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

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

$F_X(x_1; t_1)$ 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

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

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

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

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

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

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

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 $n \leq k$, 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 $\{X(t), t \geq 0\}$ 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 $\{X_n, n \geq 0\}$ 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 $\{X_n, n \geq 0\}$ is characterized by [Eq. (5.27)]

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

where $P\{x_{n+1} = j \,|\, X_n = i\}$ are known as one-step transition probabilities. If $P\{x_{n+1} = j \,|\, X_n = i\}$ 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 $\{X_n, n \geq 0\}$ be a homogeneous Markov chain with a discrete infinite state space $E = \{0, 1, 2, \ldots\}$. Then

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

regardless of the value of n. A *transition probability matrix of* $\{X_n, n \geq 0\}$ 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

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

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

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

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

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

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).

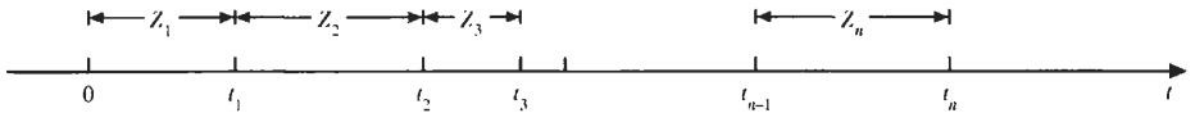

Fig. 5-1

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 $h > 0$.

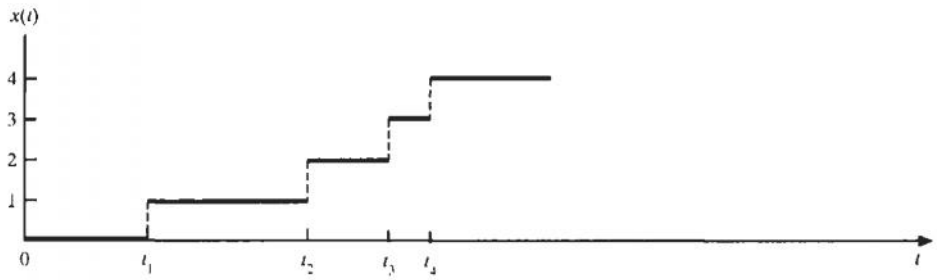

Fig. 5-2 A sample function of a counting process.

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

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

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

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 $T = \{1, 2, \ldots\}$.