INDEPENDENT \$ \mu \$ TIMES
A new proof of the basic limit theorem
of Markov chains

BY
HERMANN THORISSON

TECHNICAL REPORT NO. 243
FEBRUARY 1986

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MCS80-24649

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
INDEPENDENT μ TIMES

A new proof of the basic limit theorem
of Markov Chains

By
Hermann Thorisson

TECHNICAL REPORT NO. 243
February 1986

PREPARED UNDER THE AUSPICES

OF

NATIONAL SCIENCE FOUNDATION GRANT MCS80-24649

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
INDEPENDENT $\mu$ TIMES

A new proof of the basic limit theorem of Markov chains

Hermann Thorisson

Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg
Sweden

Department of Statistics
Stanford University
Stanford, CA 94305
USA

Abstract

A random time $T$ is an independent $\mu$ time for a Markov chain $(X_n)^\infty_0$ if $T$ is independent of $(X_{T+n})^\infty_{n=0}$ and if $(X_{T+n})^\infty_{n=0}$ is a Markov chain with initial distribution $\mu$ and the same transition probabilities as $(X_n)^\infty_0$. This concept is used to give a new and short proof of the basic limit theorem of Markov chains, improving somewhat the result in the null-recurrent case.
Introduction

Let $X = (X_n)_{n=0}^\infty$ be an irreducible, aperiodic, recurrent Markov chain on a countable state space $E$ and let $T$ be a randomized stopping time w.r.t. $X$, i.e. for each $n \geq 0$ \{T = n\} is conditionally independent of $(X_{n+k})_{k=0}^\infty$ given $(X_k)_0^n$. With $\mu$ a probability distribution on $E$, call $T$ an independent $\mu$ time if $T < \infty$ a.s.,

(i) $T$ is independent of $X_T$, and (ii) $X_T$ is governed by $\mu$.

If $\mu$ is a stationary distribution call $T$ an independent stationary time; this is a generalization of the notion of “strong uniform times” used by Aldous and Diaconis [1], [2] to prove “non-asymptotics” for certain random walks on finite groups (in that case the stationary distribution is uniform - and “strong” means “independent”).

Here independent $\mu$ times are used to prove a strong version of the basic limit theorem for recurrent Markov chains. In the positive recurrent case the limit result is the same as the one obtained by the so called coupling method. The present approach is probably not as intuitively appealing as the coupling one, but it has the advantage of covering also the null-recurrent case without additional effort. Further, it establishes the class property of positive recurrence and also the equivalence of positive recurrence on the one hand and the existence and uniqueness of a stationary distribution on the other. Finally, it yields a slightly improved limit result in the null-recurrent case.

In Section 1 we establish notation and formulate the limit theorem, in Section 2 discuss the relation between independent stationary times and coupling, in Section 3 we prove the key existence result for independent $\mu$ times and finally in Section 4 prove the limit theorem.

1. The theorem

Let $\lambda$ be the initial distribution of $X$ and $P^n = (P^n_{ij} : i, j \in E)$ the $n$-step transition matrix. We regard measures on $E$ as row-vectors, e.g. $\lambda = (\lambda_j : j \in E)$ and $\lambda_A = \sum_{j \in A} \lambda_j, A \subset E$. Thus $\lambda P^n$ is the distribution of $X_n$, i.e

$$\lambda P^n_A = \sum_{j \in A} \lambda P^n_j = \mathbb{P}(X_n \in A), \quad A \subset E.$$ 

Put for $j \in E$

$$N_j = \inf\{n \geq 1 : X_n = j\}$$

and

$$m_j = \mathbb{E}_j[N_j],$$

2
where $E_j$ indicates $X_0 = j$ a.s. Fix an arbitrary $i \in E$ and define a measure $\hat{\pi}$ on $E$ by

$$\hat{\pi}_A = \mathbb{E}_i \left[ \sum_{n=1}^{N_i} I_{\{X_n \in A\}} \right] = \sum_{n=1}^{\infty} \mathbb{P}_i (X_1 \neq i, \ldots, X_{n-1} \neq i, X_n \in A), A \subset E,$$

where $I_B = 1$ or 0 according as $B$ occurs or not. It is easily checked that

(1a) $0 < \hat{\pi}_j < \infty, j \in E,$

(1b) $\hat{\pi}_i = 1$

(1c) $\hat{\pi}_E = m_i$

(1d) $\hat{\pi} = \hat{\pi} P^n, n \geq 0.$

For $c < \infty$ put

$$\mathcal{E}_c = \{ A \subset E : \hat{\pi}_A \leq c \}.$$

**Theorem 1.** Either all states are positive recurrent, $\pi = \left( \frac{1}{m_j}, j \in E \right)$ is a unique stationary distribution and for all initial distributions $\lambda$

(2) $\mathbb{P}(X_n \in A) \to \pi_A$ uniformly in $A \subset E$ as $n \to \infty,$

or all states are null-recurrent, no stationary distribution exists and for all initial distributions $\lambda$ and all $c < \infty$

(3) $\mathbb{P}(X_n \in A) \to 0$ uniformly in $A \in \mathcal{E}_c$ as $n \to \infty.$

**Remark.** Here (3) seems to be a new result improving somewhat the classical one: $\mathbb{P}(X_n = j) \to 0$ as $n \to \infty.$ On the other hand (2) is the typical result obtained by the coupling method. With $\| \cdot \|$ denoting the total variation norm we have

(4) $\| \lambda P^n - \pi \| = 2 \sup_{A \subset E} \left( \lambda P^n_A - \pi_A \right) = 2 \sup_{A \subset E} \left( \pi_A - \lambda P^n_A \right)$

and thus (2) can be rewritten on the form

(2') $\| \lambda P^n - \pi \| \to 0$ as $n \to \infty.$

This is maybe the more appropriate form but we have chosen (2) to stress the resemblance between (2) and (3).
2. Independent stationary times and coupling

If $T$ is an independent $\mu$ time then for $k \leq n$

$$\mathbb{P}(T = k, X_n \in A) = \mathbb{P}(T = k)\mathbb{P}(X_{T+n-k} \in A) = \mathbb{P}(T = k)\mu P^{n-k}_A$$

and thus

$$\mathbb{P}(X_n \in A) = \sum_{k=0}^{n} \mathbb{P}(T = k, X_n \in A) + \mathbb{P}(T > n, X_n \in A),$$

$$\leq \sum_{k=0}^{n} \mathbb{P}(T = k)\mu P^{n-k}_A + \mathbb{P}(T > n).$$

(5)

In particular, if $\mu = \pi$ where $\pi$ is a stationary distribution then $\pi P^{n-k}_A = \pi_A$ and thus

$$\mathbb{P}(X_n \in A) - \pi_A \leq \mathbb{P}(T > n).$$

Applying (4) yields

(6) \[ \| \lambda P^n - \pi \| \leq 2\mathbb{P}(T > n) \rightarrow 0 \text{ as } n \rightarrow \infty \]

proving (2') and thus (2).

Now the inequality in (6) looks exactly like a coupling inequality and this is no coincidence: Let $X'$ be a Markov chain with initial distribution $\pi$ and independent of $T$. Then clearly $(T, (X_{T+k})_{k=0}^{\infty})$ and $(T, (X'_{T+k})_{k=0}^{\infty})$ have the same distribution and we have established a distributional coupling (see [2]) with $T$ as a coupling epoch. Hence the inequality in (6) is a coupling inequality.

We have seen that an independent stationary time can always be regarded as a coupling epoch, - the converse is obviously not true. What is more, while there always exists a coupling epoch such that (6) holds (see [3]) the same is not true for independent stationary times as can be seen from the following counter-example:

Let $X$ be positive recurrent and such that for each $n \geq 0$ there is a $j_n$ such that $P^n_{i, j_n} = 0$ (e.g. consider a random walk on the positive integers with negative drift, reflected at 0 and with bounded step-lengths). Then if $T$ is an independent stationary time we have by (i) and (ii)

$$\mathbb{P}_i(T = n)\pi_{j_n} = \mathbb{P}_i(T = n, X_n = j_n) \leq \mathbb{P}_i(X_n = j_n) = 0.$$

But $\pi_{j_n} > 0$ and thus $\mathbb{P}(T = n) = 0$ for all $n \geq 0$ contradicting $T < \infty$ a.s.
3. Independent $\mu$ times

The above counter-example shows that independent $\mu$ times do not exist in general. However, the following holds:

**Proposition 1.** If $\mu_B = 1$ where $B \subseteq E$ is finite then there exists an independent $\mu$ time.

**Remark.** If $E$ is finite, we obtain (2) from Proposition 1 and (6) by putting $\pi = \hat{\pi}/m_i$.

**Proof of the proposition.** We shall use the following well-known result:

$$\forall i, j \in E \exists n_{i,j} : P^n_{i,j} > 0 \text{ for } n \geq n_{i,j}.$$ 

Fix an $i_0 \in E$ and put

$$n_0 = \max_{j \in B} n_{i_0,j} \quad \text{and} \quad \epsilon = \min_{j \in B} P^{n_0}_{i_0,j} > 0.$$ 

Put $T_0 = 0$ and for $k \geq 1$

$$T_k = \inf\{n \geq T_{k-1} + n_0 : X_n = i_0\}.$$ 

Let $I_{k,j}, j \in E, k \geq 1$, be independent 0-1-variables that are independent of $X$ and such that

$$P(I_{k,j} = 1) = \frac{\epsilon \mu_j}{P^{n_0}_{i_0,j}}, \quad j \in E, k \geq 1.$$ 

Put $I_k = I_{k,x_{T_k+n_0}}$. Then clearly

$$(X_{T_k+n_0}, I_k) \text{ is independent of } (T_k, I_{k-1}, \ldots, I_1).$$

Further,

$$(8) \quad P(X_{T_k+n_0} = j, I_k = 1) = P(X_{T_k+n_0} = j)P(I_{k,j} = 1) = P^{n_0}_{i_0,j} \frac{\epsilon \mu_j}{P^{n_0}_{i_0,j}} = \epsilon \mu_j$$

and summing over $j \in E$ yields

$$P(I_k = 1) = \epsilon.$$ 

Put

$$T = T_K + n_0 \quad \text{where} \quad K = \inf\{k \geq 1 : I_k = 1\}.$$ 

Then $\{T = n\}$ is determined by $(X_k)_{i_0}^n$ and the $I_{k,j}$'s and thus $T$ is a randomized stopping time. Further, (7) and (8) yield the second equality in

$$P(X_T = j, T = n, K = k) = P(X_{T_k+n_0} = j, T_k+n_0 = n, I_k = 1, I_{k-1} = \ldots = I_1 = 0)$$

$$= \epsilon \mu_j P(T_k+n_0 = n, I_{k-1} = \ldots = I_1 = 0)$$

$$= \mu_j P(T_k+n_0 = n, I_k = 1, I_{k-1} = \ldots = I_1 = 0)$$

$$= \mu_j P(T = n, K = k)$$

5
while the third follows from (7) and (9). Summing over \( k \) yields \( \mathcal{P}(X_T = j, T = n) = \mu_j \mathcal{P}(T = n) \) and thus (i) and (ii) hold and the proof is complete.

4. Proof of the theorem

Take a finite \( B \subset E \) and define \( \mu \) by \( \mu_A = \frac{\hat{\pi}}{\pi_B} \). Then \( \mu \leq \hat{\pi} \) yields the inequality in

\[
\mu \mathcal{P}_A^{n-k} \leq \frac{\hat{\pi} \mathcal{P}_A^{n-k}}{\hat{\pi}_B} = \frac{\hat{\pi}_A}{\hat{\pi}_B}, \quad A \subset E, k \leq n,
\]

while the equality is due to (1d). From this, (5) and Proposition 1 we obtain

\[
\mathcal{P}(X_n \in A) \leq \frac{\hat{\pi}_A}{\hat{\pi}_B} + \mathcal{P}(T > n).
\]

Subtracting \( \hat{\pi}_A/m_i = \hat{\pi}_A/\hat{\pi}_E \) from both sides yields

\[
\sup_{A \in \xi} (\mathcal{P}(X_n \in A) - \frac{\hat{\pi}_A}{m_i}) \leq \sup_{A \in \xi} \left( \frac{\hat{\pi}_A}{\hat{\pi}_B} - \frac{\hat{\pi}_A}{\hat{\pi}_E} \right) + \mathcal{P}(T > n)
\]

\[
\leq \frac{c}{\hat{\pi}_B} - \frac{c}{\hat{\pi}_E} + \mathcal{P}(T > n)
\]

\[
\rightarrow \frac{c}{\hat{\pi}_B} - \frac{c}{\hat{\pi}_E} \text{ as } n \rightarrow \infty
\]

\[
\rightarrow 0 \quad \text{as } B \uparrow E.
\]

If \( i \) is null-recurrent \( \hat{\pi}_A/m_i = 0 \) and (3) is established. If \( i \) is positive recurrent put \( c = m_i = \hat{\pi}_E \) to obtain (2) with \( \pi = \hat{\pi}/m_i \).

Since (2) and (3) cannot hold simultaneously and since \( i \in E \) is arbitrary, either all states are null-recurrent or all states are positive recurrent. In the latter case the limit \( \pi = \hat{\pi}/m_i \) is a stationary distribution due to (1c) and (1d). Further, \( \pi \) must be independent of \( i \) and thus with \( i = j, \pi_j = 1/m_j \) due to (1b). Finally, if \( \pi' \) is a stationary distribution, then with \( \lambda = \pi' \) we have \( \lim_{n \rightarrow \infty} \mathcal{P}(X_n \in A) = \pi'_A \) and thus (3) cannot hold, i.e. \( X \) must be positive recurrent, - but then (2) holds implying \( \pi'_A = \pi_A \) so \( \pi \) is unique and the proof is complete.
References

