Probab. Th. Rel. Fields 78, 143–148 (1988)

Probability Theory and Related Fields

Future Independent Times and Markov Chains*

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

Summary. A random time T is a future independent μ time for a Markov chain $(X_n)_0^\infty$ if T is independent of $(X_{T+n})_{n=0}^\infty$ and if $(X_{T+n})_{n=0}^\infty$ is a Markov chain with initial distribution μ and the same transition probabilities as $(X_n)_0^\infty$. This concept is used (with μ the "conditional stationary measure") to give a new and short proof of the basic limit theorem of Markov chains, improving somewhat the result in the null-recurrent case.

Key words: Markov chain, future independent time, regeneration, stationary measure, coupling.

Introduction

Let $X=(X_n)_0^\infty$ be an irreducible, aperiodic, recurrent Markov chain on a countable state space E and let T be a non-negative a.s. finite random time. Call T *future independent* if T is independent of $(X_{T+k})_{k=0}^\infty$. Call T a μ *time* if $(X_{T+k})_{k=0}^\infty$ is a Markov chain with initial distribution μ and the same transition probabilities as X.

Suppose T is a *randomized stopping time*, i.e., for each $n \geqq 0$ the event $\{T=n\}$ is conditionally independent of $(X_{n+k})_{k=0}^\infty$ given $(X_k)_{k=0}^n$. Then T is a future independent μ time if

(i) T is independent of X_T and
(ii) X_T has the distribution μ,

due to the strong Markov property for randomized stopping times.

The stopping times

$$N_j=\inf\{n\geqq 1:\ X_n=j\}, \qquad j\in E,$$

are future independent δ_j times where δ_j has mass one at j. Further, let Y be a random element in E with an arbitrary distribution μ and let Y be independent

* This work was supported by the Swedish Natural Science Research Council and done while the author was visiting the Department of Statistics, Stanford University

of X. Then the randomized stopping time N_Y is a μ time. However, N_Y is not in general a future independent time. In fact, future independent μ times need not exist in general: see the counter-example at the end of Sect. 2.

If the state space E is general and X is a Harris chain then the regeneration times of X (see, e.g., Asmussen [3], Chapter VI.3) are future independent μ times where μ is the regeneration distribution. Actually, (if we allow us to use the notion "regeneration time" without a reference to a sequence of regeneration times) a future independent time T is a regeneration time in the sense of Asmussen ([3], Chapter V.1) and a time-homogeneous regeneration time in the sense of Thorisson [5], [7]. However, T is not in general a regeneration time in the traditional sense since $(X_{T+k})_{k=0}^{\infty}$ need not be independent of $(X_n)_{0 \leq n < T}$.

Call T a *stationary time* if T is a π time where π is a stationary distribution for X. Examples of future independent stationary times are the "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 "future independent").

Here future independent μ times, with μ the "conditional stationary measure", 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 (see, however, Thorisson [8]). 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 Sect. 1 we establish notation and formulate the limit theorem, in Sect. 2 we discuss the relation between future independent stationary times and coupling, in Sect. 3 we prove the key existence result for future independent μ times and, finally, in Sect. 4 we prove the limit theorem.

1. The Limit Theorem

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

$$\lambda P_A^n = \sum_{j \in A} \lambda P_j^n = \mathbb{P}(X_n \in A), \qquad A \subseteq E.$$

Put

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

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

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

where $I_B = 1$ or 0 according as B occurs or not. It is easily checked (see, e.g. Asmussen [3], Theorem I.3.2) that

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

$$\hat{\pi}_i = 1 \tag{1b}$$

$$\hat{\pi}_E = m_i \tag{1c}$$

$$\hat{\pi} = \hat{\pi} P^n, \qquad n \geqq 0, \qquad (\hat{\pi} \text{ is a stationary measure}). \tag{1d}$$

For $c < \infty$ put

$$\mathscr{E}_c = \{A \subseteq E:\ \hat{\pi}_A \leqq c\}.$$

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* λ

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

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

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

Remark 1. (2) is the typical result obtained by the coupling method. With $\|\cdot\|$ denoting the total variation norm we have

$$\|\lambda P^n - \pi\| = 2 \sup_{A \subseteq E} (\lambda P_A^n - \pi_A) = 2 \sup_{A \subseteq E} (\pi_A - \lambda P_A^n) \tag{4}$$

and thus (2) can be rewritten on the form

$$\|\lambda P^n - \pi\| \to 0 \quad \text{as} \quad n \to \infty. \tag{2'}$$

This is maybe the more appropriate form but we have chosen (2) to stress the resemblance between (2) and (3). It should, however, be observed that (2) is, due to the countable state space, logically equivalent to the seemingly weaker classical result:

$$\mathbb{P}(X_n = j) \to \pi_j \quad \text{as} \quad n \to \infty.$$

Remark 2. (3) seems to be a new result improving somewhat the classical one: $\mathbb{P}(X_n = j) \to 0$ as $n \to \infty$. In Thorisson [8] (3) is extended to null-recurrent Harris chains. In Orey [4] the following equivalent result can be found (Theorem 1.7.3): For any $\varepsilon > 0$ it holds that $\mathbb{P}(X_n \in A)/(\varepsilon + \hat{\pi}_A) \to 0$ uniformly in A as $n \to \infty$.

2. Future Independent Stationary Times and Coupling

If T is a future independent μ time then for $k \leqq n$

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

and thus

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

$$\leqq \sum_{k=0}^{n} \mathbb{P}(T=k)\, \mu P_A^{n-k} + \mathbb{P}(T>n). \tag{5}$$

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

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

Applying (4) yields

$$\|\lambda P^n - \pi\| \leqq 2\,\mathbb{P}(T>n) \to 0 \quad \text{as} \quad n \to \infty \tag{6}$$

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 π 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 [6]) with T as a coupling epoch. Hence the inequality in (6) is a coupling inequality.

We have seen that a future independent stationary time can always be regarded as a coupling epoch, – the converse is obviously not true. What is more, while in the positive recurrent case there always exists a coupling epoch such that (6) holds (see [8] or combine (2) and the maximal coupling theorem in [6]) the same is not true for future independent stationary times as can be seen from the following counter-example:

Let X be positive recurrent and such that for each $n \geqq 0$ there is a j_n such that $P_{ij_n}^n = 0$ (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 a future independent stationary time we have by (i) and (ii)

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

But $\pi_j > 0$ for all j and thus $\mathbb{P}(T=n) = 0$ for all $n \geqq 0$ contradicting $T < \infty$ a.s.

3. Future Independent μ Times

The above counter-example shows that future independent μ 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 a future independent μ time T.*

Remark 3. If E is finite, we obtain (2) from Proposition 1 and (6) by putting $\mu = \pi = \hat{\pi}/m_i$.

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

$$\forall\, i, j \in E\, \exists\, n_{ij}\colon\ P_{ij}^n > 0 \qquad \text{for } n \geqq n_{ij}.$$

Fix an $i_0 \in E$ and put

$$n_0 = \max_{j \in B} n_{i_0 j} \quad \text{and} \quad \varepsilon = \min_{j \in B} P_{i_0 j}^{n_0} > 0.$$

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

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

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

$$\mathbb{P}(I_{kj} = 1) = \frac{\varepsilon \mu_j}{P_{i_0 j}^{n_0}}, \qquad j \in E, k \geqq 1.$$

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

$$(X_{T_k + n_0}, I_k) \text{ is independent of } (T_k, I_{k-1}, \ldots, I_1). \tag{7}$$

Further,

$$\mathbb{P}(X_{T_k + n_0} = j, I_k = 1) = \mathbb{P}(X_{T_k + n_0} = j)\, \mathbb{P}(I_{kj} = 1) = P_{i_0 j}^{n_0} \frac{\varepsilon \mu_j}{P_{i_0 j}^{n_0}} = \varepsilon \mu_j \tag{8}$$

and summing over $j \in E$ yields

$$\mathbb{P}(I_k = 1) = \varepsilon. \tag{9}$$

Put

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

Then $\{T = n\}$ is determined by $(X_k)_0^n$ and the I_{kj}'s and thus T is a randomized stopping time. Further, (7) and (8) yield the second equality in

$$\begin{aligned}
\mathbb{P}(X_T = j, T = n, K = k) &= \mathbb{P}(X_{T_k + n_0} = j, T_k + n_0 = n, I_k = 1, I_{k-1} = \ldots = I_1 = 0) \\
&= \varepsilon \mu_j\, \mathbb{P}(T_k + n_0 = n, I_{k-1} = \ldots = I_1 = 0) \\
&= \mu_j\, \mathbb{P}(T_k + n_0 = n, I_k = 1, I_{k-1} = \ldots = I_1 = 0) \\
&= \mu_j\, \mathbb{P}(T = n, K = k)
\end{aligned}$$

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

4. Proof of the Theorem

Take a finite $B \subseteq E$ and let μ be the "conditional stationary measure": $\mu_A = \hat{\pi}_{B \cap A} / \hat{\pi}_B$. Then $\mu \leqq \hat{\pi} / \hat{\pi}_B$ yields the inequality in

$$\mu P_A^{n-k} \leqq \frac{\hat{\pi} P_A^{n-k}}{\hat{\pi}_B} = \frac{\hat{\pi}_A}{\hat{\pi}_B}, \qquad A \subseteq E, k \leqq n,$$

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

$$\mathbb{P}(X_n \in A) \leqq \frac{\hat{\pi}_A}{\hat{\pi}_B} + \mathbb{P}(T>n).$$

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

$$\begin{aligned}\sup_{A\in\mathscr{E}_c}\left(\mathbb{P}(X_n\in A)-\frac{\hat{\pi}_A}{m_i}\right) &\leqq \sup_{A\in\mathscr{E}_c}\left(\frac{\hat{\pi}_A}{\hat{\pi}_B}-\frac{\hat{\pi}_A}{\hat{\pi}_E}\right)+\mathbb{P}(T>n)\\ &\leqq \frac{c}{\hat{\pi}_B}-\frac{c}{\hat{\pi}_E}+\mathbb{P}(T>n)\\ &\to \frac{c}{\hat{\pi}_B}-\frac{c}{\hat{\pi}_E} \quad \text{as} \quad n\to\infty\\ &\to 0 \quad \text{as} \quad B\uparrow E.\end{aligned}$$

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, π must be independent of i and thus with $i=j$, $\pi_j = 1/m_j$ due to (1b). Finally, if π' is a stationary distribution, then with $\lambda = \pi'$ we have $\lim_{n\to\infty} \mathbb{P}(X_n \in A) = \pi'_A$ and thus (3) cannot hold, i.e. X must be positive recurrent, – but then (2) holds implying $\pi'_A = \pi_A$ so π is unique and the proof is complete.

Remark 4. The approach of this paper extends easily to continuous time, irreducible, recurrent Markov jump processes. In fact the proof of Proposition 1 becomes more elementary in that case since the "well known result" refered to at the beginning of the proof is then immediate.

References

1. Aldous, D., Diaconis, P.: Shuffling cards and stopping times. Amer. Math. Mon. **93**, 333–348 (1986)
2. Aldous, D., Diaconis, P.: Strong uniform times and random walks. Technical Report, Department of Statistics, University of California, Berkeley (1986)
3. Asmussen, S.: Applied probability and queues. New York: Wiley 1987
4. Orey, S.: Lecture notes on limit theorems for Markov chain transition probabilities. London: Van Nostrand Reinhold 1971
5. Thorisson, H.: Coupling of regenerative processes. Adv. Appl. Probab. **15**, 531–561 (1983)
6. Thorisson, H.: On maximal and distributional coupling. Ann. Probab. **14**, 873–876 (1986)
7. Thorisson, H.: Backward limits. Ann. Probab. **16** (1988)
8. Thorisson, H.: Coupling of Markov chains revisited. (In preparation)

Received March 4, 1986; received in revised form November 12, 1987