# 1. Notation

Define $\mathbb{N} = \{1, 2, \ldots\}$, the set of natural numbers. Define

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\}, \qquad \overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}, \qquad \overline{\mathbb{N}}_0 = \mathbb{N} \cup \{0\} \cup \{\infty\}$$

Let $\mathcal{X}$ be a set, equipped with the sigma-field $\mathcal{B}$. Define $\Omega = \mathcal{X}^{\mathbb{N}_0}$. The typical element of $\Omega$ is a sequence $\omega = (\omega_0, \omega_1, \omega_2, \ldots)$. Equip $\Omega$ with $\mathcal{F}_\infty$, the smallest sigma-field on $\Omega$ for which each of the maps $\omega \mapsto \omega_n$, for $n \in \mathbb{N}_0$, is $\mathcal{F}_\infty\backslash\mathcal{B}$-measurable. Write $\mathcal{F}_n$ for the smallest sigma-field on $\Omega$ for which all of the maps $\omega \mapsto \omega_i$ for $0 \le i \le n$ are $\mathcal{F}_n\backslash\mathcal{B}$-measurable.

For $0 \le n \le \infty$, write $\mathcal{M}^+(\mathcal{F}_n)$ for the set of all $\mathcal{F}_n$-measurable $[0, \infty]$-valued functions on $\Omega$. For $n < \infty$, a function belongs to $\mathcal{M}^+(\mathcal{F}_n)$ if and only if it can be represented as $f_n(\omega_0, \ldots, \omega_n)$ with $f_n$ a $\mathcal{B}^{n+1}$-measurable, nonegative function on $\mathcal{X}^{n+1}$.

A ***Markov kernel*** on $\mathcal{X}$ is a family of probability measures $\{P(x, \cdot) : x \in \mathcal{X}\}$ on $\mathcal{B}$ for which $x \mapsto P(x, B)$ is $\mathcal{B}$-measurable for each $B \in \mathcal{B}$. I will also write $P_x$ for $P(x, \cdot)$ and, for an integrable function $f$ on $\mathcal{X}$, define

$$Pf = P_x f = P_x^y f(y) = P(x, f) = \int f(y) P(x, dy).$$

When $\mathcal{X}$ is finite, one can also think of $P$ as a matrix of nonnegative numbers for which $\sum_y P(x, y) = 1$ for each $x$. The notation $Pf$ then agrees with the usual notation for the product of $P$ and a column vector $f$.

Let $\mu$ be a probability measure on $\mathcal{B}$. For each function $f(\omega) = f_n(\omega_0, \ldots, \omega_n)$ in $\mathcal{M}^+(\mathcal{F}_n)$ define

$$\begin{aligned}\mathbb{P}_\mu f &= \int \ldots \int \mu(d\omega_0) P(\omega_0, d\omega_1) P(\omega_1, d\omega_2) \ldots P(\omega_{n-1}, d\omega_n) f_n(\omega_0, \ldots, \omega_n) \\ &= \mu^{\omega_0} P_{\omega_0}^{\omega_1} P_{\omega_1}^{\omega_2} \ldots P_{\omega_{n-1}}^{\omega_n} f_n(\omega_0, \ldots, \omega_n)\end{aligned}$$

It can be shown (Pollard 2001, Section 4.8) that $\mathbb{P}_\mu$ has a unique extention to a probability measure on $\mathcal{F}_\infty$.

Write $\mathbb{P}_x$ for $\mathbb{P}_\mu$ when $\mu$ concentrates at the single point $x$

# 2. Markov chains

To begin with, suppose $\mathcal{X}$ is a countable set and $\mathcal{B}$ consists of all subsets of $\mathcal{X}$. Call $\{X_n : n \in \mathbb{N}_0\}$ a Markov chain with state space $\mathcal{X}$ and transition probabilities $P(x, y)$ if

$$\mathbb{P}\{X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\} = P(x_n, y)$$

for all all $n$, all $x_0, \ldots, x_n$, and all $y$. If $X_0$ has distribution $\mu$ then

$$\mathbb{P}\{X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\} = \mu\{x_0\} P(x_0, x_1) P(x_1, x_2) \ldots P(x_{n-1}, x_n)$$

More generally, for suitably integrable $f$,

$$\mathbb{P} f(X_0, X_1, \ldots, X_n) = \mu^{x_0} P_{x_0}^{x_1} \ldots P_{x_{n-1}}^{x_n} f(x_0, f_1, \ldots, x_n) = \mathbb{P}_\mu f.$$

Even more generally,

<1> $$\mathbb{P} f(X_0, X_1, \ldots) = \mathbb{P}_\mu f \qquad \text{at least for all } f \in \mathcal{M}^+(\mathcal{F}_\infty).$$

In other words, as a map from the underlying probability space into the sequence space $\Omega$, the infinite random vector $(X_0, X_1, \ldots)$ has distribution $\mathbb{P}_\mu$.

Property <1> also makes sense for a general state space $\mathcal{X}$. It can be used as the definition of a Markov chain with state space $\mathcal{X}$ and transition probabilities $P(x, y)$

In fact, if we are only interested in probabilities and expectations involving the $X_i$'s, we can work exclusively with $\mathbb{P}_\mu$. More formally, we could identify $X_i$ with the $i$th coordinate map, that is $X_i(\omega) = \omega_i$, so that the Markov chain is actually defined by a sequence of random variables on the sequence space $\Omega$. Under $\mathbb{P}_\mu$, the chain has initial distribution $\mu$; under $\mathbb{P}_x$, the chain starts from $x$.

## 3. The Markov property

Suppose $k < n$. Suppose $f$ and $g$ are nonnegative, measurable functions of finitely many arguments. Then

$$\begin{aligned}
&\mathbb{P}_\mu f(\omega_0, \omega_1, \ldots, \omega_k) g(\omega_k, \ldots, \omega_n) \\
&\quad = \mu^{\omega_0} P_{\omega_0}^{\omega_1} P_{\omega_1}^{\omega_2} \ldots P_{\omega_{k-1}}^{\omega_k} f(\omega_0, \ldots, \omega_k) \left( P_{\omega_k}^{\omega_{k+1}} \ldots P_{\omega_{n-1}}^{\omega_n} g(\omega_k, \ldots, \omega_n) \right).
\end{aligned}$$

The expression in parentheses at the end corresponds to an expectation for a chain started at $\omega_k$. Indeed, if we rename the dummy variables we get

$$\begin{aligned}
P_{\omega_k}^{\omega_{k+1}} &\ldots P_{\omega_{n-1}}^{\omega_n} g(\omega_k, \ldots, \omega_n) \\
&= P_{\omega_k}^{y_1} \ldots P_{y_{\ell-1}}^{y_\ell} g(\omega_k, y_1 \ldots, y_\ell) \qquad \text{where } \ell = n - k \\
&= \mathbb{P}_{\omega_k} g
\end{aligned}$$

Thus

$$\mathbb{P}_\mu f(\omega_0, \omega_1, \ldots, \omega_k) g(\omega_k, \ldots, \omega_n) = \mathbb{P}_\mu \left( f(\omega_0, \omega_1, \ldots, \omega_k) \mathbb{P}_{\omega_k} g \right)$$

A simple generating class argument extends the equality at least to all $g$ in $\mathcal{M}^+(\mathcal{F}_\infty)$. The resulting assertion is called the ***Markov property***.

## 4. Stopping times

A function $\tau : \Omega \to \overline{\mathbb{N}}_0$ is called a stopping time if $\{\tau \le k\} \in \mathcal{F}_k$ for each $k \in \overline{\mathbb{N}}_0$. Equivalently, $\{\tau = k\} \in \mathcal{F}_k$ for each $k \in \overline{\mathbb{N}}_0$.

The sigma-field corresponding to "information available at time $\tau$" is defined by

$$\mathcal{F}_\tau = \{F \in \mathcal{F}_\infty : F\{\tau \le k\} \in \mathcal{F}_k \text{ for all } k \in \overline{\mathbb{N}}_0\}.$$

It is not too difficult to show that $f \in \mathcal{M}^+(\mathcal{F}_\tau)$ if and only if

$$f(\omega) = \sum_{k \in \overline{\mathbb{N}}_0} f_k(\omega_0, \ldots, \omega_k)\{\tau = k\}$$

where $f_k$ is $\mathcal{B}^{k+1}$-measurable and nonnegative.

## 5. The strong Markov property

Suppose $f \in \mathcal{M}^+(\mathcal{F}_\tau)$ and $g \in \mathcal{M}^+(\mathcal{F}_\infty)$. Then

$$\begin{aligned}
\mathbb{P}_\mu f(\omega) g &\left( X_\tau, X_{\tau+1}, \ldots \right) \{\tau < \infty\} \\
&= \sum_{k \in \mathbb{N}_0} \mathbb{P}_\mu f(\omega)\{\tau = k\} g(\omega_k, \omega_{k+1}, \ldots) \\
&= \sum_{k \in \mathbb{N}_0} \mathbb{P}_\mu f(\omega)\{\tau = k\} \mathbb{P}_{\omega_k} g \qquad \text{by the Markov property} \\
&= \mathbb{P} \left( f(\omega)\{\tau < \infty\} \mathbb{P}_{X_\tau(\omega)} g \right).
\end{aligned}$$


9 January 2006 Stat 606, version: 9jan06 ©David Pollard

This assertion is called the ***strong Markov property***. It can also be expressed as

$$\mathbb{P}_\mu\left(g\left(X_\tau, X_{\tau+1}, \ldots\right) \mid \mathcal{F}_\tau\right) = \mathbb{P}_x g \qquad \text{on the set } \{\tau < \infty, X_\tau = x\}.$$

## 6. Exercises

Suppose $\mathcal{X}$ is a countable state space and $\alpha$ is some arbitrary, but fixed, state. Try to use the strong Markov property to establish the following assertions.

Remember that $\tau_y = \inf\{n \in \mathbb{N} : X_i = y\}$. Define $\eta_y = \sum_{n\in\mathbb{N}}\{X_n = y\}$, the number of times the chain visits state $y$.

Say that a state $\alpha$ is accessible from a state $x$, denoted by $x \curvearrowright \alpha$, if $\mathbb{P}_x\{\tau_\alpha < \infty\} > 0$.

[1] Show that the following three assertions are equivalent.

(i) $\mathbb{P}_\alpha\{\tau_\alpha < \infty\} = 1$

(ii) $\mathbb{P}_\alpha \eta_\alpha = \infty$

(iii) $\mathbb{P}_\alpha\{\eta_\alpha = \infty\} = 1$

REMARK. A state $\alpha$ that satisfies any (and hence all) of these three requirements is said to be ***recurrent***.

[2] Suppose $x \curvearrowright \alpha$ and $\mathbb{P}_x\{\tau_x < \infty\} = 1$, for two states $x, \alpha \in \mathcal{X}$. Show that $\mathbb{P}_x\{\tau_\alpha < \infty\} = \mathbb{P}_\alpha\{\tau_\alpha < \infty\} = 1$.

## 7. Notes

There are many texts that develop the standard theory for Markov chains on countable state spaces. Chung (1967) is a classic. I find Freedman (1983) very clear because many details are spelled out.

### REFERENCES

Chung, K. L. (1967), *Markov Chains with Stationary Transition Probabilities*, second edn, Springer-Verlag.

Freedman, D. (1983), *Markov Chains*, Springer-Verlag.

Pollard, D. (2001), *A User's Guide to Measure Theoretic Probability*, Cambridge University Press.