Markov processes will provide us with a large class of processes that can be used as building blocks for useful and interesting random fields. This and the following chapter are a brief introduction to Markov processes.

Although in some specialized settings we have already encountered the Markov property, e.g., Chapters 3 and 7, the general theory of Markov processes is substantially more complicated; thus, we restrict attention to a nice class of Markov processes that are known as Feller processes. In order to gel the basic ideas, we begin with the simplest case, which is that of discrete Markov chains. Our treatment of the continuous-time theory will follow, starting with Section 2.

1 Discrete Markov Chains

We start our development of the theory of Markov processes in its simplest setting: discrete Markov chains. These are discrete-time, discrete-space processes that possess the Markov property. Once some of the key concepts and methods are isolated, we proceed with our development of the more complicated continuous-time theory.

1.1 Preliminaries

If $S$ denotes a denumerable set, we say that a stochastic process $X = (X_n; \ n \geq 0)$ is a (discrete) Markov chain with state space $S$ if there exists a filtration $\mathcal{F} = (\mathcal{F}_n; \ n \geq 0)$ such that:
1. $X$ is an $S$-valued process that is adapted to $\mathcal{F}$;
2. for all $n \geq 0$ and all $a \in S$,
   \[ P(X_{n+1} = a \mid \mathcal{F}_n) = P(X_{n+1} = a \mid X_n), \text{ a.s.}^1 \]

Property 2 above is called the **Markov property** of $X$.

The following is an important exercise.

**Exercise 1.1.1** If $\mathcal{H} = (\mathcal{H}_n; n \geq 0)$ denotes the history of a Markov chain $X$, then for all $a \in S$ and all $n \geq 0$, $P(X_{n+1} = a \mid \mathcal{H}_n) = P(X_{n+1} = a \mid X_n)$, a.s. \hfill \Box

Thus, as a consequence of Exercise 1.1.1, unless a specific filtration is mentioned, the underlying filtration is tacitly assumed to be the history of $X$.

For all $n, k \geq 0$ and $a, b \in S$, define,
\[
p_{n,n+k}(a, b) = \begin{cases} 
P(X_{n+k} = b \mid X_n = a), & \text{if } P(X_n = a) > 0, \\
0, & \text{otherwise.} \end{cases}
\]

In words, $p_{n,n+k}(a, b)$ denotes the probability that our chain goes from $a$ to $b$, starting at time $n$, and ending at time $n + k$.

Recall that on $(X_n = a)$, $p_{n,n+1}(a, b) = P(X_{n+1} = b \mid X_n)$. Thus, the Markov property can be restated as follows: For all $n \geq 0$ and all $a \in S$,
\[
p_{n,n+1}(a, b) = P(X_{n+1} = b \mid \mathcal{F}_n),
\]
on $(X_n = a)$. Equivalently, we have the Markov property if and only if
\[
p_{n,n+1}(X_n, b) = P(X_{n+1} = b \mid \mathcal{F}_n), \text{ a.s.}
\]

We say that $X$ is a **time-homogeneous Markov chain** if for all $n, k \geq 0$, $p_{n,n+k} = p_{0,k}$. One can usually reduce attention to time-homogeneous Markov chains, as the following shows.

**Lemma 1.1.1** Let $X$ be an $S$-valued Markov chain. Define $Y = (Y_n; n \geq 0)$ by
\[
Y_n = (n, X_n), \quad n \geq 0.
\]

Then $Y$ is an $\mathbb{N}_0 \times S$-valued, time-homogeneous Markov chain.

**Proof** Let $\mathcal{F}$ denote the history of $X$ and note that $\mathcal{F}$ is also the history of $Y$. Clearly,
\[
P(Y_{n+k} = y \mid \mathcal{F}_n) = P(Y_{n+k} = y \mid Y_n), \quad \text{a.s.}
\]

---

1. Recall that for any random variable $Y$ and event $E$, $P(E \mid Y)$ is shorthand for $P(E \mid \sigma(Y))$, where $\sigma(Y)$ is the $\sigma$-field generated by $Y$. A similar remark holds for conditional expectations.