Egodic Theorems for discrete Markov chains

Nikolaos Halidias
Department of Mathematics
University of the Aegean
Karlovassi 83200 Samos, Greece
email: nikoshalidias@hotmail.com

February 11, 2018

Abstract

Let $X_n$ be a discrete time Markov chain with state space $S$ (countably infinite, in general) and initial probability distribution $\mu^{(0)} = (P(X_0 = i_1), P(X_0 = i_2), \cdots)$.

What is the probability of choosing in random some $k \in \mathbb{N}$ with $k \leq n$ such that $X_k = j$ where $j \in S$? This probability is the average $\frac{1}{n} \sum_{k=1}^{n} \mu_j^{(k)}$ where $\mu_j^{(k)} = P(X_k = j)$. In this note we will study the limit of this average without assuming that the chain is irreducible, using elementary mathematical tools. Finally, we study the limit of the average $\frac{1}{n} \sum_{k=1}^{n} g(X_k)$ where $g$ is a given function for a Markov chain not necessarily irreducible.

Keywords: Markov chain, average of probability distributions

2010 Mathematics Subject Classification 60J10

1 Introduction

Let $X_n$ be a discrete time Markov chain with state space $S$ (countably infinite, in general) and initial probability distribution $\mu^{(0)}$, that is $\mu_i^{(0)} = P(X_0 = i)$ where $i \in S$. We will study the limit of the average

$$\frac{1}{n} \sum_{k=1}^{n} \mu_j^{(k)}$$

This quantity gives the probability of choosing in random an integer $k$ with $k \leq n$ such that $X_k = j$. Note that, for any $i, j \in S$, we have

$$\frac{1}{n} \sum_{k=1}^{n} \mu_j^{(k)} = \frac{1}{n} \sum_{k=1}^{n} (\mu_i^{(0)} \cdot P^k)_j$$

$$= \frac{1}{n} \sum_{k=1}^{n} \sum_{i \in S} \mu_i^{(0)} P_{ij}^k$$

$$= \sum_{i \in S} \mu_i^{(0)} \frac{1}{n} \sum_{k=1}^{n} P_{ij}^k$$

(1)

Therefore, one can study the desired limit by studying the limit of the average $\frac{1}{n} \sum_{k=1}^{n} P_{ij}^k$. To do so one can use the limit theorems for $P_{ij}^m$ (see for example [2]) and the well known fact
that if \( a_n \to a \) then \( \frac{1}{n} \sum_{k=1}^{n} a_k \to a \). However, here we will give a different proof without using the limit theorems and without assuming that the chain is irreducible. Moreover, we will study the behavior of the limit \( \frac{1}{n} \sum_{k=1}^{n} g(X_k) \) for a given function \( g \), using elementary mathematical tools.

2 The main results

Let \( X_n \) be a Markov chain with (countably infinite in general) state space \( S \).

**Theorem 1** It holds that, for any \( i, j \in S \),

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mu_{j}^{(k)} = \begin{cases} \frac{1}{m_j} \sum_{i \in S} \mu_i^{(0)} f_{ij}, & \text{when } j \text{ is positive recurrent} \\ 0, & \text{otherwise} \end{cases}
\]

and

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} P_{ij}^{k} = \begin{cases} \frac{f_{ij}}{m_j}, & \text{when } j \text{ is positive recurrent} \\ 0, & \text{otherwise} \end{cases}
\]

where \( f_{ij} = P(\exists n \in \mathbb{N} : X_n = j|X_0 = i) \).

**Proof.** We know (see [2]) that when \( j \) is transient or null recurrent \( \lim_{n \to \infty} P_{ij}^{n} = 0 \). Therefore \( \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} P_{ij}^{k} = 0 \) and using (1) the result follows. Next we suppose that \( j \) is positive recurrent.

Let the random variables \( N_{j}^{k} = \begin{cases} 1, & \text{when } X_k = j \\ 0, & \text{otherwise} \end{cases} \) and \( M_j(n) = \sum_{k=1}^{n} N_{j}^{k} \). Because

\[
\mathbb{E}\left( \frac{M_j(n)}{n} \right) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}N_{j}^{k} = \frac{1}{n} \sum_{k=1}^{n} P(X_k = j) = \frac{1}{n} \sum_{k=1}^{n} \mu_{j}^{(k)}
\]

we will study the quantity \( \mathbb{E}\left( \frac{M_j(n)}{n} \right) \).

Let the event \( A_i = \{ \exists n \in \mathbb{N} : X_n = j \} \cap \{X_0 = i\} \) where \( i \in S \). Because \( P(A_i) = P(\exists n \in \mathbb{N} : X_n = j|X_0 = i) \cdot \mu_i^{(0)} \) we see that \( P(A_i) = f_{ij} \cdot \mu_i^{(0)} \) where \( f_{ij} = P(\exists n \in \mathbb{N} : X_n = j|X_0 = i) \).

We will work under the probability measure \( P_{A_i} \cdot P(\cdot|A_i) \) while the corresponding expected value will be denoted by \( \mathbb{E}_{A_i} \).

We define the following sequence of random variables,

\[
n_{1}(\omega) = \begin{cases} \min\{n \in \mathbb{N} : X_n(\omega) = j\}, & \text{when } \omega \in A_i \\ \infty, & \text{otherwise} \end{cases}
\]

\[
n_{2}(\omega) = \begin{cases} \min\{n > n_{1} : X_n(\omega) = j\}, & \text{when } \omega \in A_i \\ \infty, & \text{otherwise} \end{cases}
\]

\[
\vdots
\]

\[
n_{k}(\omega) = \begin{cases} \min\{n > n_{k-1} : X_n(\omega) = j\}, & \text{when } \omega \in A_i \\ \infty, & \text{otherwise} \end{cases}
\]
We define also $Z_m = \begin{cases} n_{m+1} - n_m, & \text{when } \omega \in A_i \\ 0, & \text{otherwise} \end{cases}$ for $m \geq 1$ which gives us the number of transitions needed to return back to $j$. Note that the sequence $Z_1, Z_2, \ldots$, is an independent and identically distributed sequence of random variables. The mean recurrent time $m_j$ is such that $m_j = E_{A_i}(Z_k)$ for every $k \geq 1$. Next we define the random variable $S_l = Z_1 + \cdots + Z_l$ with $S_0 = 0$. Note that

$$S_l + n_1 = n_{l+1} \quad \text{for every } l \geq 0 \quad (3)$$

Using the strong law of large numbers we have that

$$P_{A_i} \left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{S_n}{n} = m_j \} \right) = 1$$

Note that $M_j(n) \to \infty$ as $n \to \infty$ for almost all $\omega \in \Omega$ when $j$ is recurrent and its easy to see that $n_{M_j(k)} \leq k$ for every $k \geq 1$.

Using (3) we see that the following inequality hold

$$S_{M_j(n)-1} + n_1 \leq n \leq S_{M_j(n)} + n_1, \quad n \geq 1, \quad \text{for every } \omega \in A_i$$

Dividing by $M_j(n) > 0$ for $n > n_1$ we get

$$\frac{S_{M_j(n)-1} + n_1}{M_j(n) - 1} \leq \frac{n}{M_j(n)} \leq \frac{S_{M_j(n)}}{M_j(n)}, \quad n \geq n_2, \quad \text{for every } \omega \in A_i$$

Therefore it holds that

$$P_{A_i} \left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \} \right) = 1 \quad (4)$$

Next we will study the limit of the quantity

$$\lim_{n \to \infty} \frac{E_{A_i}(M_j(n))}{n}$$

Using the dominated convergence theorem it follows that

$$\lim_{n \to \infty} \frac{E_{A_i}(M_j(n))}{n} = E_{A_i} \left( \lim_{n \to \infty} \frac{M_j(n)}{n} \right) = E_{A_i} \left( \frac{1}{m_j} \right) = \frac{1}{m_j}$$

But, since

$$E_{A_i} \left( \frac{M_j(n)}{n} \right) = \frac{E \left( \frac{M_j(n) 1_{A_i}}{n} \right)}{P(A_i)}$$

it follows that

$$\lim_{n \to \infty} E \left( \frac{M_j(n)}{n} 1_{A_i} \right) = \frac{P(A_i)}{m_j} f_{ij} \quad \mu_i^{(0)} \frac{f_{ij}}{m_j} \quad (5)$$
Because
\[ \mathbb{E}\left( \frac{M_j(n)}{n} \right) = \sum_{i \in S} \mathbb{E}\left( \frac{M_j(n)}{n} I_{A_i} \right) \]
we obtain using \(5\)

\[
\lim_{n \to \infty} \mathbb{E}\left( \frac{M_j(n)}{n} \right) = \lim_{n \to \infty} \sum_{i \in S} \mathbb{E}\left( \frac{M_j(n)}{n} I_{A_i} \right)
= \sum_{i \in S} \lim_{n \to \infty} \mathbb{E}\left( \frac{M_j(n)}{n} I_{A_i} \right)
= \sum_{i \in S} \mu_i^{(0)} f_{ij}
= \frac{1}{m_j} \sum_{i \in S} \mu_i^{(0)} f_{ij}
\]
where we have used the dominated convergence theorem to get the second equality above.

If \(m_{ij}(n) = \mathbb{E}(M_j(n)|X_0 = i)\) then we have

\[
m_{ij}(n) = \mathbb{E}(M_j(n)|X_0 = i)
= \mathbb{E}\left( \sum_{k=1}^{n} N_j^k |X_0 = i \right)
= \sum_{k=1}^{n} \mathbb{E}(N_j^k |X_0 = i)
= \sum_{k=1}^{n} P_{ij}^k
\]

Denoting by \(A = \{ \exists k \in \mathbb{N} : X_k = j \}\), we have

\[
\mathbb{E}\left( \frac{M_j(n)}{n} |X_0 = i \right) = \mathbb{E}\left( \frac{M_j(n)}{n} I_A |X_0 = i \right) + \mathbb{E}\left( \frac{M_j(n)}{n} I_{A^c} |X_0 = i \right)
= \mathbb{E}\left( \frac{M_j(n)}{n} I_A |X_0 = i \right)
= \mathbb{E}\left( \frac{M_j(n)}{n} I_{A_i} \right)
\]
because \(M_j(n)I_{A^c} = 0\). That means that

\[
\lim_{n \to \infty} \frac{m_{ij}(n)}{n} = \lim_{n \to \infty} \mathbb{E}\left( \frac{M_j(n)}{n} |X_0 = i \right) = \frac{f_{ij}}{m_j}
\]

Therefore

\[
\lim_{n \to \infty} \frac{\sum_{k=1}^{n} P_{ij}^k}{n} = \begin{cases} \frac{f_{ij}}{m_j}, & \text{when } j \text{ is positive recurrent} \\ 0, & \text{otherwise} \end{cases}
\]
Proposition 1. It holds that, when $j$ is positive recurrent,

$$\{\omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j}\} \cup \{\omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = 0\} = \Omega \setminus E$$

with $P(E) = 0$. More precisely, it holds that

$$P\left( \left\{ \omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \right\}\right) = \sum_{i \in S} \mu_i^{(0)} \cdot f_{ij}$$

and

$$P\left( \left\{ \omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \right\}\right) = \sum_{i \in S} \mu_i^{(0)} \cdot (1 - f_{ij})$$

If $j$ is null recurrent or transient, then

$$P\left( \left\{ \omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \right\}\right) = 1$$

Proof.

Assume that $j$ is positive recurrent. Denoting by $B = \{\omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j}\}$ we can write

$$B = \bigcup_{i \in S} \{\omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j}\} \cap \{X_0 = i\} = \bigcup_{i \in S} B_i$$

and therefore $P(B) = \sum_{i \in S} P(B_i)$.

But

$$B_i = B_i \cap \{\exists k \in \mathbb{N}: X_k = j\} \cup B_i \cap \{\nexists k \in \mathbb{N}: X_k = j\}$$

so $P(B_i) = P(B_i \cap \{\exists k \in \mathbb{N}: X_k = j\}) + P(B_i \cap \{\nexists k \in \mathbb{N}: X_k = j\})$. Recalling $\mu$ we can write that

$$P(B_i \cap \{\exists k \in \mathbb{N}: X_k = j\}) = P_{A_i} \left( \left\{ \omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \right\} \right) \cdot P(A_i) = \mu_i^{(0)} \cdot f_{ij}$$

Moreover

$$P(B_i \cap \{\nexists k \in \mathbb{N}: X_k = j\}) = 0$$

since in this event $M_j(n) = 0$. Therefore $P(B_i) = \mu_i^{(0)} \cdot f_{ij}$ and thus

$$P(B) = \sum_{i \in S} \mu_i^{(0)} \cdot f_{ij}$$

Denote now $\Gamma_i = \{k \in \mathbb{N}: X_k = j\} \cap \{X_0 = i\}$. Then

$$P_{\Gamma_i} \left( \left\{ \omega \in \Omega: \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \right\} \right) = 1$$
where \( P_{\Gamma_i}(\cdot) = P(\cdot | \Gamma_i) \). Thus
\[
P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0\} \cap \Gamma_i \right) = P(\Gamma_i) = \mu_i^{(0)}(1 - f_{ij})
\]
That means that
\[
P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0\} \cap \Gamma \right) = \sum_{i \in S} \mu_i^{(0)}(1 - f_{ij})
\]
where \( \Gamma = \{ \not\exists k \in \mathbb{N} : X_k = j \} \). Thus
\[
P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0\} \right) \geq \sum_{i \in S} \mu_i^{(0)}(1 - f_{ij})
\]
The events
\[
\{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \} \quad \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \}
\]
are disjoint, therefore
\[
1 \leq P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \} \right) + P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) = P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \} \cup \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) \leq 1
\]
Therefore
\[
\{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = \frac{1}{m_j} \} \cup \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} = \Omega \setminus E
\]
with \( P(E) = 0 \) and
\[
P \left( \{\omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) = \sum_{i \in S} \mu_i^{(0)}(1 - f_{ij})
\]
Assume now that \( j \) is null recurrent and let the sequence of random variables \( Z_m = \left\{ n_{m+1} - n_m, \quad \omega \in A_i \middle| m \geq 1 \right\} \). Because \( j \) is null recurrent we have that \( \mathbb{E}(Z_m) = \infty \) for every \( m \geq 1 \). We define now the sequence \( Z^R_m = Z_m \mathbb{1}_{\{Z_m < R\}} \) for \( R > 0 \) for which it holds that \( \mathbb{E}(Z^R_m) < \infty \) for every \( m \geq 1 \). Moreover, \( \mathbb{E}(Z^R) = \mathbb{E}(Z^R_m) \) for every \( m \geq 1 \). This sequence is again an independent and identical distributed sequence of random variables. Therefore we can use the strong law of large numbers to get
\[
P_{A_i} \left( \omega \in \Omega : \lim_{n \to \infty} \frac{S^R_n}{n} = \mathbb{E}_{A_i}(Z^R_1) \right) = 1
\]
where \( S^R_n = Z^R_1 + Z^R_2 + \cdots + Z^R_n \leq S_n = Z_1 + \cdots + Z_n \) and \( A_i, P_{A_i} \) is as before. Therefore it holds that
\[
S^R_{M_j(n)-1} + n_1 \leq S_{M_j(n)-1} + n_1 \leq n
\]
So
\[
\frac{S_{n+1}^{R}}{M_{j}(n) - 1} \leq \frac{n}{M_{j}(n)} \leq \frac{1}{E(Z_{m}^{R})}, \quad n \geq n_{2}, \quad \text{for every } \omega \in A_{i}
\]
Letting \( n \to \infty \) we get that
\[
0 \leq \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \leq 1, \quad \text{almost surely, for every } R > 0
\]
under the probability measure \( P_{A_{i}} \). Note that \( Z_{m}^{R} \) is an increasing sequence in \( R \) and that \( Z_{m}^{R} \to Z_{m} \) as \( R \to \infty \) almost surely. Therefore \( E_{A_{i}}(Z_{m}^{R}) \to E_{A_{i}}(Z_{m}) = \infty \) using the monotone convergence theorem. That means that
\[
\lim_{n \to \infty} \frac{M_{j}(n)}{n} = 0 \quad \text{almost surely}
\]
derived probability measure \( P_{A_{i}} \), i.e.
\[
P_{A_{i}} \left( \left\{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_{j}(n)}{n} = 0 \right\} \right) = 1 \tag{6}
\]
Let now the event \( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \) where \( \varepsilon > 0 \). Noting that
\[
P \left( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \cap A^{c} \right) = 0
\]
where \( A = \{ \exists l \in \mathbb{N} : X_{l} = j \} \) and
\[
P \left( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \cap A \right) = \sum_{i \in S} P \left( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \cap A_{i} \right)
\]
\[
= \sum_{i \in S} P_{A_{i}} \left( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \right) P(A_{i})
\]
\[
= 0, \quad \text{see } [6]
\]
we obtain
\[
P \left( \left\{ \omega \in \Omega : \limsup_{n \to \infty} \frac{M_{j}(n)}{n} \geq \varepsilon \right\} \right) = 0
\]
Because \( \frac{M_{j}(n)}{n} \geq 0 \) it follows the desired result.

Finally we assume that \( j \) is transient. It is well known that \( P(M_{j} < \infty | X_{0} = i) = 1 \) for every state \( i \), where \( M_{j} = \lim_{n \to \infty} M_{j}(n) \). Therefore
\[
P(M_{j} < \infty) = \sum_{i \in S} P(M_{j} < \infty | X_{0} = i) \cdot P(X_{0} = i) = \sum_{i \in S} \mu_{i}^{(0)} = 1
\]
Moreover
\[
\Omega = \left( \bigcup_{N=0}^{\infty} B_{N} \right) \cup B_{\infty}
\]
where $B_N = \{ M_j = N \}$ and $B_\infty = \{ M_j = \infty \}$. Thus
\[
\sum_{N=0}^\infty P(B_N) = 1
\]
since $P(B_\infty) = 0$.

Therefore we can write
\[
\{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} = \left( \bigcup_{N=0}^\infty \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) \cap B_N \cup \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \cap B_\infty
\]
Thus
\[
P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) = \sum_{N=0}^\infty P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \cap B_N \right)
\]
since $P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \cap B_\infty \right) \leq P(B_\infty) = 0$. But
\[
P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right)
= P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \cap B_N \right) P(B_N)
= P(B_N)
\]
since it holds that $P\left( \{ \omega \in \Omega : \lim_{n \to \infty} \frac{M_j(n)}{n} = 0 \} \right) = 1$. Since $\sum_{N=0}^\infty P(B_N) = 1$ we obtain the desired result. \hfill \square

**Corollary 1** If $g : S \to \mathbb{R}$ is such that
\[
\sum_{i \in S} |g(i)| < \infty
\]
then it holds that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}g(X_k) = \sum_{j \in C} g(j) \sum_{i \in S} \mu_i^{(0)} f_{ij}
\]
where $C \subseteq S$ is the subset of $S$ of positive recurrent states.

**Proof.** Note that $g(X_k) = \sum_{j \in S} g(j) \mathbb{1}_{\{X_k = j\}}$. Therefore
\[
\frac{1}{n} \sum_{k=1}^n \mathbb{E}g(X_k) = \frac{1}{n} \sum_{k=1}^n g(j) \sum_{i \in S} \mathbb{1}_{\{X_k = j\}}
= \sum_{j \in S} g(j) \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k = j\}}
= \sum_{j \in S} g(j) \mathbb{E}\left( \frac{M_j(n)}{n} \right)
\]
We have interchange the sums $\sum_{k=1}^n \sum_{j \in S}$ because the series is absolutely convergent since $\sum_{i \in S} |g(i)| < \infty$.

So

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E} g(X_k) = \lim_{n \to \infty} \sum_{j \in S} g(j) \mathbb{E} \left( \frac{M_j(n)}{n} \right) = \sum_{j \in C} \frac{g(j)}{m_j} \sum_{i \in S} \mu_i^{(0)} f_{ij}$$

We have used the dominated convergence theorem to interchange the limit with the sum in the second equality above.

\[ \square \]

**Corollary 2** Given a function $g : S \to \mathbb{R}$ such that

$$\sum_{i \in S} |g(i)| < \infty$$

it holds that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n g(X_k) = \sum_{j \in C} \frac{g(j)}{m_j} \mathbb{1}_{A^j} \text{ almost surely}$$

where $A^j = \{ \omega \in \Omega : \exists l \in \mathbb{N} : X_l = j \}$.

**Proof.** Note that

$$\frac{1}{n} \sum_{k=1}^n g(X_k) = \frac{1}{n} \sum_{k=1}^n \sum_{j \in S} g(j) \mathbb{1}_{\{X_k=j\}} = \sum_{j \in S} g(j) \frac{M_j(n)}{n}$$

$$= \sum_{j \in C} g(j) \frac{M_j(n)}{n} \mathbb{1}_{A^j} + \sum_{j \in NR} g(j) \frac{M_j(n)}{n} + \sum_{j \in T} g(j) \frac{M_j(n)}{n}$$

where $C \subseteq S$ is the subset of positive recurrent states of $S$, $NR \subseteq S$ is the subset of null recurrent states of $S$, $T \subseteq S$ is the subset of transient states of $S$ and $A^j = \{ \omega \in \Omega : \exists l \in \mathbb{N} : X_l = j \}$. The condition on $g$, i.e. $\sum_{i \in S} |g(i)| < \infty$ is needed in order to interchange the sums to get the second equation above.

Note that $A^j = \bigcup_{l \in S} \{ \exists \ l \in \mathbb{N} : X_l = j \} \cap \{ X_0 = i \}$ and therefore $P(A^j) = \sum_{i \in S} \mu_i^{(0)} f_{ij}$.

Finally, using proposition \[\square\] we obtain the desired result, i.e.

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n g(X_k) = \sum_{j \in C} \frac{g(j)}{m_j} \mathbb{1}_{A^j}, \text{ almost surely}$$

\[ \square \]

This corollary is closely related to Birkhoff’s ergodic theorem (see for example \[4\] and \[7\]).
References

[1] R. Ash, *Basic Probability Theory*, Wiley, 1970.

[2] R. Ash - C. Doleans-Dade, *Probability and Measure Theory*, Elsevier, 2000.

[3] P. Bremaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*, Springer, 1999.

[4] R. Durrett, *Probability: Theory and Examples*, Cambridge University Press, 2010.

[5] G. Grimmett - D. Stirzaker, *Probability and Random Processes*, Oxford University Press, 2001.

[6] P. Hoel - S. Port - C. Stone, *Introduction to Stochastic Processes*, Houghton Mifflin Company, 1972.

[7] I. Gikhman - A. Skorohod, *Introduction to the Theory of Random Processes*, 1969.

[8] J. Jacod - P. Protter, *Probability Essentials*, Springer, 2004.

[9] S. Karlin - H. Taylor, *A First Course in Stochastic Processes*, Academic Press, 1975.

[10] J. Kemeny - J. Snell, *Finite Markov Chains*, Springer, 1976.

[11] J. Norris, *Markov Chains*, Cambridge University Press, 1997.

[12] D. Revuz, *Markov Chains*, North Holland, 1984.

[13] D. Stroock, *An Introduction to Markov Processes*, Springer, 1999.