Part One

BASIC CONCEPTS AND TOOLS

COPYRIGHTED MATERIAL

1

Stochastic processes

1.1 Introduction

The theme of this book is Bayesian Analysis of Stochastic Process Models. In this first chapter, we shall provide the basic concepts needed in defining and analyzing stochastic processes. In particular, we shall review what stochastic processes are, their most important characteristics, the important classes of processes that shall be analyzed in later chapters, and the main inference and decision-making tasks that we shall be facing. We also set up the basic notation that will be followed in the rest of the book. This treatment is necessarily brief, as we cover material which is well known from, for example, the texts that we provide in our final discussion.

1.2 Key concepts in stochastic processes

Stochastic processes model systems that evolve randomly in time, space or space-time. This evolution will be described through an index $t \in T$. Consider a random experiment with sample space Ω, endowed with a σ-algebra $\mathcal{F}$ and a base probability measure P. Associating numerical values with the elements of that space, we may define a family of random variables $\{X_t, t \in T\}$, which will be a stochastic process. This idea is formalized in our first definition that covers our object of interest in this book.

Definition 1.1: *A stochastic process $\{X_t, t \in T\}$ is a collection of random variables X_t, indexed by a set T, taking values in a common measurable space S endowed with an appropriate σ-algebra.*

T could be a set of times, when we have a temporal stochastic process; a set of spatial coordinates, when we have a spatial process; or a set of both time and spatial coordinates, when we deal with a spatio-temporal process. In this book, in general,

Bayesian Analysis of Stochastic Process Models, First Edition. David Rios Insua, Fabrizio Ruggeri and Michael P. Wiper.
© 2012 John Wiley & Sons, Ltd. Published 2012 by John Wiley & Sons, Ltd.

P1: TIX/XYZ P2: ABC
JWST172-c01 JWST172-Ruggeri March 3, 2012 15:53

we shall focus on stochastic processes indexed by time, and will call T the *space of times*. When T is discrete, we shall say that the process is *in discrete time* and will denote time through n and represent the process through $\{X_n,\ n = 0, 1, 2, \ldots\}$. When T is continuous, we shall say that the process is in *continuous time*. We shall usually assume that $T = [0, \infty)$ in this case. The values adopted by the process will be called the *states* of the process and will belong to the *state space* S. Again, S may be either discrete or continuous.

At least two visions of a stochastic process can be given. First, for each $\omega \in \Omega$, we may rewrite $X_t(\omega) = g_\omega(t)$ and we have a function of t which is a realization or a sample function of the stochastic process and describes a possible evolution of the process through time. Second, for any given t, X_t is a random variable. To completely describe the stochastic process, we need a joint description of the family of random variables $\{X_t, t \in T\}$, not just the individual random variables. To do this, we may provide a description based on the joint distribution of the random variables at any discrete subset of times, that is, for any $\{t_1, \ldots, t_n\}$ with $t_1 < \cdots < t_n$, and for any $\{x_1, \ldots, x_n\}$, we provide

$$P\left(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n\right).$$

Appropriate consistency conditions over these finite-dimensional families of distributions will ensure the definition of the stochastic process, via the Kolmogorov extension theorem, as in, for example, Øksendal (2003).

Theorem 1.1: *Let $T \subseteq [0, \infty)$. Suppose that, for any $\{t_1, \ldots, t_n\}$ with $t_1 < \cdots < t_n$, the random variables $X_{t_1}, \ldots, X_{t_n}$ satisfy the following consistency conditions:*

1. *For all permutations π of $1, \ldots, n$ and $x_1, \ldots, x_n$ we have that $P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n) = P(X_{t_{\pi(1)}} \leq x_{\pi(1)}, \ldots, X_{t_{\pi(n)}} \leq x_{\pi(n)})$.*
2. *For all $x_1, \ldots, x_n$ and $t_{n+1}, \ldots, t_{n+m}$, we have $P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n) = P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n, X_{t_{n+1}} < \infty, \ldots, X_{t_{n+m}} < \infty)$.*

Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and a stochastic process $X_t : T \times \Omega \to \mathbb{R}^n$ having the families $X_{t_1}, \ldots, X_{t_n}$ as finite-dimensional distributions.

Clearly, the simplest case will hold when these random variables are independent, but this is the territory of standard inference and decision analysis. Stochastic processes adopt their special characteristics when these variables are dependent.

Much as with moments for standard distributions, we shall use some tools to summarize a stochastic process. The most relevant are, assuming all the involved moments exist:

Definition 1.2: *For a given stochastic process $\{X_t, t \in T\}$ the mean function is*

$$\mu_X(t) = E[X_t].$$

The autocorrelation function of the process is the function

$$R_X(t_1, t_2) = E[X_{t_1} X_{t_2}].$$

Finally, the autocovariance function of the process is

$$C_X\,(t_1, t_2) = E[(X_{t_1} - \mu_X(t_1))(X_{t_2} - \mu_X(t_2))].$$

It should be noted that these moments are merely summaries of the stochastic process and do not characterize it, in general.

An important concept is that of a stationary process, that is a process whose characterization is independent of the time at which the observation of the process is initiated.

Definition 1.3: *We say that the stochastic process* $\{X_t, t \in T\}$ *is strictly stationary if for any* n, $t_1, t_2, \ldots, t_n$ *and* τ, $(X_{t_1}, \ldots, X_{t_n})$ *has the same distribution as* $(X_{t_1+\tau}, \ldots, X_{t_n+\tau})$.

A process which does not satisfy the conditions of Definition 1.3 will be called nonstationary. Stationarity is a typical feature of a system which has been running for a long time and has stabilized its behavior.

The required condition of equal joint distributions in Definition 1.3 has important parameterization implications when $n = 1, 2$. In the first case, we have that all X_t variables have the same common distribution, independent of time. In the second case, we have that the joint distribution depends on the time differences between the chosen times, but not on the particular times chosen, that is,

$$F_{X_{t_1}, X_{t_2}}(x_1, x_2) = F_{X_0, X_{t_2 - t_1}}(x_1, x_2).$$

Therefore, we easily see the following.

Proposition 1.1: *For a strictly stationary stochastic process* $\{X_t, t \in T\}$, *the mean function is constant, that is,*

$$\mu_X(t) = \mu_X, \forall t. \tag{1.1}$$

Also, the autocorrelation function of the process is a function of the time differences, that is,

$$R_X(t_1, t_2) = R(t_2 - t_1). \tag{1.2}$$

Finally, the autocovariance function is given by

$$C_X(t_1, t_2) = R(t_2 - t_1) - \mu_X^2,$$

assuming all relevant moments exist.

A process that fulfills conditions (1.1) and (1.2) is commonly known as a weakly stationary process. Such a process is not necessarily strictly stationary, whereas a strictly stationary process will be weakly stationary if first and second moments exist.

Example 1.1: A first-order autoregressive, or AR(1), process is defined through

$$X_n = \phi_0 + \phi_1 X_{n-1} + \epsilon_n,$$

where ϵ_n is a sequence of independent and identically distributed (IID) normal random variables with zero mean and variance σ^2. This process is weakly, but not strictly, stationary if $|\phi_1| < 1$. Then, we have $\mu_X = \phi_0 + \phi_1 \mu_X$, which implies that $\mu_X = \frac{\phi_0}{1-\phi_1}$. If $|\phi_1| \geq 1$, the process is not stationary. △

When dealing with a stochastic process, we shall sometimes be interested in its transition behavior, that is, given some observations of the process, we aim at forecasting some of its properties a certain time t ahead in the future. To do this, it is important to provide the so called *transition functions*. These are the conditional probability distributions based on the available information about the process, relative to a specific value of the parameter t_0.

Definition 1.4: *Let $t_0, t_1 \in T$ be such that $t_0 \leq t_1$. The* conditional transition distribution function *is defined by*

$$F\left(x_0, x_1; t_0, t_1\right) = P\left(X_{t_1} \leq x_1 \mid X_{t_0} \leq x_0\right).$$

When the process is discrete in time and space, we shall use the transition probabilities defined, for $m \leq n$, through

$$P_{ij}^{(m,n)} = P\left(X_n = j \mid X_m = i\right).$$

When the process is stationary, the transition distribution function will depend only on the time differences $t = t_1 - t_0$,

$$F\left(x_0, x; t_0, t_0 + t\right) = F\left(x_0, x; 0, t\right), \ \forall t_0 \in T.$$

For convenience, the previous expression will sometimes be written as $F\left(x_0, x; t\right)$. Analogously, for the discrete process $\{X_n\}_n$ we shall use the expression $P_{ij}^{(n)}$.

Letting $t \to \infty$, we may consider the long-term limiting behavior of the process, typically associated with the stationary distribution. When this distribution exists, computations are usually much simpler than doing short-term predictions based on the use of the transition functions. These limit distributions reflect a parallelism with the laws of large numbers, for the case of IID observations, in that

$$\frac{1}{n}\sum_{i=1}^{n} X_{t_i} \to E[X_\infty]$$

when $t_n \to \infty$, for some limiting random variable X_∞. This is the terrain of ergodic theorems and ergodic processes, see, e.g., Walters (2000).

In particular, for a given stochastic process, we may be interested in studying the so-called time averages. For example, we may define the mean time average, which is the random variable defined by

$$\mu_X(T) = \frac{1}{T}\int_0^T X_t \, \mathrm{d}t.$$

If the process is stationary, interchanging expectation with integration, we have

$$E[\mu_X(T)] = \frac{1}{T}E\left[\int_0^T X_t \mathrm{d}t\right] = \frac{1}{T}\int_0^T E[X_t]\mathrm{d}t = \frac{1}{T}\int_0^T \mu_X = \mu_X.$$

This motivates the following definition.

Definition 1.5: *The process X_t is said to be mean ergodic if:*

1. $\mu_X(T) \to \mu_X$*, for some* μ_X*, and*
2. $var(\mu_X(T)) \to 0$.

An autocovariance ergodic process can be defined in a similar way. Clearly, for a stochastic process to be ergodic, it has to be stationary. The converse is not true.

1.3 Main classes of stochastic processes

Here, we define the main types of stochastic processes that we shall study in this book. We start with Markov chains and Markov processes, which will serve as a model for many of the other processes analyzed in later chapters and are studied in detail in Chapters 3 and 4.

1.3.1 Markovian processes

Except for the case of independence, the simplest dependence form among the random variables in a stochastic process is the Markovian one.

Definition 1.6: *Consider a set of time instants $\{t_0, t_1, \ldots, t_n, t\}$ with $t_0 < t_1 < \cdots < t_n < t$ and $t, t_i \in T$. A stochastic process $\{X_t, t \in T\}$ is Markovian if the distribution* *of X_t conditional on the values of $X_{t_1}, \ldots, X_{t_n}$ depends only on X_{t_n}, that is, the most recent known value of the process*

$$\begin{aligned} &P\left(X_t \leq x \mid X_{t_n} \leq x_n, X_{t_{n-1}} \leq x_{n-1}, \ldots, X_{t_0} \leq x_0\right) \\ &\quad = P\left(X_t \leq x \mid X_{t_n} \leq x_n\right) = F\left(x_n, x; t_n, t\right). \end{aligned} \tag{1.3}$$

As a consequence of the previous relation, we have

$$F\left(x_0, x; t_0, t_0 + t\right) = \int_{y \in S} F\left(y, x; \tau, t\right) \mathrm{d}F\left(x_0, y; t_0, \tau\right) \tag{1.4}$$

with $t_0 < \tau < t$.

If the stochastic process is discrete in both time and space, then (1.3) and (1.4) adopt the following form: For $n > n_1 > \cdots > n_k$, we have

$$\begin{aligned} &P\left(X_n = j \mid X_{n_1} = i_1, X_{n_2} = i_2, \ldots, X_{n_k} = i_{n_k}\right) = \\ &\qquad\qquad P\left(X_n = j \mid X_{n_1} = i_1\right) = p_{i_1 j}^{(n_1, n)}. \end{aligned}$$

Using this property and taking r such that $m < r < n$, we have

$$\begin{aligned} p_{ij}^{(m,n)} &= P\left(X_n = j \mid X_m = i\right) \\ &= \sum_{k \in S} P\left(X_n = j \mid X_r = k\right) P\left(X_r = k \mid X_m = i\right). \end{aligned} \tag{1.5}$$

Equations (1.4) and (1.5) are called the *Chapman–Kolmogorov equations* for the continuous and discrete cases, respectively. In this book we shall refer to discrete state space Markov processes as Markov chains and will use the term Markov process to refer to processes with continuous state spaces and the Markovian property.

Discrete time Markov chains

Markov chains with discrete time space are an important class of stochastic processes whose analysis serves as a guide to the study of other more complex processes. The main features of such chains are outlined in the following text. Their full analysis is provided in Chapter 3.

Consider a discrete state space Markov chain, $\{X_n\}$. Let $p_{ij}^{(m,n)}$ be defined as in (1.5), being the probability that the process is at time n in j, when it was in i at time m. If $n = m + 1$, we have

$$p_{ij}^{(m,m+1)} = P\left(X_{m+1} = j \mid X_m = i\right),$$

which is known as the one-step *transition probability*. When $p_{ij}^{(m,m+1)}$ is independent of m, the process is stationary and the chain is called *time homogeneous*. Otherwise, the process is called time inhomogeneous. Using the notation

$$p_{ij} = P\,(X_{m+1} = j \mid X_m = i)$$
$$p_{ij}^n = P\,(X_{n+m} = j \mid X_m = i)$$

for every m, the Chapman–Kolmogorov equations are now

$$p_{ij}^{n+m} = \sum_{k \in S} p_{ik}^n p_{kj}^m \tag{1.6}$$

for every $n, m \geq 0$ and i, j. The *n-step transition probability matrix* is defined as $\mathbf{P}^{(n)}$, with elements p_{ij}^n. Equation (1.6) is written $\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)} \cdot \mathbf{P}^{(m)}$. These matrices fully characterize the transition behavior of an homogeneous Markov chain. When $n = 1$, we shall usually write $\mathbf{P}$ instead of $\mathbf{P}^{(1)}$ and shall refer to the *transition matrix* instead of the one-step transition matrix.

Example 1.2: A famous problem in stochastic processes is the gambler's ruin problem. A gambler with an initial stake, $x_0 \in \mathbb{N}$, plays a coin tossing game where at each turn, if the coin comes up heads, she wins a unit and if the coin comes up tails, she loses a unit. The gambler continues to play until she either is bankrupted or her current holdings reach some fixed amount m. Let X_n represent the amount of money held by the gambler after n steps. Assume that the coin tosses are IID with probability of heads p at each turn. Then, $\{X_n\}$ is a time homogeneous Markov chain with $p_{00} = p_{mm} = 1$, $p_{i\,i+1} = p$ and $p_{i\,i-1} = 1 - p$, for $i = 1, \ldots, m-1$ and $p_{ij} = 0$ for $i \in \{0, \ldots, m\}$ and $j \neq i$. △

The analysis of the stationary behavior of an homogeneous Markov chain requires studying the relations among states as follows.

Definition 1.7: *A state j is reachable from a state i if $p_{ij}^n > 0$, for some n. We say that two states that are mutually reachable, communicate, and belong to the same communication class.*

If all states in a chain communicate among themselves, so that there is just one communication class, we shall say that the Markov chain is irreducible. In the case of the gambler's ruin problem of Example 1.2, we can see that there are three communication classes: $\{0\}$, $\{1, \ldots, m-1\}$, and $\{m\}$.

Definition 1.8: *Given a state i, let p_i be the probability that, starting from state i, the process returns to such state. We say that state i is recurrent if $p_i = 1$ and transitory if $p_i < 1$.*

We may easily see that if state i is recurrent and communicates with another state j, then j is recurrent. In the case of gambler's ruin, only the states $\{0\}$ and $\{m\}$ are recurrent.

P1: TIX/XYZ P2: ABC
JWST172-c01 JWST172-Ruggeri March 3, 2012 15:53

Definition 1.9: *A state i has* period *k if* $p_{ii}^n = 0$ *whenever n is not divisible by k and k is the biggest integer with this property. A state with period one is aperiodic.*

We may also see easily that if i has period k and states i and j communicate, then state j has period k. In the gambler's ruin problem, states $\{0, m\}$ are aperiodic and the remaining states have period two.

Definition 1.10: *A state i is positive recurrent if, starting at i, the expected time until return to i is finite.*

Positive recurrence is also a class property in the sense that, if i is positively recurrent and states i and j communicate, then state j is also positively recurrent. We may also prove that in a Markov chain with a finite number of states all recurrent states are positive recurrent. The final key definition is the following.

Definition 1.11: *A positive recurrent, aperiodic state is called ergodic.*

We then have the following important limiting result for a Markov chain, whose proof may be seen in, for example, Ross (1995).

Theorem 1.2: *For an ergodic and irreducible Markov chain, then* $\pi_j = \lim_{n\to\infty} p_{ij}^n$, *which is independent of i* π_j *is the unique nonnegative solution of* $\pi_j = \sum_i \pi_i p_{ij}$, $j \geq 0$, *with* $\sum_{i=0}^{\infty} \pi_i = 1$.

Continuous time Markov chains

Here, we describe only the homogeneous case. Continuous time Markov chains are stochastic processes with discrete-state space and continuous space time such that whenever a system enters in state i, it remains there for an exponentially distributed time with mean $1/\lambda_i$, and when it abandons this state, it goes to state $j \neq i$ with probability p_{ij}, where $\sum_{j \neq i} p_{ij} = 1$.

The required transition and limited behavior of these processes and some generalizations are presented in Chapter 4.

1.3.2 Poisson process

Poisson processes are continuous time, discrete space processes that we shall analyze in detail in Chapter 5. Here, we shall distinguish between homogeneous and nonhomogeneous Poisson processes.

Definition 1.12: *Suppose that the stochastic process* $\{X_t\}_{t \in T}$ *describes the number of events of a certain type produced until time t and has the following properties:*

1. *The number of events in nonoverlapping intervals are independent.*
2. *There is a constant* λ *such that the probabilities of occurrence of events over 'small' intervals of duration* $\triangle t$ *are:*
   - $P(\textit{number of events in } (t, t+\triangle t] = 1) = \lambda\triangle t + o(\triangle t)$.
   - $P(\textit{number of events in } (t, t+\triangle t] > 1) = o(\triangle t)$, *where* $o(\triangle t)$ *is such that* $o(\triangle t)/\triangle t \to 0$ *when* $\triangle t \to 0$.

Then, we say that $\{X_t\}$ *is an homogeneous Poisson process with parameter* λ, *characterized by the fact that* $X_t \sim Po(\lambda t)$.

For such a process, it can be proved that the times between successive events are IID random variables with distribution $\text{Ex}(\lambda)$.

The Poisson process is a particular case of many important generic types of processes. Among others, it is an example of a renewal process, that is, a process describing the number of events of a phenomenon of interest occurring until a certain time such that the times between events are IID random variables (exponential in the case of the Poisson process). Poisson processes are also a special case of continuous time Markov chains, with transition probabilities $p_{i,i+1} = 1, \forall i$ and $\lambda_i = \lambda$.

Nonhomogeneous Poisson processes

Nonhomogeneous Poisson processes are characterized by the intensity function $\lambda(t)$ or the mean function $m(t) = \int_0^t \lambda(s)\text{d}s$; we consider, in general, a time-dependent intensity function but it could be space and space-time dependent as well. Note that, when $\lambda(t) = \lambda$, we have an homogeneous Poisson process. For a nonhomogeneous Poisson process, the number of events occurring in the interval $(t, t+s]$ will have a $\text{Po}(m(t+s) - m(t))$ distribution.

1.3.3 Gaussian processes

The Gaussian process is continuous in both time and state spaces. Let $\{X_t\}$ be a stochastic process such that for any n times $\{t_1, t_2, \ldots, t_n\}$ the joint distribution of $X_{t_i}, i = 1, 2, .., n$, is n-variate normal. Then, the process is *Gaussian*. Moreover, if for any finite set of time instants $\{t_i\},\ i = 1, 2, \ldots$ the random variables are mutually independent and X_t is normally distributed for every t, we call it a *purely random Gaussian process*.

Because of the specific properties of the normal distribution, we may easily specify many properties of a Gaussian process. For example, if a Gaussian process is weakly stationary, then it is strictly stationary.

1.3.4 Brownian motion

This continuous time and state-space process has the following properties:

1. The process $\{X_t, t \geq 0\}$ has independent, stationary increments: for $t_1, t_2 \in T$ and $t_1 < t_2$, the distribution of $X_{t_2} - X_{t_1}$ is the same of $X_{t_2+h} - X_{t_1+h}$ for every $h > 0$

and, for nonoverlapping intervals (t_1, t_2) and (t_3, t_4), with $t_1 < t_2 < t_3 < t_4$, the random variables $X_{t_2} - X_{t_1}$ and $X_{t_4} - X_{t_3}$ are independent.
2. For any time interval (t_1, t_2), the random variable $X_{t_2} - X_{t_1}$ has distribution $\mathrm{N}(0, \sigma^2(t_2 - t_1))$.

1.3.5 Diffusion processes

Diffusion processes are Markov processes with certain continuous path properties which emerge as solution of stochastic differential equations. Specifically,

Definition 1.13: *A continuous time and state process is a diffusion process if it is a Markov process $\{X_t\}$ with transition density $p(s, t; x, y)$ such that there are two functions $\mu(t, x)$ and $\beta^2(t, x)$, known as the drift and the diffusion coefficients, such that*

$$\int_{|x-y|\le\epsilon} p(t, t+\Delta t; x, y)\mathrm{d}y = \mathrm{o}(\Delta \mathrm{t}),$$
$$\int_{|x-y|\le\epsilon} (y-x)p(t, t+\Delta t; x, y)\mathrm{d}y = \mu(\mathrm{t}, \mathrm{x}) + \mathrm{o}(\Delta \mathrm{t}),$$
$$\int_{|x-y|\le\epsilon} (y-x)^2 p(t, t+\Delta t; x, y)\mathrm{d}y = \beta^2(\mathrm{t}, \mathrm{x}) + \mathrm{o}(\Delta \mathrm{t}).$$

The previous three types of processes are dealt with in Chapter 6.

1.4 Inference, prediction, and decision-making

Given the key definitions and results concerning stochastic processes, we can now informally set up the statistical and decision-making problems that we shall deal with in the following chapters.

Clearly, stochastic processes will be characterized by their initial value and the values of their parameters, which may be finite or infinite dimensional.

Example 1.3: In the case of the gambler's ruin problem of Example 1.2 the process is parameterized by p, the probability of heads. More generally, for a stationary finite Markov chain model with states $1, 2, \ldots, k$, the parameters will be the transition probabilities $(p_{11}, \ldots, p_{k,k})$, where p_{ij} satisfy that $p_{ij} \ge 0$ and $\sum_j p_{ij} = 1$.

The AR(1) process of Example 1.1 is parameterized through the parameters ϕ_0 and ϕ_1.

A nonhomogeneous Poisson process with intensity function $\lambda(t) = M\beta t^{\beta-1}$, corresponding to a Power Law model, is a finite parametric model with parameters (M, β).

A normal dynamic linear model (DLM) with univariate observations X_n, is described by

$$\begin{aligned} \theta_0 | D_0 &\sim \mathrm{N}\,(m_0, C_0) \\ \theta_n | \theta_{n-1} &\sim \mathrm{N}\,(\boldsymbol{G}_n \theta_{n-1}, \boldsymbol{W}_n) \\ X_n | \theta_n &\sim \mathrm{N}\,(F_n' \theta_n, V_n) \end{aligned}$$

where, for each n, F_n is a known vector of dimension $m \times 1$, $\boldsymbol{G}_n$ is a known $m \times m$ matrix, V_n is a known variance, and $\boldsymbol{W}_n$ is a known $m \times m$ variance matrix. The parameters are now $\{\theta_0, \theta_1, \ldots\}$. △

Inference problems for stochastic processes are stated as follows. Assume we have observations of the stochastic process, which will typically be observations $X_{t_1}, \ldots, X_{t_n}$ at time points $t_1, \ldots, t_n$. Sometimes we could have continuous observations in terms of one, or more, trajectories within a given interval. Our aim in inference is then to summarize the available information about these parameters so as to provide point or set estimates or test hypotheses about them. It is important to emphasize that this available information comes from both the observed data and any available prior information.

More important in the context of stochastic processes is the task of forecasting the future behavior of the process, in both the transitory and limiting cases, that is, at a fixed future time and in the long term, respectively.

We shall also be interested in several decision-making problems in relation with stochastic processes. Typically, they will imply making a decision from a set of available ones, once we have taken the process observations. A reward will be obtained depending on the decision made and the future behavior of the process. We aim at obtaining the optimal solution in some sense.

This book explores how the problems of inference, forecasting, and decision-making with underlying stochastic processes may be dealt with using Bayesian techniques. In the following chapter, we review the most important features of the Bayesian approach, concentrating on the standard IID paradigm while in the later chapters, we concentrate on the analysis of some of the specific stochastic processes outlined earlier in Section 1.3.

1.5 Discussion

In this chapter, we have provided the key results and definitions for stochastic processes that will be needed in the rest of this book. Most of these results are of a probabilistic nature, as is usual in the majority of books in this field. Many texts provide very complete outlines of the probabilistic aspects of stochastic processes. For examples, see Karlin and Taylor (1975, 1981), Ross (1995), and Lawler (2006), to name a few.

There are also texts focusing on some of the specific processes that we have mentioned. For example, Norris (1998) or Ching and Ng (2010) are full-length books on Markov chains; Stroock (2005) deals with Markov processes; Poisson processes are studied in Kingman (1993); Rasmussen and Williams (2005) study Gaussian processes, whereas diffusions are studied by Rogers and Williams (2000a, 2000b).

As we have observed previously, there is less literature dedicated to inference for stochastic processes. A quick introduction may be seen in Lehoczky (1990) and both Bosq and Nguyen (1996), and Bhat and Miller (2002) provide applied approaches very much in the spirit of this book, although from a frequentist point of view. Prabhu and Basawa (1991), Prakasa Rao (1996), and Rao (2000) are much more theoretical.

Finally, we noted earlier that the index T of a stochastic process need not always be a set of times. Rue and Held (2005) illustrate the case of spatial processes, when T is a spatial set.

References

Bhat, U.N. and Miller, G. (2002). *Elements of Applied Stochastic Processes* (3rd edn.). New York: John Wiley & Sons, Inc.

Bosq, D. and Nguyen, H.T. (1996) *A Course in Stochastic Processes: Stochastic Models and Statistical Inference*. Dordrecht: Kluwer.

Ching, W. and Ng, N.K. (2010) *Markov Chains: Models, Algorithms and Applications*. Berlin: Springer.

Karlin, S. and Taylor, H.M. (1975) *A First Course in Stochastic Processes* (2nd edn.). New York: Academic Press.

Karlin, S. and Taylor, H.M. (1981) *A Second Course in Stochastic Processes*. New York: Academic Press.

Kingman, J.F.C. (1993) *Poisson Processes*. Oxford: Oxford University Press.

Lawler, G.F. (2006) *Introduction to Stochastic Processes* (2nd edn.). New York: Chapman and Hall.

Lehoczky, J. (1990) Statistical methods. In *Stochastic Models*, D.P. Heyman and M.J. Sobel (Eds.). Amsterdam: North-Holland.

Norris, J.R. (1998) *Markov Chains*. Cambridge: Cambridge University Press.

Øksendal, B. (2003) *Stochastic Differential Equations: An Introduction with Applications*. Berlin: Springer.

Prabhu, N.U. and Basawa, I.V. (1991) *Statistical Inference in Stochastic Processes*. New York: Marcel Dekker.

Prakasa Rao, B.L.S. (1996) *Stochastic Processes and Statistical Inference*. New Delhi: New Age International.

Rao, M.M. (2000) *Stochastic Processes: Inference Theory*. Dordrecht: Kluwer.

Rasmussen, C.E. and Williams, C.K.I. (2005) *Gaussian Processes for Machine Learning*. Cambridge, MA: The MIT Press.

Rogers, L.C.G. and Williams, D. (2000a) *Diffusions, Markov Processes and Martingales: Volume 1 Foundations*. Cambridge: Cambridge University Press.

Rogers, L.C.G. and Williams, D. (2000b) *Diffusions, Markov Processes and Martingales: Volume 2 Ito Calculus*. Cambridge: Cambridge University Press.

Ross, S. (1995) *Stochastic Processes*. New York: John Wiley & Sons, Inc.

Rue, H. and Held, L. (2005) *Gaussian Markov Random Fields: Theory and Applications*. Boca Raton: Chapman and Hall.

Stroock, D.W. (2005) *An Introduction to Markov Processes*. Berlin: Springer.

Walters, P. (2000) *Introduction to Ergodic Theory*. Berlin: Springer.