The Annals of Mathematical Statistics
1972, Vol. 43, No. 2, 621-626

LIMIT THEOREMS FOR MARKOV TRANSITION FUNCTIONS[1]

By Richard Isaac

Herbert H. Lehman College and Graduate Center, CUNY

New L_1 and pointwise limit theorems are proved for the transition functions $P^n(x, E)$ of a Markov process with σ-finite invariant measure π satisfying a recurrence condition. Also given are related results about the operators on functions and measures induced by these transition functions. The method depends upon the application of martingale theorems, and the principal restriction concerns the structure of a certain σ-field.

1. Introduction. We study limiting behavior of stationary transition probabilities $P^n(x, E)$ for Markov processes with σ-finite stationary measure π, satisfying

(1.1) $\pi(E) > 0$ implies $P\{X_n \in E$ infinitely often $|X_0 = x\} = 1$ a.e. (π).

The background for our investigations is that of e.g. [5], [11]. Specifically, the following ideas are needed: The process $\{X_n, n \geqq 0\}$ is defined on a measurable space (Ω, Σ) and can be embedded in a process $\{X_n, -\infty < n < +\infty\}$. π and the transition probabilities induce a measure π_1 on bilateral sequence space Ω_1 consisting of the sample sequences $\omega = (\cdots, x_{-1}, x_0, x_1, \cdots)$, and π_1 is invariant relative to the $1 - 1$ invertible shift T. (1.1) implies T is conservative and ergodic relative to π_1. For each fixed n, π is the projection of π_1 on x_n-coordinate space.

The following condition (condition of T. E. Harris) has been considered in much recent work [11].

(1.1′) Same as (1.1), except that "a.e. (π)" is replaced by "for all $x \in \Omega$."

(1.1′) implies indecomposability of the process in the sense of Doeblin, and is a considerable strengthening of (1.1). Pointwise limit theorems (see Corollary 2.2) were proved for $P^n(x, E)$ under (1.1′) in [7].

In this paper we present pointwise limit theorems as well as certain L_1 limit theorems for $P^n(x, E)$. Related results obtained give information about the iterations μP^n and $P^n g$ on certain classes of measures and functions respectively (the operators P^n are defined in (2.6) and (2.15). Other than (1.1) or (1.1′), our conditions involve restrictions on the left tail σ-field $\tau_{-\infty}$ of the process (defined below). Since limiting behavior of $P^n(x, E)$ as $n \to +\infty$ is

Received March 26, 1971.

[1] Research supported by an NSF grant and a British Science Research Council grant at Westfield College, London.

determined by the structure of $\tau_{-\infty}$ it is suggestive to think of the asymptotic future determined by the asymptotic past.

Let $\tau_{-\infty} = \bigcap_{n \geqq 0} \mathscr{B}(\cdots, X_{-n-1}, X_{-n})$ where $\mathscr{B}(\cdot)$ is the σ-field generated by the given set of variables. Similarly $\tau_{+\infty} = \bigcap_{n \geqq 0} \mathscr{B}(X_n, X_{n+1}, \cdots)$. These are the left and right tail σ-fields of the process $\{X_n, -\infty < n < +\infty\}$. Under (1.1′) it is known, [8], that $\tau_{+\infty} = \tau_{-\infty} = \tau$ and τ is X_0-measurable and atomic, the atoms consisting of a finite number of cyclically moving classes. Although this fails to be true under (1.1), alternative behavior seems to be of a very limited nature (in a forthcoming article we hope to analyze these possibilities). For these reasons, (2.1) and (2.2) are more general than one might at first think, for if $\tau_{-\infty}$ is atomic and consists of a finite number of cyclically moving classes, behavior of $P^n(x, E)$ can be completely described by restriction to each atom and consideration of the related aperiodic process for which (2.1) or (2.2) holds.

The martingale method of this paper appears to be new in examining these questions; see also [6].

Let $\Delta_n = \mathscr{B}(\cdots, X_{-n-1}, X_{-n})$ and let f be integrable. Set $f_n = E(f \mid \Delta_n)$. Then $T\Delta_n = \Delta_{n-1}$ and a simple computation using the invariance of π_1 shows

$$T^k f_n = T^k E(f \mid \Delta_n) = E(T^k f \mid T^{-k}\Delta_n) = E(T^k f \mid \Delta_{n-k}) .$$

Moreover, setting $f = 1_{X_0 \in E}$, the indicator of $X_0 \in E$, $\pi(E) < \infty$, the Markov property yields

$$(1.2) \qquad T^n f_n = E(T^n f \mid \cdots, X_{-1}, X_0) = E(1_{X_n \in E} \mid X_0) = P^n(X_0, E) .$$

Property (1.2) suggests a study of the backward martingale f_n as a means of learning about $P^n(x, E)$. Observe that $\pi_1(\omega : T^n f_n(\omega) < y) = \pi(x : P^n(x, E) < y)$ for all real numbers y.

LEMMA 1.1. *For any integrable f, the sequence $f_n = E(f \mid \Delta_n)$ is a martingale with respect to the decreasing sequence of σ-fields Δ_n. If $\tau_{-\infty}$ is σ-finite,* $\lim_{n\to-\infty} f_n = E(f \mid \tau_{-\infty})$ a.e. (π_1). *If $A \in \tau_{-\infty}$ is such that $B \in \tau_{-\infty}$ and $B \subseteqq A$ implies $\tau_1(B) = 0$ or ∞, then* $\lim_{n\to-\infty} f_n = 0$ a.e. (π_1) *on A.*

PROOF. The martingale relation is an immediate consequence of the definition of f_n. For $\pi_1(\Omega_1) = 1$ the second assertion follows from the backward martingale convergence theorem [1]; if $\tau_{-\infty}$ is σ-finite, restrict f_n to each one of a countable number of sets of finite measure to which the martingale theorem just cited can be applied. The final assertion appears in [9].

2. Limit theorems. Now consider those processes satisfying (1.1) and one of the following conditions:

$$(2.1) \qquad \pi(\Omega) = 1 \quad \text{and} \quad \tau_{-\infty} \quad \text{is trivial, or}$$

$$(2.2) \qquad \pi(\Omega) = \infty \quad \text{and} \quad \tau_{-\infty} \quad \text{consists only of sets of measure } 0 \text{ or } \infty .$$

THEOREM 2.1. *Under the condition* (1.1) *and either* (2.1) *or* (2.2), *if E is a fixed set, then whatever subsequence $\{n_i\}$ is given, there is a further subsequence $\{m_i\}$ such that*

(2.3) *Under* (2.1), $P^{m_i}(x, E) \to \pi(E)$ a.e. (π) .

(2.4) *Under* (2.2), *if* $\pi(E) < \infty$, $P^{m_i}(x, E) \to 0$ a.e. (π) .

PROOF. *Proof of* (2.3). Let $\varepsilon > 0$; by Lemma 1.1, if $f = 1_{X_0 \in E}$, f_n converges and by ergodicity the limit must be $\pi(E)$ a.e. The sets $B_n = \{f_n > \pi(E) + \varepsilon\}$ satisfy $\pi_1(B_n) \to 0$, and $f_n - \pi(E)$ converges to zero in the mean. Then

$$
\begin{aligned}
(2.5)\quad 0 \leftarrow \int |f_n - \pi(E)|\, d\pi_1 &= \int T^n|f_n - \pi(E)|\, d\pi_1 \\
&= \int |T^n f_n - \pi(E)|\, d\pi_1 = \int |P^n(X_0, E) - \pi(E)|\, d\pi_1 \\
&= \int |P^n(x, E) - \pi(E)|\, d\pi ,
\end{aligned}
$$

so $P^n(x, E) \to \pi(E)$ in mean (π). For any subsequence $\{n_i\}$ there is then a further subsequence $\{m_i\}$ such that $P^{m_i}(x, E) \to \pi(E)$ a.e. (π).

Proof of (2.4). The sets $B_n = \{f_n > \varepsilon\}$ for $\varepsilon > 0$ fixed satisfy $\pi_1(B_n) \to 0$ as above; this follows from the relation

$$\varepsilon\pi_1(B_n) < \int_{B_n} f_n\, d\pi_1 = \int_{B_n} f\, d\pi_1$$

where the right side tends to zero because f is integrable and $\limsup_n B_n$ is π_1-null. Then f_n converges to zero in π_1 measure, and there is a subsequence $\{m_i\}$ of $\{n_i\}$ such that $\sum_{i=1}^{\infty} \pi_1(B_{m_i}) < \infty$, hence $\sum_{i=1}^{\infty} \pi_1(T^{-m_i}B_{m_i}) < \infty$. This means that the set $\limsup_i T^{-m_i}B_{m_i}$ is π_1 null; that is, $\limsup_i T^{m_i}f_{m_i} = \limsup_i P^{m_i}(X_0, E) \leqq \varepsilon$ a.e. (π_1), or $\limsup_i P^{m_i}(x, E) \leqq \varepsilon$ a.e. (π).
Since ε is arbitrary, a decreasing sequence ε_k converging to 0 may be chosen, and the argument may be repeated to yield a subsequence, say m_i, with $P^{m_i}(x, E) \to 0$ a.e. (π).

Define the measure

$$(2.6)\qquad \mu P^n = \mu_n = \int P^n(x, \cdot)\mu(dx)$$

where μ is a probability measure.

COROLLARY 2.1. *Let μ be any probability measure absolutely continuous with respect to π, and let E be fixed.*

(2.7) *Under* (1.1) *and* (2.1), $P^n(x, E) \to \pi(E)$ *in* $L_1(\mu)$, $\mu_n(E) \to \pi(E)$ *for all* E, *and for each* a.e. (π) *bounded* g, $\int g\, d\mu_n \to \int g\, d\pi$.

(2.8) *Under* (1.1) *and* (2.2), *if* $\pi(E) < \infty$, $P^n(x, E) \to 0$ *in* $L_1(\mu)$, *that is,* $\mu_n(E) \to 0$, *and for each* a.e. *bounded* g *in* $L_1(\pi)$, $\int g\, d\mu_n \to 0$.

PROOF. Proof of (2.7). (2.5) showed $L_1(\pi)$ convergence of $P^n(x, E)$ to $\pi(E)$, and there is convergence in π measure. Then, by finiteness and absolute continuity, there is convergence in μ measure and hence convergence in $L_1(\mu)$. Then

$$|\mu_n(E) - \pi(E)| \leqq \int |P^n(x, E) - \pi(E)|\mu(dx) \to 0 .$$

Now let g be a.e. bounded. $\int g\, d\phi$ is a continuous linear functional on the space of all signed measures with finite total variation and $\mu_n \to \pi$ in the sense of weak convergence of measures ([3] page 308) so that $\int g\, d\mu_n \to \int g\, d\pi$ (alternatively, there is a simple direct proof of this last assertion).

Proof of (2.8). $\pi_1(f_n > \varepsilon) \to 0$ for ε fixed, and then $\pi_1(T^n f_n > \varepsilon) \to 0$ or $\pi(P^n(x, E) > \varepsilon) \to 0$, so by finiteness of μ and absolute continuity, $\mu(P^n(x, E) > \varepsilon) \to 0$. $P^n(x, E)$ then converges to 0 in $L_1(\mu)$ since μ is finite, and

$$0 \leftarrow \int P^n(x, E)\mu(dx) = \mu_n(E) .$$

Now it is sufficient to consider positive g. Let $g \leqq L$ except on a π-null set N, and let g be in $L_1(\pi)$. Then $\mu(N) = 0$ and $P^n(x, N) = 0$ for all n a.e. (π) by stationarity so that $\mu_n(N) = 0$ for all n. Then

$$\int g\, d\mu_n = \int_A g\, d\mu_n + \int_B g\, d\mu_n \leqq \varepsilon + L\mu_n(B) \tag{2.9}$$

where $A = \{g \leqq \varepsilon\}$, $B = \{\varepsilon < g \leqq L\}$ and we have used the fact that $\mu_n(N) = 0$. Since $g \in L_1(\pi)$, $\pi(B) < \infty$ and $\mu_n(B) \to 0$ by the preceding part, hence (2.9) $\to 0$.

Now the question of almost everywhere convergence will be considered. We have found it necessary to impose (2.10) below, a restriction related to condition (D′) of Doeblin (see [7] page 210; also [2]). We have no example in which (2.1) or (2.2) holds but (2.10) does not hold, and so (2.10) may be superfluous, though necessary for the assertions of Theorem 2.2 to be valid.

According to [4], if E is any set, (1.1) implies $\limsup_n P^n(x, E) = c$, constant a.e. (π).

(2.10) For every decreasing sequence of sets A_i with $\pi(A_i) \downarrow 0$, $\lim_i \limsup_n P^n(x, A_i) = 0$ a.e. (π).

For each i, the lim sup is a.e. constant by the preceding remark; (2.10) requires that these constants converge to zero.

THEOREM 2.2.

(2.11) *Under* (1.1), (2.1), *and* (2.10), $P^n(x, E) \to \pi(E)$ a.e. (π).

(2.12) *Under* (1.1), (2.2), *and* (2.10), *if* $\pi(E) < \infty$, $P^n(x, E) \to 0$ a.e. (π).

PROOF. First we prove (2.12). Let $\varepsilon > 0$ and $B_n = [P^n(x, E) > \varepsilon]$. The proof of (2.8) shows $\pi(B_n) \to 0$, so there is a subsequence m_j such that the sets

$$A_{m_j} = B_{m_j} \cup B_{m_{j+1}} \cup \cdots$$

form a decreasing sequence with $\pi(A_{m_j}) \downarrow 0$. If $i \geqq j$, $P^{m_i}(x, E) \leqq \varepsilon$ on A'_{m_j}. Let i and j be fixed with $i \geqq j$ and let r be any integer. Then,

$$P^{r+m_i}(x, E) = (\textstyle\int_{A'_{m_j}} + \int_{A_{m_j}})P^r(x, dy)P^{m_i}(y, E) \leqq \varepsilon + P^r(x, A_{m_j}) ,$$

implying,

$$\limsup_n P^n(x, E) = \limsup_r P^{r+m_i}(x, E) \leqq \varepsilon + \limsup_r P^r(x, A_{m_j}) .$$

Using (2.10) and letting $j \to \infty$ shows $\limsup_n P^n(x, E) \leqq \varepsilon$ on a π-full set. Since ε is arbitrary, (2.12) readily follows. The proof of (2.11) is almost identical except for the obvious changes.

LEMMA 2.1. *Processes satisfying* (1.1′) *satisfy* (2.10).

PROOF. It follows from a result of Doeblin, [2], that processes satisfying (1.1′) have the property: If $\{A_i\}$ is a decreasing sequence with

$$\lim_i \limsup_n P^n(x, A_i) = \gamma \text{ a.e.}$$

then there is a subsequence $\{m_i\}$ and a set V, $0 < \pi(V)$, with

$$\liminf_i P^{m_i}(x, A_i) \geqq \gamma$$

for all $x \in V$ (see e.g. [7] p 212). Now let A_i be the decreasing sequence of (2.10). By Fatou's lemma,

$$0 = \lim_i \pi(A_i) \geqq \int \liminf P^{m_i}(x, A_i)\, d\pi \geqq \gamma\pi(V) , \qquad \text{implying } \gamma = 0 .$$

LEMMA 2.2. *Processes satisfying* (1.1′) *with* $\pi(\Omega) = \infty$ *satisfy* (2.2).

PROOF. (1.1′) implies the equality of $\tau_{-\infty}$ and $\tau_{+\infty}$ and that they are X_0-measurable and atomic with a finite number of atoms, each of infinite measure [8], [11].

COROLLARY 2.2. [7]. *For processes satisfying* (1.1′),

(2.13) *Under* (1.1′) *and* (2.1), $P^n(x, E) \to \pi(E)$, *for all* $x \in \Omega$.

(2.14) *Under* (1.1′), *if* $\pi(E) < \infty$ *and* $\pi(\Omega) = \infty$, $P^n(x, E) \to 0$ *for all* $x \in \Omega$.

PROOF. By Lemmas 2.1, 2.2, and Theorem 2.2, it follows that the assertions hold a.e. (π) for each E, say on a set $V = V(E)$. Since (1.1′) guarantees that eaɔh point will enter V and stay in it, a standard argument involving the Chapman-Kolmogorov equation completes the proof.

(2.15) Let $P^n g = \int P^n(\cdot, dy)g(y)$.

THEOREM 2.3. *Let* (1.1), (2.2) *and* (2.10) *be satisfied, and let g be* a.e. *bounded and in* $L_1(\pi)$. *Then*

(2.16) $$P^n g \to 0 \text{ a.e. } (\pi) .$$

PROOF. Let $g \leqq L$ except on a π-null set N. Since $P^n(x, N) = 0$ for all n a.e. (π), we may assume $g \leqq L$ everywhere in (2.17).
Let $A = \{|g| \leqq \varepsilon\}$, $B = \{\varepsilon < |g| \leqq L\}$; then

$$(2.17) \qquad |P^n g(x)| \leqq \textstyle\int_A + \int_B \{P^n(x, dy)|g(y)|\} \leqq \varepsilon + LP^n(x, B), \quad \text{and} \quad P^n(x, B) \to 0 \text{ a.e.}$$

by Theorem 2.2, concluding the proof.

Whether (2.16) holds for any g in $L_1(\pi)$ is an open question.

We conclude with an example of a process satisfying (1.1), (2.1) and (2.10) but not (1.1'). Let Z_n be a sequence of independent variables, $-\infty < n + \infty$, with common distribution: $P(Z_n = 0) = P(Z_n = 1) = \frac{1}{2}$. Let X_n on $\Omega = [0, 1]$ be given by the binary expansion: $X_n = \cdot Z_n Z_{n-1} Z_{n-2} \cdots$. X_n is Markov, $\pi =$ Lebesgue measure on Borel sets, (1.1) is satisfied, but not (1.1') since the orbit of each point is countable, and $\tau_{-\infty}$ is the tail σ-field of independent variables, so is trivial. It follows readily that each point spreads its conditional mass in so regular a way that the limit on each set is Lebesgue measure, and so (2.10) is satisfied.

Acknowledgment. The author is grateful to Professor S. J. Taylor and the Mathematics Department at Westfield College for their hospitality.

REFERENCES

[1] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.
[2] DOEBLIN, W. (1940). Elements d'une theorie generale des chaines simple constantes de Markoff. *Ann. Sci. Ecole Norm. Sup.* III, *Ser.* 57, 61-111.
[3] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators* **1**. Interscience, New York.
[4] ISAAC, R. (1966). On regular functions for certain Markov processes. *Proc. Amer. Math. Soc.* **17**, 1308-1313.
[5] ISAAC, R. (1968). Some topics in the theory of recurrent Markov processes. *Duke Math. J.* **35** 641-652.
[6] ISAAC, R. (1972). Theorems for conditional expectations, with applications to Markov processes. To appear.
[7] JAIN, N. C. (1966). Some limit theorems for a general Markov process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **6** 206-223.
[8] JAMISON, B. and OREY, S. (1967). Markov chains recurrent in the sense of Harris. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 41-48.
[9] JERISON, M. (1959). Martingale formulation of ergodic theorems. *Proc. Amer. Math. Soc.* **10** 531-539.
[10] MUNROE, M.E. (1953). *Introduction to Measure and Integration*. Addison-Wesley, Reading.
[11] OREY, S. (1968). Limit theorems for Markov chain transition probability functions. Mimeographed notes, Univ. of Minnesota.