Introduction to Monte Carlo Methods
Lecture Notes

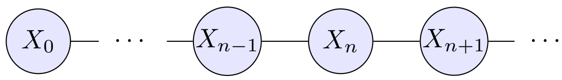
Chapter 3
Markov Chains

Qing Zhou[*]

Contents

[*]UCLA Department of Statistics (email: zhou@stat.ucla.edu).

1. Definitions and Examples



A discrete-time Markov chain (M.C.), $\{X_t : t = 0, 1, \cdots\}$, is a stochastic process with the Markov property:

$$P(X_{n+1} = j | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i)$$

for all time index n and all states $i_0, \cdots, i_{n-1}, i, j$.

State space is the range of possible values for the random variables X_t.

Assume the state space is finite: $\{0, 1, \cdots, N\}$ or countable: $\{0, 1, 2, \cdots\}$.

$\Rightarrow$ Discrete state discrete time Markov chain.

1.1. One-step transition probabilities

For a Markov chain, $P(X_{n+1} = j | X_n = i)$ is called a one-step transition probability. We assume that this probability does *not* depend on n, i.e.,

$$P(X_{n+1} = j | X_n = i) = p_{ij} \qquad \text{for } n = 0, 1, \ldots$$

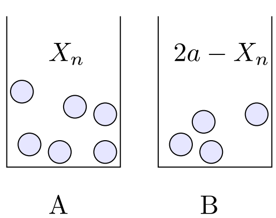
is the same for all time indices. In this case, $\{X_t\}$ is called a time homogeneous Markov chain.

Transition matrix: Put all transition probabilities (p_{ij}) into an $(N+1) \times (N+1)$ matrix,

$$\mathbb{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0N} \\ p_{10} & p_{11} & \cdots & p_{1N} \\ \vdots & \vdots & \vdots & \vdots \\ p_{N0} & p_{N1} & \cdots & p_{NN} \end{bmatrix} \qquad \begin{array}{ll} p_{ij} \geq 0, & \forall i, j. \\ \sum_{j=0}^{N} p_{ij} = 1, & \forall i. \end{array}$$

For countable space, replace N by ∞.

Example 1. The Ehrenfest urn model.



A total of $2a$ balls in two containers A and B. At each step, a ball is randomly selected from the $2a$ balls and moved to the other container. Let X_n be the # of balls in urn A at step n. ($X_n \in \{0, 1, \cdots, 2a\}$).

Then $\{X_0, X_1, \cdots\}$ is an M.C.

State space: $\{0, 1, 2, \cdots, 2a\}$

Transition probabilities:

$$P(X_{n+1} = i+1 | X_n = i) = P(\text{A ball is chosen from urn B} \mid X_n = i) = \frac{2a-i}{2a},$$
$$P(X_{n+1} = i-1 | X_n = i) = P(\text{A ball is chosen from urn A} \mid X_n = i) = \frac{i}{2a},$$
$$P(X_{n+1} = j | X_n = i) = 0, \qquad j \notin \{i+1, i-1\}.$$

$\Rightarrow \mathbb{P} = (p_{ij})_{(2a+1)\times(2a+1)}$:

$$p_{ij} = \begin{cases} \frac{2a-i}{2a}, & \text{if } j = i+1; \\ \frac{i}{2a}, & \text{if } j = i-1; \\ 0, & \text{otherwise.} \end{cases}$$

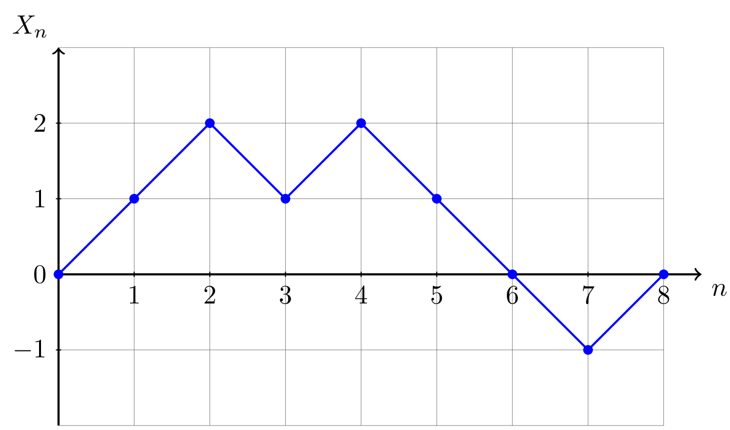
Example 2. 1D random walk.

Let $Y_1, Y_2, \cdots, Y_n, \cdots$ be iid: $P(Y_i = 1) = p$, $P(Y_i = -1) = q$, $p + q = 1$.

Define $\begin{cases} X_0 = 0, \\ X_n = X_{n-1} + Y_n, \quad \text{for } n = 1, 2, \cdots. \end{cases}$

$\Rightarrow X_n = Y_1 + Y_2 \cdots + Y_n = \sum_{i=1}^{n} Y_i$.

Then $\{X_n\}$ is an M.C.



Proof. To show that X_t is a Markov chain:

$$
\begin{aligned}
&P(X_{n+1} = j | X_1 = i_1, \cdots, X_{n-1} = i_{n-1}, X_n = i) \\
&= P(Y_{n+1} = j - i) \\
&= \begin{cases} p, & \text{if } j = i + 1; \\ q, & \text{if } j = i - 1; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

This shows that $P(X_{n+1} = j | X_1 = i_1, \cdots, X_{n-1} = i_{n-1}, X_n = i)$ only depends on $X_n = i$. □

1.2. Joint probability

The joint distribution of a Markov chain is completely determined by its one-step transition matrix and the initial probability distribution $P(X_0)$:

Theorem 1. *If $\{X_t : t = 0, 1, \cdots, n\}$ is an M.C., then*

$$P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = P(X_0 = i_0) \cdot p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}} p_{i_{n-1} i_n}.$$

Proof.

$$
\begin{aligned}
&P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) \\
&= P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}) P(X_n = i_n | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\
&= P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}) \underbrace{P(X_n = i_n | X_{n-1} = i_{n-1})}_{p_{i_{n-1} i_n}} \\
&= P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}) p_{i_{n-1} i_n}.
\end{aligned}
$$

Now by induction we get

$$P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = P(X_0 = i_0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

□

Example 3. The state space of a Markov chain is $\mathcal{S} = \{1, 2, 3\}$ and the transition probability matrix

$$\mathbb{P} = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0.5 & 0.5 \end{bmatrix}.$$

The initial state X_0 is drawn uniformly over $\mathcal{S}$, i.e. $P(X_0 = i) = 1/3$ for $i = 1, 2, 3$.

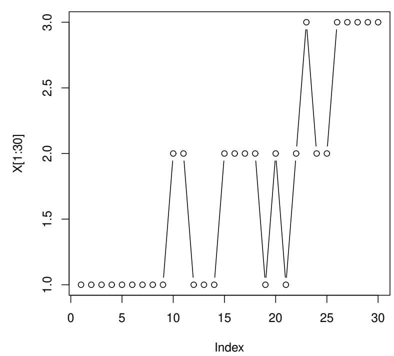
A few examples of joint probabilities:

$$P(X_0 = 1, X_1 = 1, X_2 = 2) = P(X_0 = 1) \times p_{11} \times p_{12} = 1/3 \times 0.7 \times 0.3;$$
$$P(X_0 = 3, X_1 = 2, X_2 = 1) = P(X_0 = 3) \times p_{32} \times p_{21} = 1/3 \times 0.5 \times 0.1;$$
$$P(X_0 = 3, X_1 = 3, X_2 = 1) = P(X_0 = 3) \times p_{33} \times p_{31} = 1/3 \times 0.5 \times 0 = 0.$$

The following code simulates from this Markov chain for $t = 0, 1, \ldots, n-1$. The plot shows the simulated values for the first 30 random variables.

```
S=1:3 #state space
#transition matrix:
P=matrix(c(0.7,0.1,0,0.3,0.6,0.5,0,0.3,0.5),3,3)

n=100
X=numeric(n)
X[1]=sample(S,size=1,prob=c(1/3,1/3,1/3))
for(t in 2:n)
{
  X[t]=sample(S,size=1,prob=P[X[t-1],])
}
```



1.3. n-step transition probabilities

We call $p_{ij}^{(n)} \triangleq P(X_{t+n} = j | X_t = i)$ an n-step transition probability. Put them into the n-step transition matrix $\mathbb{P}^{(n)} = (p_{ij}^{(n)})$.

Theorem 2. *The n-step transition probabilities of an M.C. satisfy*

$$p_{ij}^{(n)} = \sum_{k=0}^{N} p_{ik} p_{kj}^{(n-1)}, \qquad n \geq 2. \tag{1}$$

In matrix notation,

$$\mathbb{P}^{(n)} = \underbrace{\mathbb{P} \times \mathbb{P} \times \cdots \times \mathbb{P}}_{n \text{ factors}} = \mathbb{P}^n \ .$$

Proof. We show (1) by averaging over X_1 in the conditional distribution of

(X_n, X_1) given $X_0 = i$, i.e. $[X_n, X_1 \mid X_0 = i]$:

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j|X_0 = i) = \sum_{k=0}^{N} P(X_n = j, X_1 = k|X_0 = i) \\ &= \sum_{k=0}^{N} P(X_1 = k|X_0 = i)P(X_n = j|X_1 = k, X_0 = i) \\ &\overset{(2)}{=} \sum_{k=0}^{N} p_{ik} P(X_n = j|X_1 = k) = \sum_{k=0}^{N} p_{ik} p_{kj}^{(n-1)}. \end{aligned}$$

To get equality (2), we used the fact that X_n is independent of X_0 given X_1:

$$P(X_n = j|X_1 = k, X_0 = i) = P(X_n = j|X_1 = k).$$

In matrix notation, (1) means $\mathbb{P}^{(n)} = \mathbb{P} \times \mathbb{P}^{(n-1)}$. Now by induction on $\mathbb{P}^{(n-1)}$, we arrive at $\mathbb{P}^{(n)} = \mathbb{P}^n$. □

Example 4. Two-state M.C.

Transition matrix $\mathbb{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$, $0 < a,\ b < 1$.

The n-step transition matrix of this Markov chain is

$$\mathbb{P}^n = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{(a+b)} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}. \tag{2}$$

Proof. By induction. Define

$$\mathbb{A} = \begin{pmatrix} b & a \\ b & a \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

Then (2) becomes

$$\mathbb{P}^n = (a+b)^{-1}[\mathbb{A} + (1-a-b)^n \mathbb{B}].$$

Simple calculation shows that

$$\mathbb{AP} = \begin{pmatrix} b & a \\ b & a \end{pmatrix} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \begin{pmatrix} b & a \\ b & a \end{pmatrix} = \mathbb{A},$$

$$\mathbb{BP} = \begin{pmatrix} a & -a \\ -b & b \end{pmatrix} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = (1-a-b)\mathbb{B}.$$

1. It is easy to verify that (2) holds for $n = 1$:

$$\frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{1-a-b}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \mathbb{P}.$$

2. Assume that (2) holds for n. Then

$$\begin{aligned}\mathbb{P}^{n+1} &= \mathbb{P}^n\mathbb{P} = (a+b)^{-1}[\mathbb{A} + (1-a-b)^n\mathbb{B}]\mathbb{P} \\ &= (a+b)^{-1}[\mathbb{A}\mathbb{P} + (1-a-b)^n\mathbb{B}\mathbb{P}] \\ &= (a+b)^{-1}[\mathbb{A} + (1-a-b)^{n+1}\mathbb{B}],\end{aligned}$$

which shows that (2) holds for $n+1$ as well.

This competes the proof. □

$\because |1-a-b| < 1$ when $0 < a,\ b < 1 \Rightarrow (1-a-b)^n \to 0$, as $n \to \infty$.

$$\therefore \lim_{n\to\infty} \mathbb{P}^n = \tfrac{1}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix} = \begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{pmatrix} = \mathbb{P}^{(\infty)}.$$

$$\lim_{n\to\infty} P(X_n = 1|X_0 = 0) = p_{01}^{(\infty)} = \frac{a}{a+b},\ \lim_{n\to\infty} P(X_n = 1|X_0 = 1) = \frac{a}{a+b}.$$

In the long run, this M.C. will be in state 1 with probability $\frac{a}{a+b}$ and in state 0 with probability $\frac{b}{a+b}$, *independent* of the initial state (X_0). That is,

$$\pi = (\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b}\right)$$

is the *limiting distribution* of this Markov chain. As shown in Example 5, π is also a stationary distribution of this Markov chain.

2. Limiting Behavior of Markov Chains

2.1. Stationary distribution

Definition 1. let $\mathbb{P} = (p_{ij})$ be the transition matrix of a Markov chain on $\{0, 1, \cdots, N\}$, then any distribution $\pi = (\pi_0, \pi_1, \cdots, \pi_N)$ that satisfies the following set of equations is a stationary distribution of this Markov chain:

$$\begin{cases} \pi_j = \sum_{k=0}^{N} \pi_k \cdot p_{kj}, \quad j = 0, 1, \cdots, N; \ \text{(in matrix notation: } \pi = \pi\,\mathbb{P}) \\ \sum_{k=0}^{N} \pi_k = 1. \end{cases} \tag{3}$$

For countable state space, let $N = \infty$. In the above matrix multiplication, π is regarded as a row vector.

Meaning of stationary distributions: If $X_t \sim \pi$ for any time point t, then $X_{t+n} \sim \pi$ for all $n = 1, 2, \ldots$.

It is sufficient to verify this statement for $n = 1$: If $P(X_t = i) = \pi_i$ for all $i = 0, \ldots, N$, we want to verify that $P(X_{t+1} = i) = \pi_i$ for all i as well. This is done as follows: For any state j,

$$P(X_{t+1} = j) = \sum_{k=0}^{N} P(X_t = k) P(X_{t+1} = j | X_t = k) = \sum_{k=0}^{N} \pi_k \cdot p_{kj} = \pi_j, \tag{4}$$

where last equality is from the first line in (3).

The derivation in (4) shows a useful relationship between the marginal distributions of a Markov chain $\{X_t\}$: Let $D_t = (\theta_{t0}, \ldots, \theta_{tN})$ denote the marginal distribution of X_t for $t = 0, 1, \ldots$, i.e. $P(X_t = i) = \theta_{ti}$ for all i. Then

$$\theta_{t+1,j} = \sum_{k=0}^{N} \theta_{tk} \cdot p_{kj}, \qquad j = 0, \ldots, N,$$

or equivalently in matrix notation,

$$D_{t+1} = D_t\,\mathbb{P}. \tag{5}$$

This can be viewed as an iterative algorithm that generates a sequence of distributions D_t. A stationary distribution π is a fixed point of this iteration: If $D_t = \pi$, then $D_{t+1} = \pi$.

Example 5. Two-state M.C.

$\mathbb{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$. Let $\pi = [\pi_0, \pi_1]$.

$\pi = \pi \cdot \mathbb{P}$

$\Rightarrow \; [\pi_0 \quad \pi_1] = [\pi_0 \quad \pi_1] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = [(1-a)\pi_0 + b\pi_1, \; \pi_0 a + \pi_1(1-b)]$

$$\pi_0 = (1-a)\pi_0 + b\pi_1; \tag{6}$$
$$\Rightarrow \quad \pi_1 = \pi_0 a + \pi_1(1-b); \quad \to \text{redundant}$$
$$\pi_0 + \pi_1 = 1. \tag{7}$$

From (6) and (7) $\Rightarrow \begin{cases} \pi_0 = \dfrac{b}{a+b} \\ \pi_1 = \dfrac{a}{a+b} \end{cases}$.

Example 6. $\mathbb{P} = \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.05 & 0.7 & 0.25 \\ 0.05 & 0.5 & 0.45 \end{bmatrix}$, $\pi = (\pi_0, \; \pi_1, \; \pi_2)$.

$$\left.\begin{array}{c} (\pi_0 \quad \pi_1 \quad \pi_2)\mathbb{P} = (\pi_0 \quad \pi_1 \quad \pi_2) \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{array}\right\} \Rightarrow \begin{cases} \pi_0 = \dfrac{1}{13} \\ \pi_1 = \dfrac{5}{8} \\ \pi_2 = \dfrac{31}{104} \end{cases}.$$

2.2. Irreducible Markov chains

Definition 2. State j is accessible from state i if $p_{ij}^{(n)} > 0$ for some $n \geq 0$.

If two states i and j are accessible to each other, then i and j communicate: $i \leftrightarrow j$.

Definition 3. A Markov chain is irreducible if all states communicate with each other.

Example 7.

$$\mathbb{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$0 \leftrightarrow 1$

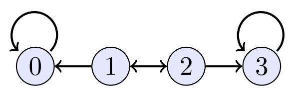
$2 \leftrightarrow 3 \leftrightarrow 4 \Rightarrow 2 \leftrightarrow 4$ (by transitivity)

Partition states into two classes, $\{0, 1\}$ and $\{2, 3, 4\}$: states between the two classes do not communicate.

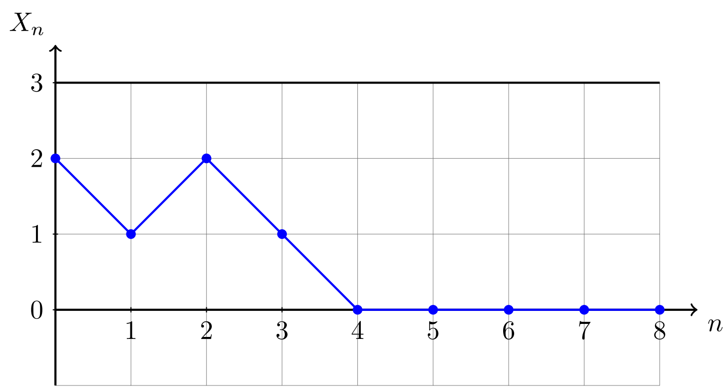
M.C. is reducible.

Example 8 (Random walk with absorbing boundaries)**.**

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



States form three classes: $\{0\}, \{1, 2\}, \{3\}$. Therefore, this Markov chain is reducible.



2.3. Periodicity of a Markov chain

Definition 4. The period of state i, written $d(i)$, is the greatest common divisor (g.c.d.) of all integers $n \geq 1$ for which $p_{ii}^{(n)} > 0$, that is

$$d(i) = \text{g.c.d.}\{n \geq 1 : p_{ii}^{(n)} > 0\}.$$

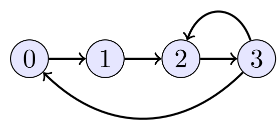
In Example 8, $d(1) = d(2) = 2$, $d(0) = d(3) = 1$. ($p_{ii} > 0 \;\Rightarrow\; d(i) = 1$)

Example 9. N-state M.C. with state space $\{1, 2, \cdots, N\}$.

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} \qquad d(i) = N, \quad \forall i = 1, 2, \cdots, N.$$

Example 10.

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$



$p_{00} = 0$, $p_{00}^{(2)} = 0$, $p_{00}^{(3)} = 0$, $p_{00}^{(4)} = \frac{1}{2}$, $p_{00}^{(5)} = 0$, $p_{00}^{(6)} = \frac{1}{4}$, $\cdots$,

$p_{00}^{(n)} > 0$ if $n = 4, 6, 8, \cdots$. Thus, $d(0) = \text{g.c.d.}\{4, 6, 8, \ldots\} = 2$.

Definition 5. An M.C. in which every state has period 1 is called aperiodic.

Proposition 1. *If a Markov chain is irreducible and $p_{ii} > 0$ for some state i, then the Markov chain is aperiodic.*

Proof. Since $p_{ii} > 0$, $d(i) = 1$. For any $j \neq i$, since the Markov chain is irreducible, there are directed paths $j \mapsto i$ and $i \mapsto j$, which form a directed cycle $j \mapsto i \mapsto j$. Suppose the length of this cycle is $k \geq 2$. Since $p_{ii} > 0$, we may form another directed cycle $j \mapsto i \mapsto i \mapsto j$ which has length $k + 1$. Therefore $d(j)$ is a common divisor of both k and $k + 1$ and thus $d(j) = 1$ for all j. □

2.4. The basic limit theorem

Theorem 3. *Suppose an irreducible and aperiodic Markov chain $X_0, X_1, \ldots$ has a finite state space $\mathcal{S} = \{0, 1, \ldots, N\}$. Let $\pi = (\pi_0, \ldots, \pi_N)$ be a stationary distribution of the Markov chain. Then for any initial state $X_0 = i \in \mathcal{S}$,*

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j \qquad \text{for all } j \in \mathcal{S},$$
$$\frac{1}{n} \sum_{t=1}^{n} h(X_i) \xrightarrow{a.s.} \mathbb{E}_\pi[h(X)].$$

Here, π is called the limiting distribution of this Markov chain. This theorem also implies that the stationary distribution is unique under its assumptions.

Interpretation of the limiting distribution:

1. $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} P(X_n = j | X_0 = i)$.

   In the long run, the probability of the chain in state j is π_j, irrespective of the starting state;

2. The fraction of time in state j for $t = 1, \cdots, n-1, n$, fixing $X_0 = i$, is $F_j = \frac{1}{n}\sum_{t=1}^{n} I(X_t = j)$.
   The mean fraction:

$$\begin{aligned}\mathbb{E}[F_j | X_0 = i] &= \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[I(X_t = j)|X_0 = i] \\ &= \frac{1}{n}\sum_{t=1}^{n} P(X_t = j | X_0 = i) \\ &= \frac{1}{n}\sum_{t=1}^{n} p_{ij}^{(t)} \xrightarrow{n\to\infty} \pi_j.\end{aligned}$$

Example 11. Use the three-state Markov chain in Example 3 to demonstrate the limiting theorem. State space $\mathcal{S} = \{1, 2, 3\}$ and the transition probability matrix

$$\mathbb{P} = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0.5 & 0.5 \end{bmatrix}.$$

It is easy to verify that

$$\pi = (\pi_1, \pi_2, \pi_3) = (5/29, 15/29, 9/29) \approx (0.173, 0.517, 0.310)$$

is a stationary distribution: $\pi\mathbb{P} = \pi$. For a function $h(x) = x$, its expectation with respect to π is

$$\mathbb{E}_\pi[h] = \pi_1 \times 1 + \pi_2 \times 2 + \pi_3 \times 3 \approx 2.138.$$

Based on Theorem 3, we can simulate this Markov chain $X_1, \ldots, X_n$ long enough (n large) and use sample average $\frac{1}{n}\sum_i X_i$ to estimate $\mathbb{E}_\pi[h]$. Running the code in Example 3 with $n = 5,000$, we get the following results:

```
> mean(X)
[1] 2.1604
> sum(X==1)/n
[1] 0.1572
> sum(X==2)/n
[1] 0.5252
> sum(X==3)/n
[1] 0.3176
```

We also see that the frequencies of $\{X_i\}$ in the three states approximate π.

To demonstrate convergence to the limit distribution π, we use the relation in (5) to calculate the marginal distribution $D_t = (\theta_{t1}, \theta_{t2}, \theta_{t3})$ for each X_t, starting with $X_1 = 1$, i.e. $D_1 = (1, 0, 0)$. We also record the absolute deviation between D_t and π:

$$r_t = \sum_{i=1}^{3} |\theta_{ti} - \pi_i|.$$

The following code calculates D_t and r_t for $t = 1, \ldots, 50$. It is seen that $r_t \approx 0$ for $t \geq 20$, showing that the distribution of X_t (`D[t,]`) is already very close to π (`dst`) after 20 iterations.

```
#transition matrix:
P=matrix(c(0.7,0.1,0,0.3,0.6,0.5,0,0.3,0.5),3,3)

n=50;
dst=c(5/29,15/29,9/29); #stationary distribution pi
D=matrix(0,nrow=n,ncol=3);
D[1,]=c(1,0,0); # initial state X_1=1
for(t in 2:n)
{
  D[t,]=D[t-1,]%*%P; #distribution of X_t
}

#calculate absolute deviation between D[t,] and dst
r=numeric(n);
for(t in 1:n){r[t]=sum(abs(D[t,]-dst));}
plot(1:n,r,xlab="t",ylab="r_t")
```

```
> D[1:20,]
          [,1]      [,2]      [,3]
 [1,] 1.0000000 0.0000000 0.0000000
 [2,] 0.7000000 0.3000000 0.0000000
 [3,] 0.5200000 0.3900000 0.0900000
 [4,] 0.4030000 0.4350000 0.1620000
 [5,] 0.3256000 0.4629000 0.2115000
 [6,] 0.2742100 0.4811700 0.2446200
 [7,] 0.2400640 0.4932750 0.2666610
 [8,] 0.2173723 0.5013147 0.2813130
 [9,] 0.2022921 0.5066570 0.2910509
[10,] 0.1922702 0.5102073 0.2975226
[11,] 0.1856098 0.5125667 0.3018235
[12,] 0.1811836 0.5141347 0.3046817
[13,] 0.1782420 0.5151768 0.3065813
[14,] 0.1762870 0.5158693 0.3078437
[15,] 0.1749879 0.5163295 0.3086826
[16,] 0.1741245 0.5166354 0.3092402
[17,] 0.1735507 0.5168386 0.3096107
[18,] 0.1731693 0.5169737 0.3098569
[19,] 0.1729159 0.5170635 0.3100206
[20,] 0.1727475 0.5171232 0.3101293
> dst
[1] 0.1724138 0.5172414 0.3103448
```