k = 1, 2, \ldots, N-1 \tag{5.131}$$

which is a second-order homogeneous linear constant-coefficient difference equation. Next, we have

$$P(0) = 1 \qquad \text{and} \qquad P(N) = 0 \tag{5.132}$$

since if $k = 0$, absorption at 0 is a sure event, and if $k = N$, absorption at N has occurred and absorption at 0 is impossible. Thus, finding $P(k)$ reduces to solving Eq. (5.131) subject to the boundary conditions given by Eq. (5.132). Let $P(k) = r^k$. Then Eq. (5.131) becomes

$$r^{k+1} - \frac{1}{p} r^k + \frac{q}{p} r^{k-1} = 0 \qquad p + q = 1$$

Setting $k = 1$ (and noting that $p + q = 1$), we get

$$r^2 - \frac{1}{p} r + \frac{q}{p} = (r-1)\left(r - \frac{q}{p}\right) = 0$$

from which we get $r = 1$ and $r = q/p$. Thus,

$$P(k) = c_1 + c_2\left(\frac{q}{p}\right)^k \qquad q \neq p \tag{5.133}$$

where c_1 and c_2 are arbitrary constants. Now, by Eq. (5.132),

$$P(0) = 1 \rightarrow c_1 + c_2 = 1$$

$$P(N) = 0 \rightarrow c_1 + c_2\left(\frac{q}{p}\right)^N = 0$$

Solving for c_1 and c_2, we obtain

$$c_1 = \frac{-(q/p)^N}{1-(q/p)^N} \qquad c_2 = \frac{1}{1-(q/p)^N}$$

Hence

$$P(k) = \frac{(q/p)^k - (q/p)^N}{1-(q/p)^N} \qquad q \neq p \tag{5.134}$$

Note that if $N \gg k$,

$$P(k) = \begin{cases} 1 & q > p \\ \left(\dfrac{q}{p}\right)^k & p > q \end{cases} \tag{5.135}$$

Setting $r = q/p$ in Eq. (5.134), we have

$$P(k) = \frac{r^k - r^N}{1 - r^N} \xrightarrow[r \to 1]{} 1 - \frac{k}{N}$$

Thus, when $p = q = \frac{1}{2}$,

$$P(k) = 1 - \frac{k}{N} \tag{5.136}$$

5.44. Show that Eq. (5.134) is consistent with Eq. (5.128).

Substituting $k = 1$ and $N = 3$ in Eq. (*5.134*), and noting that $p + q = 1$, we have

$$P(1) = \frac{(q/p) - (q/p)^3}{1 - (q/p)^3} = \frac{q(p^2 - q^2)}{(p^3 - q^3)}$$
$$= \frac{q(p + q)}{p^2 + pq + q^2} = \frac{q}{(p + q)^2 - pq} = \frac{q}{1 - pq}$$

Now from Eq. (*5.128*), we have

$$u_{10} = \frac{q}{1 - pq} = P(1)$$

5.45. Consider the simple random walk $X(n)$ with state space $E = \{0, 1, 2, \ldots, N\}$, where 0 and N are absorbing states (Prob. 5.38). Let r.v. T_k denote the time (or number of steps) to absorption of $X(n)$ when $X_0 = k$, $k = 0, 1, \ldots, N$. Find $E(T_k)$.

Let $Y(k) = E(T_k)$. Clearly, if $k = 0$ or $k = N$, then absorption is immediate, and we have

$$Y(0) = Y(N) = 0 \tag{5.137}$$

Let the probability that absorption takes m steps when $X_0 = k$ be defined by

$$P(k, m) = P(T_k = m) \qquad m = 1, 2, \ldots \tag{5.138}$$

Then, we have (Fig. 5-12)

$$P(k, m) = pP(k + 1, m - 1) + qP(k - 1, m - 1) \tag{5.139}$$

and $$Y(k) = E(T_k) = \sum_{m=1}^{\infty} mP(k, m) = p\sum_{m=1}^{\infty} mP(k + 1, m - 1) + q\sum_{m=1}^{\infty} mP(k - 1, m - 1)$$

Setting $m - 1 = i$, we get

$$Y(k) = p\sum_{i=0}^{\infty}(i + 1)P(k + 1, i) + q\sum_{i=0}^{\infty}(i + 1)P(k - 1, i)$$
$$= p\sum_{i=0}^{\infty} iP(k + 1, i) + q\sum_{i=0}^{\infty} iP(k - 1, i) + p\sum_{i=0}^{\infty} P(k + 1, i) + q\sum_{i=0}^{\infty} P(k - 1, i)$$

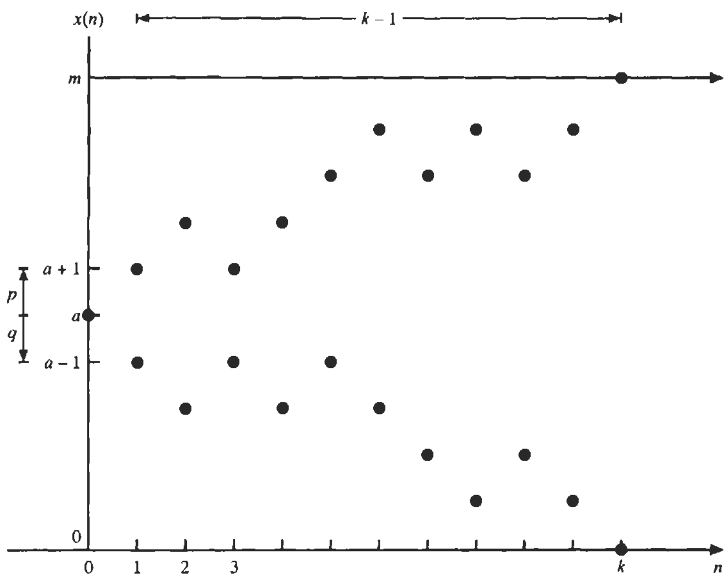

Fig. 5-12 Simple random walk with absorbing barriers.

Now by the result of Prob. 5.39, we see that absorption is certain; therefore

$$\sum_{i=0}^{\infty} P(k+1, i) = \sum_{i=0}^{\infty} P(k-1, i) = 1$$

Thus

$$Y(k) = pY(k+1) + qY(k-1) + p + q$$

or

$$Y(k) = pY(k+1) + qY(k-1) + 1 \qquad k = 1, 2, \ldots, N-1 \tag{5.140}$$

Rewriting Eq. (5.140), we have

$$Y(k+1) - \frac{1}{p} Y(k) + \frac{q}{p} Y(k-1) = -\frac{1}{p} \tag{5.141}$$

Thus, finding $P(k)$ reduces to solving Eq. (5.141) subject to the boundary conditions given by Eq. (5.137). Let the general solution of Eq. (5.141) be

$$Y(k) = Y_h(k) + Y_p(k)$$

where $Y_h(k)$ is the homogeneous solution satisfying

$$Y_h(k+1) - \frac{1}{p} Y_h(k) + \frac{q}{p} Y_h(k-1) = 0 \tag{5.142}$$

and $Y_p(k)$ is the particular solution satisfying

$$Y_p(k+1) - \frac{1}{p} Y_p(k) + \frac{q}{p} Y_p(k-1) = -\frac{1}{p} \tag{5.143}$$

Let $Y_p(k) = \alpha k$, where α is a constant. Then Eq. (5.143) becomes

$$(k+1)\alpha - \frac{1}{p} k\alpha + \frac{q}{p}(k-1)\alpha = -\frac{1}{p}$$

from which we get $\alpha = 1/(q-p)$ and

$$Y_p(k) = \frac{k}{q-p} \qquad p \neq q \tag{5.144}$$

Since Eq. (5.142) is the same as Eq. (5.131), by Eq. (5.133), we obtain

$$Y_h(k) = c_1 + c_2\left(\frac{q}{p}\right)^k \qquad q \neq p \tag{5.145}$$

where c_1 and c_2 are arbitrary constants. Hence, the general solution of Eq. (5.141) is

$$Y(k) = c_1 + c_2\left(\frac{q}{p}\right)^k + \frac{k}{q-p} \qquad q \neq p \tag{5.146}$$

Now, by Eq. (5.137),

$$Y(0) = 0 \rightarrow c_1 + c_2 = 0$$

$$Y(N) = 0 \rightarrow c_1 + c_2\left(\frac{q}{p}\right)^N + \frac{N}{q-p} = 0$$

Solving for c_1 and c_2, we obtain

$$c_1 = \frac{-N/(q-p)}{1-(q/p)^N} \qquad c_2 = \frac{N/(q-p)}{1-(q/p)^N}$$

Substituting these values in Eq. (5.146), we obtain (for $p \neq q$)

$$Y(k) = E(T_k) = \frac{1}{q-p}\left(k - N\left[\frac{1-(q/p)^k}{1-(q/p)^N}\right]\right) \tag{5.147}$$

When $p = q = \frac{1}{2}$, we have

$$Y(k) = E(T_k) = k(N-k) \qquad p = q = \tfrac{1}{2} \tag{5.148}$$

5.46. Consider a Markov chain with two states and transition probability matrix

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(a) Find the stationary distribution $\hat{\mathbf{p}}$ of the chain.
(b) Find $\lim_{n \to \infty} P^n$.

(a) By definition (5.52),

$$\hat{\mathbf{p}} P = \hat{\mathbf{p}}$$

or
$$[p_1 \quad p_2] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [p_1 \quad p_2]$$

which yields $p_1 = p_2$. Since $p_1 + p_2 = 1$, we obtain

$$\hat{\mathbf{p}} = [\tfrac{1}{2} \quad \tfrac{1}{2}]$$

(b) Now
$$P^n = \begin{cases} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & n = 1, 3, 5, \ldots \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & n = 2, 4, 6, \ldots \end{cases}$$

and $\lim_{n \to \infty} P^n$ does not exist.

5.47. Consider a Markov chain with two states and transition probability matrix

$$P = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

(a) Find the stationary distribution $\hat{\mathbf{p}}$ of the chain.
(b) Find $\lim_{n \to \infty} P^n$.
(c) Find $\lim_{n \to \infty} P^n$ by first evaluating P^n.

(a) By definition (5.52), we have

$$\hat{\mathbf{p}} P = \hat{\mathbf{p}}$$

or
$$[p_1 \quad p_2] \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [p_1 \quad p_2]$$

which yields

$$\tfrac{3}{4} p_1 + \tfrac{1}{2} p_2 = p_1$$
$$\tfrac{1}{4} p_1 + \tfrac{1}{2} p_2 = p_2$$

Each of these equations is equivalent to $p_1 = 2p_2$. Since $p_1 + p_2 = 1$, we obtain

$$\hat{\mathbf{p}} = [\tfrac{2}{3} \quad \tfrac{1}{3}]$$

(b) Since the Markov chain is regular, by Eq. (5.53), we obtain

$$\lim_{n \to \infty} P^n = \lim_{n \to \infty} \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}^n = \begin{bmatrix} \hat{\mathbf{p}} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

(c) Setting $a = \frac{1}{4}$ and $b = \frac{1}{2}$ in Eq. (5.120) (Prob. 5.30), we get

$$P^n = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} - \left(\tfrac{1}{4}\right)^n \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

Since $\lim_{n \to \infty} \left(\frac{1}{4}\right)^n = 0$, we obtain

$$\lim_{n \to \infty} P^n = \lim_{n \to \infty} \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}^n = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

POISSON PROCESSES

5.48. Let T_n denote the arrival time of the nth customer at a service station. Let Z_n denote the time interval between the arrival of the nth customer and the $(n-1)$st customer; that is,

$$Z_n = T_n - T_{n-1} \qquad n \geq 1 \tag{5.149}$$

and $T_0 = 0$. Let $\{X(t),\ t \geq 0\}$ be the counting process associated with $\{T_n,\ n \geq 0\}$. Show that if $X(t)$ has stationary increments, then Z_n, $n = 1, 2, \ldots$, are identically distributed r.v.'s.

We have

$$P(Z_n > z) = 1 - P(Z_n \leq z) = 1 - F_{Z_n}(z)$$

By Eq. (5.149),
$$P(Z_n > z) = P(T_n - T_{n-1} > z) = P(T_n > T_{n-1} + z)$$

Suppose that the observed value of T_{n-1} is t_{n-1}. The event $(T_n > T_{n-1} + z \mid T_{n-1} = t_{n-1})$ occurs if and only if $X(t)$ does not change count during the time interval $(t_{n-1}, t_{n-1} + z)$ (Fig. 5-13). Thus,

$$\begin{aligned} P(Z_n > z \mid T_{n-1} = t_{n-1}) &= P(T_n > T_{n-1} + z \mid T_{n-1} = t_{n-1}) \\ &= P[X(t_{n-1} + z) - X(t_{n-1}) = 0] \end{aligned}$$

or
$$P(Z_n \leq z \mid T_{n-1} = t_{n-1}) = 1 - P[X(t_{n-1} + z) - X(t_{n-1}) = 0] \tag{5.150}$$

Since $X(t)$ has stationary increments, the probability on the right-hand side of Eq. (5.150) is a function only of the time difference z. Thus

$$P(Z_n \leq z \mid T_{n-1} = t_{n-1}) = 1 - P[X(z) = 0] \tag{5.151}$$

which shows that the conditional distribution function on the left-hand side of Eq. (5.151) is independent of the particular value of n in this case, and hence we have

$$F_{Z_n}(z) = P(Z_n \leq z) = 1 - P[X(z) = 0] \tag{5.152}$$

which shows that the cdf of Z_n is independent of n. Thus we conclude that the Z_n's are identically distributed r.v.'s.

Fig. 5-13

5.49. Show that Definition 5.6.2 implies Definition 5.6.1.

Let $p_n(t) = P[X(t) = n]$. Then, by condition 2 of Definition 5.6.2, we have

$$\begin{aligned} p_0(t + \Delta t) = P[X(t + \Delta t) = 0] &= P[X(t) = 0,\ X(t + \Delta t) - X(0) = 0] \\ &= P[X(t) = 0]P[X(t + \Delta t) - X(t) = 0] \end{aligned}$$

Now, by Eq. (5.59), we have

$$P[X(t+\Delta t)-X(t)=0]=1-\lambda\,\Delta t+o(\Delta t)$$

Thus,
$$p_0(t+\Delta t)=p_0(t)[1-\lambda\,\Delta t+o(\Delta t)]$$

or
$$\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t}=-\lambda p_0(t)+\frac{o(\Delta t)}{\Delta t}$$

Letting $\Delta t \to 0$, and by Eq. (5.58), we obtain

$$p_0'(t)=-\lambda p_0(t) \tag{5.153}$$

Solving the above differential equation, we get

$$p_0(t)=ke^{-\lambda t}$$

where k is an integration constant. Since $p_0(0)=P[X(0)=0]=1$, we obtain

$$p_0(t)=e^{-\lambda t} \tag{5.154}$$

Similarly, for $n > 0$,

$$\begin{aligned}p_n(t+\Delta t)&=P[X(t+\Delta t)=n]\\&=P[X(t)=n,\ X(t+\Delta t)-X(0)=0]\\&\quad+P[X(t)=n-1,\ X(t+\Delta t)-X(0)=1]+\sum_{k=2}^{n}P[X(t)=n-k,\ X(t+\Delta t)-X(0)=k]\end{aligned}$$

Now, by condition 4 of Definition 5.6.2, the last term in the above expression is $o(\Delta t)$. Thus, by conditions 2 and 3 of Definition 5.6.2, we have

$$p_n(t+\Delta t)=p_n(t)[1-\lambda\,\Delta t+o(\Delta t)]+p_{n-1}(t)[\lambda\,\Delta t+o(\Delta t)]+o(\Delta t)$$

Thus
$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=-\lambda p_n(t)+\lambda p_{n-1}(t)+\frac{o(\Delta t)}{\Delta t}$$

and letting $\Delta t \to 0$ yields

$$p_n'(t)+\lambda p_n(t)=\lambda p_{n-1}(t) \tag{5.155}$$

Multiplying both sides by $e^{\lambda t}$, we get

$$e^{\lambda t}[p_n'(t)+\lambda p_n(t)]=\lambda e^{\lambda t}p_{n-1}(t)$$

Hence
$$\frac{d}{dt}[e^{\lambda t}p_n(t)]=\lambda e^{\lambda t}p_{n-1}(t) \tag{5.156}$$

Then by Eq. (5.154), we have

$$\frac{d}{dt}[e^{\lambda t}p_1(t)]=\lambda$$

or
$$p_1(t)=(\lambda t+c)e^{-\lambda t}$$

where c is an integration constant. Since $p_1(0)=P[X(0)=1]=0$, we obtain

$$p_1(t)=\lambda te^{-\lambda t} \tag{5.157}$$

To show that

$$p_n(t)=e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

we use mathematical induction. Assume that it is true for $n-1$; that is,

$$p_{n-1}(t)=e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

Substituting the above expression into Eq. (5.156), we have

$$\frac{d}{dt}[e^{\lambda t}p_n(t)] = \frac{\lambda^n t^{n-1}}{(n-1)!}$$

Integrating, we get

$$e^{\lambda t}p_n(t) = \frac{(\lambda t)^n}{n!} + c_1$$

Since $p_n(0) = 0$, $c_1 = 0$, and we obtain

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \tag{5.158}$$

which is Eq. (5.55) of Definition 5.6.1. Thus we conclude that Definition 5.6.2 implies Definition 5.6.1.

5.50. Verify Eq. (5.59).

We note first that $X(t)$ can assume only nonnegative integer values; therefore, the same is true for the counting increment $X(t + \Delta t) - X(t)$. Thus, summing over all possible values of the increment, we get

$$\begin{aligned}\sum_{k=0}^{\infty} P[X(t+\Delta t) - X(t) = k] &= P[X(t+\Delta t) - X(t) = 0] \\ &\quad + P[X(t+\Delta t) - X(t) = 1] + P[X(t+\Delta t) - X(t) \geq 2] \\ &= 1\end{aligned}$$

Substituting conditions 3 and 4 of Definition 5.6.2 into the above equation, we obtain

$$P[X(t+\Delta t) - X(t) = 0] = 1 - \lambda\,\Delta t + o(\Delta t)$$

5.51. (*a*) Using the Poison probability distribution in Eq. (5.158), obtain an analytical expression for the correction term $o(\Delta t)$ in the expression (condition 3 of Definition 5.6.2)

$$P[X(t+\Delta t) - X(t) = 1] = \lambda\,\Delta t + o(\Delta t) \tag{5.159}$$

(*b*) Show that this correction term does have the property of Eq. (5.58); that is,

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

(*a*) Since the Poisson process $X(t)$ has stationary increments, Eq. (5.159) can be rewritten as

$$P[X(\Delta t) = 1] = p_1(\Delta t) = \lambda\,\Delta t + o(\Delta t) \tag{5.160}$$

Using Eq. (5.158) [or Eq. (5.157)], we have

$$\begin{aligned}p_1(\Delta t) &= \lambda\,\Delta t\, e^{-\lambda\Delta t} = \lambda\,\Delta t(1 + e^{-\lambda\Delta t} - 1) \\ &= \lambda\,\Delta t + \lambda\,\Delta t(e^{-\lambda\Delta t} - 1)\end{aligned}$$

Equating the above expression with Eq. (5.160), we get

$$\lambda\,\Delta t + o(\Delta t) = \lambda\,\Delta t + \lambda\,\Delta t(e^{-\lambda\Delta t} - 1)$$

from which we obtain

$$o(\Delta t) = \lambda\,\Delta t(e^{-\lambda\Delta t} - 1) \tag{5.161}$$

(*b*) From Eq. (5.161), we have

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\lambda\,\Delta t(e^{-\lambda\Delta t} - 1)}{\Delta t} = \lim_{\Delta t \to 0} \lambda(e^{-\lambda\Delta t} - 1) = 0$$

5.52. Find the autocorrelation function $R_X(t, s)$ and the autocovariance function $K_X(t, s)$ of a Poisson process $X(t)$ with rate λ.

From Eqs. (5.56) and (5.57),

$$E[X(t)] = \lambda t \qquad \text{Var}[X(t)] = \lambda t$$

Now, the Poisson process $X(t)$ is a random process with stationary independent increments and $X(0) = 0$. Thus, by Eq. (5.103) (Prob. 5.23), we obtain

$$K_X(t, s) = \sigma_1^2 \min(t, s) = \lambda \min(t, s) \tag{5.162}$$

since $\sigma_1^2 = \text{Var}[X(1)] = \lambda$. Next, since $E[X(t)]E[X(s)] = \lambda^2 ts$, by Eq. (5.10), we obtain

$$R_X(t, s) = \lambda \min(t, s) + \lambda^2 ts \tag{5.163}$$

5.53. Show that the time intervals between successive events (or interarrival times) in a Poisson process $X(t)$ with rate λ are independent and identically distributed exponential r.v.'s with parameter λ.

Let $Z_1, Z_2, \ldots$ be the r.v.'s representing the lengths of interarrival times in the Poisson process $X(t)$. First, notice that $\{Z_1 > t\}$ takes place if and only if no event of the Poisson process occur in the interval $(0, t)$, and thus by Eq. (5.154),

$$P(Z_1 > t) = P\{X(t) = 0\} = e^{-\lambda t}$$

or

$$F_{Z_1}(t) = P(Z_1 \le t) = 1 - e^{-\lambda t}$$

Hence Z_1 is an exponential r.v. with parameter λ [Eq. (2.49)]. Let $f_1(t)$ be the pdf of Z_1. Then we have

$$\begin{aligned} P(Z_2 > t) &= \int P(Z_2 > t \mid Z_1 = \tau) f_1(\tau)\, d\tau \\ &= \int P[X(t + \tau) - X(\tau) = 0] f_1(\tau)\, d\tau \\ &= e^{-\lambda t} \int f_1(\tau)\, d\tau = e^{-\lambda t} \end{aligned} \tag{5.164}$$

which indicates that Z_2 is also an exponential r.v. with parameter λ and is independent of Z_1. Repeating the same argument, we conclude that $Z_1, Z_2, \ldots$ are iid exponential r.v.'s with parameter λ.

5.54. Let T_n denote the time of the nth event of a Poisson process $X(t)$ with rate λ. Show that T_n is a gamma r.v. with parameters (n, λ).

Clearly,

$$T_n = Z_1 + Z_2 + \cdots + Z_n$$

where Z_n, $n = 1, 2, \ldots$, are the interarrival times defined by Eq. (5.149). From Prob. 5.53, we know that Z_n are iid exponential r.v.'s with parameter λ. Now, using the result of Prob. 4.33, we see that T_n is a gamma r.v. with parameters (n, λ), and its pdf is given by [Eq. (2.76)]:

$$f_{T_n}(t) = \begin{cases} \lambda e^{-\lambda t} \dfrac{(\lambda t)^{n-1}}{(n-1)!} & t > 0 \\ 0 & t < 0 \end{cases} \tag{5.165}$$

The random process $\{T_n, n \ge 1\}$ is often called an *arrival process*.

5.55. Suppose t is not a point at which an event occurs in a Poisson process $X(t)$ with rate λ. Let $W(t)$ be the r.v. representing the time until the next occurrence of an event. Show that the distribution of $W(t)$ is independent of t and $W(t)$ is an exponential r.v. with parameter λ.

Let s $(0 \le s < t)$ be the point at which the last event [say the $(n-1)$st event] occurred (Fig. 5-14). The event $\{W(t) > \tau\}$ is equivalent to the event

$$\{Z_n > t - s + \tau \mid Z_n > t - s\}$$

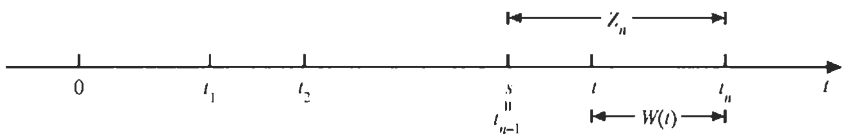

Fig. 5-14

Thus, using Eq. (*5.164*), we have

$$P[W(t) > \tau] = P(Z_n > t - s + \tau \mid Z_n > t - s)$$
$$= \frac{P(Z_n > t - s + \tau)}{P(Z_n > t - s)} = \frac{e^{-\lambda(t-s+\tau)}}{e^{-\lambda(t-s)}} = e^{-\lambda\tau}$$

and
$$P[W(t) \le \tau] = 1 - e^{\lambda\tau} \tag{5.166}$$

which indicates that $W(t)$ is an exponential r.v. with parameter λ and is independent of t. Note that $W(t)$ is often called a *waiting time*.

5.56. Patients arrive at the doctor's office according to a Poisson process with rate $\lambda = \frac{1}{10}$ minute. The doctor will not see a patient until at least three patients are in the waiting room.

(*a*) Find the expected waiting time until the first patient is admitted to see the doctor.

(*b*) What is the probability that nobody is admitted to see the doctor in the first hour?

(*a*) Let T_n denote the arrival time of the nth patient at the doctor's office. Then

$$T_n = Z_1 + Z_2 + \cdots + Z_n$$

where Z_n, $n = 1, 2, \ldots$, are iid exponential r.v.'s with parameter $\lambda = \frac{1}{10}$. By Eqs. (*4.108*) and (*2.50*),

$$E(T_n) = E\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=i}^{n} E(Z_i) = n\,\frac{1}{\lambda} \tag{5.167}$$

The expected waiting time until the first patient is admitted to see the doctor is

$$E(T_3) = 3(10) = 30 \text{ minutes}$$

(*b*) Let $X(t)$ be the Poisson process with parameter $\lambda = \frac{1}{10}$. The probability that nobody is admitted to see the doctor in the first hour is the same as the probability that at most two patients arrive in the first 60 minutes. Thus, by Eq. (*5.55*),

$$P[X(60) - X(0) \le 2] = P[X(60) - X(0) = 0] + P[X(60) - X(0) = 1] + P[X(60) - X(0) = 2]$$
$$= e^{-60/10} + e^{-60/10}\left(\tfrac{60}{10}\right) + e^{-60/10}\tfrac{1}{2}\left(\tfrac{60}{10}\right)^2$$
$$= e^{-6}(1 + 6 + 18) \approx 0.062$$

5.57. Let T_n denote the time of the nth event of a Poisson process $X(t)$ with rate λ. Suppose that one event has occurred in the interval $(0, t)$. Show that the conditional distribution of arrival time T_1 is uniform over $(0, t)$.

For $\tau \le t$,

$$P[T_1 \le \tau \mid X(t) = 1] = \frac{P[T_1 \le \tau, X(t) = 1]}{P[X(t) = 1]}$$
$$= \frac{P[X(\tau) = 1, X(t) - X(\tau) = 0]}{P[X(t) = 1]}$$
$$= \frac{P[X(\tau) = 1]P[X(t) - X(\tau) = 0]}{P[X(t) = 1]}$$
$$= \frac{\lambda\tau e^{-\lambda\tau}e^{-\lambda(t-\tau)}}{\lambda t e^{-\lambda t}} = \frac{\tau}{t} \tag{5.168}$$

which indicates that T_1 is uniform over $(0, t)$ [see Eq. (*2.45*)].

5.58. Consider a Poisson process $X(t)$ with rate λ, and suppose that each time an event occurs, it is classified as either a type 1 or a type 2 event. Suppose further that the event is classified as a type 1 event with probability p and a type 2 event with probability $1 - p$. Let $X_1(t)$ and $X_2(t)$ denote the number of type 1 and type 2 events, respectively, occurring in $(0, t)$. Show that $\{X_1(t), t \geq 0\}$ and $\{X_2(t), t \geq 0\}$ are both Poisson processes with rates λp and $\lambda(1 - p)$, respectively. Furthermore, the two processes are independent.

We have

$$X(t) = X_1(t) + X_2(t)$$

First we calculate the joint probability $P[X_1(t) = k, X_2(t) = m]$.

$$P[X_1(t) = k, X_2(t) = m] = \sum_{n=0}^{\infty} P[X_1(t) = k, X_2(t) = m \mid X(t) = n]P[X(t) = n]$$

Note that

$$P[X_1(t) = k, X_2(t) = m \mid X(t) = n] = 0 \qquad \text{when } n \neq k + m$$

Thus, using Eq. (5.158), we obtain

$$\begin{aligned} P[X_1(t) = k, X_2(t) = m] &= P[X_1(t) = k, X_2(t) = m \mid X(t) = k + m]P[X(t) = k + m] \\ &= P[X_1(t) = k, X_2(t) = m \mid X(t) = k + m]e^{-\lambda t}\frac{(\lambda t)^{k+m}}{(k+m)!} \end{aligned}$$

Now, given that $k + m$ events occurred, since each event has probability p of being a type 1 event and probability $1 - p$ of being a type 2 event, it follows that

$$P[X_1(t) = k, X_2(t) = m \mid X(t) = k + m] = \binom{k+m}{k}p^k(1-p)^m$$

Thus,

$$\begin{aligned} P[X_1(t) = k, X_2(t) = m] &= \binom{k+m}{k}p^k(1-p)^m e^{-\lambda t}\frac{(\lambda t)^{k+m}}{(k+m)!} \\ &= \frac{(k+m)!}{k!\,m!}p^k(1-p)^m e^{-\lambda t}\frac{(\lambda t)^{k+m}}{(k+m)!} \\ &= e^{-\lambda pt}\frac{(\lambda pt)^k}{k!}e^{-\lambda(1-p)t}\frac{[\lambda(1-p)t]^m}{m!} \end{aligned} \tag{5.169}$$

Then

$$\begin{aligned} P[X_1(t) = k] &= \sum_{m=1}^{\infty} P[X_1(t) = k, X_2(t) = m] \\ &= e^{-\lambda pt}\frac{(\lambda pt)^k}{k!}e^{-\lambda(1-p)t}\sum_{m=1}^{\infty}\frac{[\lambda(1-p)t]^m}{m!} \\ &= e^{-\lambda pt}\frac{(\lambda pt)^k}{k!}e^{-\lambda(1-p)t}e^{\lambda(1-p)t} \\ &= e^{-\lambda pt}\frac{(\lambda pt)^k}{k!} \end{aligned} \tag{5.170}$$

which indicates that $X_1(t)$ is a Poisson process with rate λp. Similarly, we can obtain

$$\begin{aligned} P[X_2(t) = m] &= \sum_{k=1}^{\infty} P[X_1(t) = k, X_2(t) = m] \\ &= e^{-\lambda(1-p)t}\frac{[\lambda(1-p)t]^m}{m!} \end{aligned} \tag{5.171}$$

and so $X_2(t)$ is a Poisson process with rate $\lambda(1 - p)$. Finally, from Eqs. (5.170), (5.171), and (5.169), we see that

$$P[X_1(t) = k, X_2(t) = m] = P[X_1(t) = k]P[X_2(t) = m]$$

Hence, $X_1(t)$ and $X_2(t)$ are independent.

WIENER PROCESSES

5.59. Let $X_1, \ldots, X_n$ be jointly normal r.v.'s. Show that the joint characteristic function of $X_1, \ldots, X_n$ is given by

$$\Psi_{X_1 \cdots X_n}(\omega_1, \ldots, \omega_n) = \exp\left(j \sum_{i=1}^{n} \omega_i \mu_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \omega_i \omega_k \sigma_{ik} \right) \tag{5.172}$$

where $\mu_i = E(X_i)$ and $\sigma_{ik} = \text{Cov}(X_i, X_k)$.

Let $$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

By definition (*4.50*), the characteristic function of Y is

$$\Psi_Y(\omega) = E[e^{j\omega(a_1 X_1 + \cdots + a_n X_n)}] = \Psi_{X_1 \cdots X_n}(\omega a_1, \ldots, \omega a_n) \tag{5.173}$$

Now, by the results of Prob. 4.55, we see that Y is a normal r.v. with mean and variance given by [Eqs. (*4.108*) and (*4.111*)]

$$\mu_Y = E(Y) = \sum_{i=1}^{n} a_i E(X_i) = \sum_{i=1}^{n} a_i \mu_i \tag{5.174}$$

$$\sigma_Y^2 = \text{Var}(Y) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_i a_k \,\text{Cov}(X_i, X_k) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_i a_k \sigma_{ik} \tag{5.175}$$

Thus, by Eq. (*4.125*),

$$\begin{aligned} \Psi_Y(\omega) &= \exp[j\omega\mu_Y - \tfrac{1}{2}\sigma_Y^2 \omega^2] \\ &= \exp\left(j\omega \sum_{i=1}^{n} a_i \mu_i - \tfrac{1}{2}\omega^2 \sum_{i=1}^{n} \sum_{k=1}^{n} a_i a_k \sigma_{ik} \right) \end{aligned} \tag{5.176}$$

Equating Eqs. (*5.176*) and (*5.173*) and setting $\omega = 1$, we get

$$\Psi_{X_1 \cdots X_n}(a_1, \ldots, a_n) = \exp\left(j \sum_{i=1}^{n} a_i \mu_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} a_i a_k \sigma_{ik} \right)$$

By replacing a_i's with ω_i's, we obtain Eq. (*5.172*); that is,

$$\Psi_{X_1 \cdots X_n}(\omega_1, \ldots, \omega_n) = \exp\left(j \sum_{i=1}^{n} \omega_i \mu_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \omega_i \omega_k \sigma_{ik} \right)$$

Let $$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} \qquad \boldsymbol{\omega} = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} \qquad K = [\sigma_{ik}] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix}$$

Then we can write

$$\sum_{i=1}^{n} \omega_i \mu_i = \boldsymbol{\omega}^T \boldsymbol{\mu} \qquad \sum_{i=1}^{n} \sum_{k=1}^{n} \omega_i \omega_k \sigma_{ik} = \boldsymbol{\omega}^T K \boldsymbol{\omega}$$

and Eq. (*5.172*) can be expressed more compactly as

$$\Psi_{X_1 \cdots X_n}(\omega_1, \ldots, \omega_n) = \exp(j\boldsymbol{\omega}^T \boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{\omega}^T K \boldsymbol{\omega}) \tag{5.177}$$

5.60. Let $X_1, \ldots, X_n$ be jointly normal r.v.'s Let

$$\begin{aligned} Y_1 &= a_{11} X_1 + \cdots + a_{1n} X_n \\ &\vdots \\ Y_m &= a_{m1} X_1 + \cdots + a_{mn} X_n \end{aligned} \tag{5.178}$$

where a_{ik} $(i = 1, \ldots, m; j = 1, \ldots, n)$ are constants. Show that $Y_1, \ldots, Y_m$ are also jointly normal r.v.'s.

Let
$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix} \qquad A = [a_{ik}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Then Eq. (*5.178*) can be expressed as

$$\mathbf{Y} = A\mathbf{X} \tag{5.179}$$

Let
$$\boldsymbol{\mu}_{\mathbf{X}} = E(\mathbf{X}) = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} \qquad \boldsymbol{\omega} = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_m \end{bmatrix} \qquad K_{\mathbf{X}} = [\sigma_{ik}] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix}$$

Then the characteristic function for $\mathbf{Y}$ can be written as

$$\begin{aligned} \Psi_{\mathbf{Y}}(\omega_1, \ldots, \omega_m) &= E(e^{j\boldsymbol{\omega}^T\mathbf{Y}}) = E(e^{j\boldsymbol{\omega}^T A\mathbf{X}}) \\ &= E[e^{j(A^T\boldsymbol{\omega})^T\mathbf{X}}] = \Psi_{\mathbf{X}}(A^T\boldsymbol{\omega}) \end{aligned}$$

Since $\mathbf{X}$ is a normal random vector, by Eq. (*5.177*) we can write

$$\begin{aligned} \Psi_{\mathbf{X}}(A^T\boldsymbol{\omega}) &= \exp[j(A^T\boldsymbol{\omega})^T\boldsymbol{\mu}_{\mathbf{X}} - \tfrac{1}{2}(A^T\boldsymbol{\omega})^T K_{\mathbf{X}}(A^T\boldsymbol{\omega})] \\ &= \exp[j\boldsymbol{\omega}^T A\boldsymbol{\mu}_{\mathbf{X}} - \tfrac{1}{2}\boldsymbol{\omega}^T A K_{\mathbf{X}} A^T\boldsymbol{\omega}] \end{aligned}$$

Thus
$$\Psi_{\mathbf{Y}}(\omega_1, \ldots, \omega_m) = \exp(j\boldsymbol{\omega}^T\boldsymbol{\mu}_{\mathbf{Y}} - \tfrac{1}{2}\boldsymbol{\omega}^T K_{\mathbf{Y}}\boldsymbol{\omega}) \tag{5.180}$$

where
$$\boldsymbol{\mu}_{\mathbf{Y}} = A\boldsymbol{\mu}_{\mathbf{X}} \qquad K_{\mathbf{Y}} = A K_{\mathbf{X}} A^T \tag{5.181}$$

Comparing Eqs. (*5.177*) and (*5.180*), we see that Eq. (*5.180*) is the characteristic function of a random vector $\mathbf{Y}$. Hence, we conclude that $Y_1, \ldots, Y_m$ are also jointly normal r.v.'s

Note that on the basis of the above result, we can say that a random process $\{X(t), t \in T\}$ is a normal process if every finite linear combination of the r.v.'s $X(t_i)$, $t_i \in T$ is normally distributed.

5.61. Show that a Wiener process $X(t)$ is a normal process.

Consider an arbitrary linear combination

$$\sum_{i=1}^{n} a_i X(t_i) = a_1 X(t_1) + a_2 X(t_2) + \cdots + a_n X(t_n) \tag{5.182}$$

where $0 \le t_1 < \cdots < t_n$ and a_i are real constants. Now we write

$$\begin{aligned} \sum_{i=1}^{n} a_i X(t_i) &= (a_1 + \cdots + a_n)[X(t_1) - X(0)] + (a_2 + \cdots + a_n)[X(t_2) - X(t_1)] \\ &\quad + \cdots + (a_{n-1} + a_n)[X(t_{n-1}) - X(t_{n-2})] + a_n[X(t_n) - X(t_{n-1})] \end{aligned} \tag{5.183}$$

Now from conditions 1 and 2 of Definition 5.7.1, the right-hand side of Eq. (*5.183*) is a linear combination of independent normal r.v.'s. Thus, based on the result of Prob. 5.60, the left-hand side of Eq. (*5.183*) is also a normal r.v.; that is, every finite linear combination of the r.v.'s $X(t_i)$ is a normal r.v. Thus we conclude that the Wiener process $X(t)$ is a normal process.

5.62. A random process $\{X(t), t \in T\}$ is said to be *continuous in probability* if for every $\varepsilon > 0$ and $t \in T$,

$$\lim_{h \to 0} P\{|X(t+h) - X(t)| > \varepsilon\} = 0 \tag{5.184}$$

Show that a Wiener process $X(t)$ is continuous in probability.

From Chebyshev inequality (*2.97*), we have

$$P\{|X(t+h) - X(t)| > \varepsilon\} \le \frac{\text{Var}[X(t+h) - X(t)]}{\varepsilon^2} \qquad \varepsilon > 0$$

Since $X(t)$ has stationary increments, we have

$$\text{Var}[X(t+h)-X(t)] = \text{Var}[X(h)] = \sigma^2 h$$

in view of Eq. (5.63). Hence,

$$\lim_{h\to 0} P\{|X(t+h)-X(t)| > \varepsilon\} = \lim_{h\to 0} \frac{\sigma^2 h}{\varepsilon^2} = 0$$

Thus the Wiener process $X(t)$ is continuous in probability.

Supplementary Problems

5.63. Consider a random process $X(n) = \{X_n, n \geq 1\}$, where

$$X_n = Z_1 + Z_2 + \cdots + Z_n$$

and Z_n are iid r.v.'s with zero mean and variance σ^2. Is $X(n)$ stationary?

Ans. No.

5.64. Consider a random process $X(t)$ defined by

$$X(t) = Y\cos(\omega t + \Theta)$$

where Y and Θ are independent r.v.'s and are uniformly distributed over $(-A, A)$ and $(-\pi, \pi)$, respectively.

(a) Find the mean of $X(t)$.

(b) Find the autocorrelation function $R_X(t, s)$ of $X(t)$.

Ans. (a) $E[X(t)] = 0$; (b) $R_X(t, s) = \frac{1}{6}A^2 \cos\omega(t-s)$

5.65. Suppose that a random process $X(t)$ is wide-sense stationary with autocorrelation

$$R_X(t, t+\tau) = e^{-|\tau|/2}$$

(a) Find the second moment of the r.v. $X(5)$.

(b) Find the second moment of the r.v. $X(5) - X(3)$.

Ans. (a) $E[X^2(5)] = 1$; (b) $E\{[X(5) - X(3)]^2\} = 2(1 - e^{-1})$

5.66. Consider a random process $X(t)$ defined by

$$X(t) = U\cos t + (V+1)\sin t \qquad -\infty < t < \infty$$

where U and V are independent r.v.'s for which

$$E(U) = E(V) = 0 \qquad E(U^2) = E(V^2) = 1$$

(a) Find the autocovariance function $K_X(t, s)$ of $X(t)$.

(b) Is $X(t)$ WSS?

Ans. (a) $K_X(t, s) = \cos(s - t)$; (b) No.

5.67. Consider the random processes

$$X(t) = A_0\cos(\omega_0 t + \Theta) \qquad Y(t) = A_1\cos(\omega_1 t + \Phi)$$

where A_0, A_1, ω_0, and ω_1 are constants, and r.v.'s Θ and Φ are independent and uniformly distributed over $(-\pi, \pi)$.

(a) Find the cross-correlation function of $R_{XY}(t, t+\tau)$ of $X(t)$ and $Y(t)$.

(b) Repeat (a) if $\Theta = \phi$.

Ans. (a) $R_{XY}(t, t+\tau)] = 0$

(b) $R_{XY}(t, t+\tau) = \dfrac{A_0 A_1}{2} \cos[(\omega_1 - \omega_0)t + \omega_1 \tau]$

5.68. Given a Markov chain $\{X_n, n \geq 0\}$, find the joint pmf

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)$$

Hint: Use Eq. (5.32).

Ans. $p_{i_0}(0)p_{i_0 i_1}p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$

5.69. Let $\{X_n, n \geq 0\}$ be a homogeneous Markov chain. Show that

$$P(X_{n+1} = k_1, \ldots, X_{n+m} = k_m \mid X_0 = i_0, \ldots, X_n = i) = P(X_1 = k_1, \ldots, X_m = k_m \mid X_0 = i)$$

Hint: Use the Markov property (5.27) and the homogeneity property.

5.70. Verify Eq. (5.37).

Hint: Write Eq. (5.39) in terms of components.

5.71. Find P^n for the following transition probability matrices:

(a) $P = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}$ (b) $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}$

Ans. (a) $P^n = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + (0.5)^n \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix}$ (b) $P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.6 & 0.4 & 0 \end{bmatrix} + (0.5)^n \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -0.6 & -0.4 & 1 \end{bmatrix}$

5.72. A certain product is made by two companies, A and B, that control the entire market. Currently, A and B have 60 percent and 40 percent, respectively, of the total market. Each year, A loses $\frac{2}{3}$ of its market share to B, while B loses $\frac{1}{2}$ of its share to A. Find the relative proportion of the market that each hold after 2 years.

Ans. A has 43.3 percent and B has 56.7 percent.

5.73. Consider a Markov chain with state $\{0, 1, 2\}$ and transition probability matrix

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}$$

Is state 0 periodic?

Hint: Draw the state transition diagram.

Ans. No.

5.74. Verify Eq. (5.51).

Hint: Let $\tilde{N} = [N_{jk}]$, where N_{jk} is the number of times the state $k(\in B)$ is occupied until absorption takes place when $X(n)$ starts in state $j(\in B)$. Then $T_j = \sum_{k=m+1}^{N} N_{jk}$; calculate $E(N_{jk})$.

5.75. Consider a Markov chain with transition probability matrix

$$P = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.4 & 0.5 & 0.1 \\ 0.6 & 0 & 0.4 \end{bmatrix}$$

Find the steady-state probabilities.

Ans. $\hat{\mathbf{p}} = [\frac{5}{9} \quad \frac{2}{9} \quad \frac{2}{9}]$

5.76. Let $X(t)$ be a Poisson process with rate λ. Find $E[X^2(t)]$.

Ans. $\lambda t + \lambda^2 t^2$

5.77. Let $X(t)$ be a Poisson process with rate λ. Find $E\{[X(t) - X(s)]^2\}$ for $t > s$.

Hint: Use the independent stationary increments condition and the result of Prob. 5.76.

Ans. $\lambda(t - s) + \lambda^2(t - s)^2$

5.78. Let $X(t)$ be a Poisson process with rate λ. Find

$$P[X(t - d) = k \mid X(t) = j] \qquad d > 0$$

Ans. $\dfrac{j!}{k!(j-k)!}\left(\dfrac{t-d}{t}\right)^k\left(\dfrac{d}{t}\right)^{j-k}$

5.79. Let T_n denote the time of the nth event of a Poisson process with rate λ. Find the variance of T_n.

Ans. n/λ^2

5.80. Assume that customers arrive at a bank in accordance with a Poisson process with rate $\lambda = 6$ per hour, and suppose that each customer is a man with probability $\frac{2}{3}$ and a woman with probability $\frac{1}{3}$. Now suppose that 10 men arrived in the first 2 hours. How many woman would you expect to have arrived in the first 2 hours?

Ans. 4

5.81. Let $X_1, \ldots, X_n$ be jointly normal r.v.'s. Let

$$Y_i = X_i + c_i \qquad i = 1, \ldots, n$$

where c_i are constants. Show that $Y_1, \ldots, Y_n$ are also jointly normal r.v.'s.

Hint: See Prob. 5.60.

5.82. Derive Eq. (*5.63*).

Hint: Use condition (1) of a Wiener process and Eq. (*5.102*) of Prob. 5.22.