We begin this course with the theory of Markov chains. Let \((S, \mathcal{S})\), be any measurable space. Usually \(S\) is a finite set, a countable set, or \(\mathbb{R}^n\). For the most part we will confine our attention to discrete time processes. In the continuous time, continuous state space setting, Markov chains are known as Markov Processes.

### 1.1 Markov Property and Existence

**Definition 1.1** A sequence of random variables \((X_n)\) is called a Markov chain if the past and future of the process are conditionally independent given the present.

**Example 1.1** A random walk is an example of a Markov chain.

**Definition 1.2** A function \(p : S \times S \to \mathbb{R}\) is called a Markov kernel if

1) For each \(x \in S\), the mapping \(A \mapsto p(x, A)\) is a probability distribution on \((S, \mathcal{S})\).

2) For each \(A \in \mathcal{S}\), the mapping \(x \mapsto p(x, A)\) is an \(\mathcal{S}\)-measurable function.

**Definition 1.3** The Markov kernel \(p_n\) is called a transition probability function, or t.p.f, for \((X_n)\) if

\[
P(X_{n+1} \in B | \mathcal{F}_n) = p_n(X_n, B)
\]

for each \(B \in \mathcal{S}\).

In other words, \(p_n(x, B)\) is the probability that the next step in the chain lies in \(B\) given that the current state is \(x\). In the absence of the subscript \(n\) we call \(p\) a homogeneous transition probability function. Henceforth, in these notes we assume that \((X_n)\) has the homogeneous t.p.f. \(p\).

**Theorem 1.1** (Ionescu-Tulcea) Given a measurable space \((S, \mathcal{S})\) with distribution \(\mu\), and a transition probability function \(p\), there exists a Markov chain on the space and its distribution, \(P_\mu\), is unique on \((S \times S \times \ldots, \mathcal{S}^\infty)\). Here \(\mathcal{S}^\infty\) is the product \(\sigma\)-field.

It often convenient to suppose that \((X_n)\) is a coordinate process. That is, for \(w \in \Omega = \{(w_0, w_1, \ldots) : w_i \in S\}\) we set

\[
X_n(w) = w_n
\]

**Proof Sketch:** Under regularity assumptions on \(S\) this is a consequence of Kolmogorov’s Extension Theorem.
Notice that if we define,
\[
P_{\mu}(X_0 \in A_0) = \mu(A_0)
\]
\[
P_{\mu}(X_0 \in A_0, X_1 \in A_1) = \int_{A_0} \mu(dx_0)p(x_0, A_1)
\]
\[
P_{\mu}(X_0 \in A_0, X_1 \in A_1, X_2 \in A_2) = \int_{A_0} \mu(dx_0)\int_{A_1} p(x_0, dx_1)p(x_1, A_2)
\]
and so on, then we have a sequence of distributions on \(S, S \times S, \ldots\) that is consistent in the sense of Kolmogorov’s Extension Theorem. Measure theory then tells us that there exists a distribution on \(S \times S \times \ldots\) such that the first \(n\) coordinates are distributed as above on \(S^n\).

### 1.2 Some General Facts

To find the distribution of \(X_n\) we first regard the Markov kernel, \(p(\cdot, \cdot)\) as an operator on measures,
\[
\mu p(B) := \int \mu(dx)p(x, B)
\]
Thus \(\mu p\) is a new probability distribution. It is the distribution of \(X_1\) for a Markov chain, \((X_0, X_1, \ldots)\), with \(X_0 \sim \mu\) and t.p.f, \(p\). Similarly,
\[
\mu p^n(B) = \int \mu(dx)p^n(x, B) = \text{distribution of } X_n
\]
where \(p^n(x, B) = P^n_{\mu}(X_n \in B | X_0 = x)\).

When \(S\) is countable we typically denote the elements of \(S\) by \(i, j, k\), etc. In this case we define the transition matrix, \(P\), by
\[
P_{ij} = p(i, \{j\})
\]
the probability of transitioning from state \(i\) to state \(j\) given that the current position is \(i\). We can also identify the initial distribution, \(\mu\) with a row vector,
\[
\mu_i = \mu(\{i\})
\]
Clearly the matrix \(P\) must satisfy, \(\sum_j P_{ij} = 1\), for each \(i \in S\).

Applying this notation to the discussion at the begining of the section we conclude that if \((X_n)\) is a Markov chain with countable state space transition matrix \(P\) then
\[
P_{\mu}(X_n = j | X_0 = i) = P^n_{ij}
\]
and if \(X_0 \sim \mu\), and \(P^n\) denotes the \(n\)th matrix power of \(P\), then
\[
\mu P^n = \text{distribution of } X_n.
\]

On a general state space \((S, S)\) and for a suitable \(f : S \rightarrow \mathbb{R}\), say bounded measurable or non-negative measurable, define
\[
p f(x) := \int_S f(y)p(x, dy)
\]
Claim 1.2

\[ pf(x_n) = E_\mu[f(X_{n+1})|X_n = x_n] = E_\mu[f(X_{n+1})|X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0] \]

**Proof:** See Durrett, section 5.1. \( \square \)

Similarly we have that

\[ p^m f(x_n) = E_\mu[f(X_{n+m})|X_n = x_n] \]

In the case that \( S \) is countable the action of \( p \) on \( f \) can again be interpreted as a matrix vector operation, \( Pf \).

Now consider those functions \( h \) such that \( ph = h \). These functions are called *harmonic functions* because of a close relationship with the harmonic functions of Analysis. Applying the result of the claim, if \( h \) is harmonic then

\[ E_\mu[h(X_{n+1})|X_n = x_n] = ph(x_n) = h(x_n) \]

Thus for any initial distribution \( \mu \), \((h(X_n))\) is an \( \mathcal{F}_n \)-martingale.

**Example 1.2** Let \( B \) be the set of all *absorbing states*, meaning that \( p(b, b) = 1 \) for all \( b \in B \). Call \( B \) the *boundary* of the state space \( S \). Let \( A \) be some subset of \( B \) and define

\[ h_A(x) = P_\mu(X_n \in A \text{ eventually}). \]

Then (see Durrett, Section 5.2, Exercise 2.6)

1) \( h_A \) is a \( p \)-harmonic function.

2) if \( P_\mu(X_n \in B \text{ eventually}) = 1 \) then \( h_A \) is the unique \( p \)-harmonic function whose boundary values are given by \( 1_A \), the indicator function of \( A \).