The Annals of Probability
1973, Vol. 1, No. 6, 995-1013

ERGODIC BEHAVIOR FOR NONNEGATIVE KERNELS

BY RICHARD W. MADSEN AND PATRICIA S. CONN

University of Missouri, Columbia and Purdue University, Fort Wayne

Ergodic behavior for Markov chains can be determined by studying the properties of the corresponding sequence of stochastic transition kernels. Dobrushin's ergodic coefficient has been useful for this purpose. In this paper we define pointwise strongly and weakly ergodic behavior for sequences of nonnegative kernels and use Dobrushin's ergodic coefficient to give sufficient conditions for these two types of behavior. Applications are given to sequential probability ratio tests.

0. Summary. Two types of ergodic behavior for non-homogeneous Markov chains were considered by Paz [9] and Madsen [8]. In [8] the existence of stochatic transition density kernels $P_n(x, y)$ was assumed and the ergodic behavior was related to the ergodic coefficient defined by Dobrushin in [4]. In this paper we define two types of ergodic behavior for the case where the kernels are not stochastic, but only nonnegative. Further, by transforming nonnegative kernels into stochastic ones, we can use Dobrushin's ergodic coefficient to give sufficient conditions for these two types of ergodic behavior to hold. (See Sections 3 and 4 below).

Conn in [2] defined two types of ergodic behavior for nonnegative kernels in essentially the same way as in Definitions 1.1 and 1.2, but considered only kernels uniformly bounded above and below by positive numbers. Blum and Reichaw [1] generalized the use of the ergodic coefficient of Dobrushin to the case of nonnegative kernels rather than stochastic kernels. They made use of the ergodic coefficient to give conditions under which the sequence of superpositions of kernels is Cauchy.

In Section 5, we give some other sufficient conditions for ergodic behavior which may be easier to verify than those given in Sections 3 and 4. Section 6 gives applications of these results to ordinary and generalized sequential probability ratio tests.

1. Introduction. Throughout this paper we make various assumptions about the kernels under consideration. Let $(S, \mathscr{B}, \mu)$ be a σ-finite measure space and let $\{M_n\}$ be a sequence of nonnegative measurable kernels defined on $S \times S$. Assume that the kernels are sufficiently well behaved so that sequential superpositions defined by

$$(1.1) \qquad M_{m,m+n}(x, y) = \int_S \cdots \int_S M_m(x, z_1)M_{m+1}(z_1, z_2) \cdots M_{m+n}(z_n, y)\mu(dz_1) \cdots \mu(dz_n)$$

Received October 15, 1971; revised May 21, 1973.

AMS 1970 *subject classifications*. Primary 60J35, Secondary 62L10.

Key words and phrases. Ergodic behavior, sequence of nonnegative kernels, stochastic kernels, eigenfunctions.

exist for all x and $y \in S$. (For notational convenience, all integrals will be assumed to be over the whole space S unless otherwise indicated.) We also assume throughout that for nonnegative $f \in L_1(\mu)$, $g_n(y) = \int f(x)M_n(x, y)\mu(dx)$ is integrable for all n. Note that a sufficient condition for g_n to be integrable is that there exist a finite B_n such that

$$M_n(x, S) \equiv \int M_n(x, y)\mu(dy) \leqq B_n .$$

Further, if M_n is a stochastic kernel, the number 1 is such a bound. Finally we assume that $M_n(x, S) > 0$ for all n and all $x \in S$.

A non-homogeneous Markov chain is determined by a sequence $\{P_n(\cdot, \cdot)\}$ of stochastic transition kernels and an initial probability distribution over the state space S. Assuming a density f_0 exists for this probability distribution, the density function

$$f_n(y) = \int f_{n-1}(x)P_n(x, y)\mu(dx) = \int f_0(x)P_{1,n}(x, y)(dx)$$

gives the probability density over S at time n.

Two basic types of ergodic behavior can be described in terms of the sequence of functions $\{f_n\}$. If the effect of the initial probability distribution is lost as $n \to \infty$, the behavior is "weakly" ergodic. If, in addition, the sequence of functions $\{f_n\}$ converges as $n \to \infty$, the behavior is "strongly" ergodic. Various authors (for example [2], [4], [5], [8], and [9]) define ergodic behavior in different ways. We will consider pointwise behavior in ways made explicit in Definitions 1.1 and 1.2.

Let M_n be a sequence of kernels and let f_0 be a nonnegative function satisfying

$$0 < \int f_0(x)\mu(dx) < \infty . \tag{1.2}$$

Such functions f_0 will be called *starting functions* and the collection of all starting functions will be denoted by $\mathscr{S}$. For such functions we define, for $m < n$,

$$f_{m,n}(y) = \int f_0(x)M_{m,n}(x, y)\mu(dx)$$

and

$$f^*_{m,n}(y) = f_{m,n}(y)/\int f_{m,n}(y)\mu(dy) .$$

For notational convenience we will let $f_n \equiv f_{1,n}$.

Definition 1.1. A sequence of nonnegative kernels $\{M_n(x, y)\}$ will be called pointwise weakly ergodic (PWE) if for all m,

$$|f^*_{m,n}(y) - g^*_{m,n}(y)| \to_n 0$$

uniformly in $y \in S$ and $f_0, g_0 \in \mathscr{S} \times \mathscr{S}$.

Definition 1.2. A sequence of nonnegative kernels $\{M_n(x, y)\}$ will be called pointwise strongly ergodic (PSE) if there exists a function $q(y)$ such that for all m

$$|f^*_{m,n}(y) - q(y)| \to_n 0$$

uniformly in $y \in S$ and $f_0 \in \mathscr{S}$.

REMARK. The conditions are required to hold for all m so that PWE or PSE behavior does not come about because of the effect of one (or indeed a finite number) of kernels in the sequence. However, the proofs given in subsequent sections in no way depend on m, so without loss of generality we take $m = 1$.

2. Results for stochastic kernels. For stochastic kernels $P(x, y)$, the ergodic coefficient of Dobrushin [4] is defined as

$$\alpha(P) = 1 - \sup_{x,z,A\in\mathscr{B}} |\textstyle\int_A [P(x, y) - P(z, y)]\mu(dy)|$$

or equivalently

$$\begin{aligned}\alpha(P) &= 1 - \sup_{x,z} \textstyle\int [P(x, y) - P(z, y)]^+\mu(dy) \\ &= 1 - \tfrac{1}{2}\sup_{x,z} \textstyle\int |P(x, y) - P(z, y)|\mu(dy) .\end{aligned}$$

For convenience we often consider $\delta(P) \equiv 1 - \alpha(P)$.

We now state as lemmas some known results concerning $\delta(P)$. For further detail, the reader is referred to [4] or [9].

LEMMA 2.1. *If $P(x, y)$ is a stochastic kernel, then $0 \leqq \delta(P) \leqq 1$.*

LEMMA 2.2. *If P and Q are stochastic kernels and if PQ denotes the superposition of P and Q, then $\delta(PQ) \leqq \delta(P)\delta(Q)$.*

THEOREM 2.1. *Let $\{P_n\}$ be a sequence of stochastic kernels such that for all n, $0 \leqq P_n \leqq \Delta_n < \infty$. Then for any $m < n$,*

$$\sup_x P_{m,n}(x, y) - \inf_x P_{m,n}(x, y) \leqq \Delta_n \delta(P_{m,n-1}) .$$

PROOF. We have

$$\begin{aligned}&[\sup_x P_{m,n}(x, y) - \inf_x P_{m,n}(x, y)] \\ &\quad= \sup_{x,z} \{\textstyle\int [P_{m,n-1}(x, u)P_n(u, y) - P_{m,n-1}(z, u)P_n(u, y)]\mu(du)\} \\ &\quad= \sup_{x,z} \textstyle\int \{[P_{m,n-1}(x, u) - P_{m,n-1}(z, u)]^+ \\ &\qquad\quad - [P_{m,n-1}(x, u) - P_{m,n-1}(z, u)]^-\}P_n(u, y)\mu(du) \\ &\quad\leqq \sup_{x,z} \{\sup_u P_n(u, y) \textstyle\int [P_{m,n-1}(x, u) - P_{m,n-1}(z, u)]^+\mu(du) \\ &\qquad\quad - \inf_u P_n(u, y) \textstyle\int [P_{m,n-1}(x, u) - P_{m,n-1}(z, u)]^-\mu(du)\} \\ &\quad\leqq \sup_u P_n(u, y) \sup_{x,z} \textstyle\int [P_{m,n-1}(x, u) - P_{m,n-1}(z, u)]^+\mu(du) \\ &\quad\leqq \Delta_n \delta(P_{m,n-1}) . \qquad \Box\end{aligned}$$

COROLLARY 2.1. *If $\{P_n\}$ is a sequence of stochastic kernels satisfying $0 \leqq P_n \leqq \Delta$ and if for each n $\delta(P_n) \leqq \delta$, then for all $m < n$,*

$$\sup_x P_{m,n}(x, y) - \inf_x P_{m,n}(x, y) \leqq \Delta\delta^{n-m} .$$

PROOF. This follows from Lemma 2.2 and Theorem 2.1. □

Let P^n denote the n-fold superposition of P with itself. The next corollary shows that under certain conditions $P^n(x, y)$ converges, uniformly in x, to a function $g(y)$ which is a left eigenfunction of $P(x, y)$ corresponding to the eigenvalue 1. These results are not surprising in view of the well-known results in the matrix case as in Doob [5] or Feller [6], for example.

REMARK. Theorem 2.1 and Corollary 2.1 describe pointwise ergodic behavior for stochastic kernels and so are somewhat different than the results in [4] or [8] which describe ergodic behavior in the L_1 sense.

COROLLARY 2.2. *If P is a stochastic kernel satisfying $0 \leqq P \leqq \Delta$ and $\delta(P) < 1$, and if $\sup_x P^k(x, y)$ is an integrable function of y for some k, say k_0, then there exists a function $q(y)$ such that*

(a) $|P^n(x, y) - q(y)| \to_n 0$ *uniformly in x*

(b) $\int q(x)P(x, y)\mu(dx) = q(y)$.

PROOF. It follows from Corollary 2.1 that

$$\sup_x P^n(x, y) - \inf_x P^n(x, y) \leqq \Delta\delta(P)^{n-1} . \tag{2.1}$$

which tends to zero as $n \to \infty$.

Consider the sequence of functions $\{\sup_x P^n(x, \cdot)\}$. This is a non-increasing sequence since

$$\begin{aligned} \sup_x P^{n+1}(x, y) &= \sup_x \int P(x, z)P^n(z, y)\mu(dz) \\ &\leqq \sup_x [\sup_z P^n(z, y) \int P(x, z)\mu(dz)] \\ &= \sup_z P^n(z, y) . \end{aligned}$$

Hence for each y, $\{\sup_x P^n(x, y)\}$ is a non-increasing sequence which is bounded below. Thus the pointwise limit, call it q, exists. Similarly, for each y, $\{\inf_x P^n(x, y)\}$ is a non-decreasing sequence which is bounded above by the number $q(y)$, hence its pointwise limit also exists. In fact, in view of (2.1) $q(y)$ must be the common limit.

Now (a) holds since

$$|P^n(x, y) - q(y)| \leqq \sup_x P^n(x, y) - \inf_x P^n(x, y) \leqq \Delta\delta(P)^{n-1} .$$

To prove (b), consider for $n > k_0$

$$P^{n+1}(x, y) = \int P^n(x, z)P(z, y)\mu(dz)$$

and apply Lebesgue's dominated convergence theorem. Then by (a) formula (b) holds. □

REMARK. If $\mu(S) < \infty$, then the condition $P(x, y) \leqq \Delta$ implies the integrability of $\sup_x P(x, y)$. Also, under the hypotheses of Corollary 2.2, it is possible to show that a homogeneous Markov chain with transition kernel $P(x, y)$ is strongly ergodic in the sense defined in [8]. In view of this, it follows that q integrates to 1.

3. Weak ergodicity. In Section 2 μ is assumed to be σ-finite. We now assume $\mu(S) < \infty$. The following conditions are listed for convenient referencing in what follows.

CONDITION 1.

$$0 \leqq M_n(x, y) \leqq \Delta . \tag{3.1}$$

CONDITION 2.

$$M_n(x, S) \geqq v > 0 \qquad \text{for all } x \tag{3.2a}$$

$$M_n(S, y) \geqq v > 0 . \qquad \text{for all } y \tag{3.2b}$$

CONDITION 3. For each kernel M_n there is an eigenvalue λ_n with corresponding right and left eigenfunctions Φ_n and ψ_n satisfying

$$0 < b \leqq \Phi_n(x) \leqq B < \infty \tag{3.3}$$

$$0 < d \leqq \psi_n(y) \leqq D < \infty . \tag{3.4}$$

CONDITION 4. The sequence $\{P_n\}$ defined by

$$P_n(x, y) = M_n(x, y)\Phi_n(y)/\lambda_n \Phi_n(x) \tag{3.5}$$

satisfies for some δ,

$$\delta(P_n) \leqq \delta < 1 . \tag{3.6}$$

CONDITION 5. The sequence $\{Q_n\}$ defined by

$$Q_n(x, y) = \psi_n(y)M_n(y, x)/\lambda_n \psi_n(x) \tag{3.7}$$

satisfies for some δ,

$$\delta(Q_n) \leqq \delta < 1 . \tag{3.8}$$

REMARK. It follows from Conn's work [2] that if Condition 1 is replaced by

CONDITION 1^1.

$$0 < \Delta_l \leqq M_n \leqq \Delta_u < \infty \tag{3.9}$$

then all of Conditions 2 through 5 are satisfied. Harris [7] showed that for primitive kernels the eigenvalue λ_n described in Condition 3 is simple and that both (3.3) and (3.4) are satisfied.

LEMMA 3.1.

(a) (3.1) *implies* $M_{m,m+n} \leqq \Delta^{n+1}\mu(S)^n$

(b) (3.1) *and* (3.2a) *imply that for* f_0 *a starting function*

$$0 < \int f_n(x)\mu(dx) < \infty \quad \text{and} \quad f_n^*(y) = f_n(y)/\int f_n(x)\mu(dx) \leqq \Delta/v \tag{3.10}$$

(c) (3.1), (3.2a), *and* (3.4) *or* (3.1), (3.2b), *and* (3.3) *imply*

$$0 < v \leqq \lambda_n \leqq \Delta\mu(S) \tag{3.11}$$

(d) (3.1), (3.2a), *and* (3.4) *imply*

$$Q_n \leqq \Delta D/vd \tag{3.12}$$

(e) (3.1), (3.2a), (3.3), *and* (3.4) *imply* $P_n \leqq \Delta B/vb$.

PROOF. (a) follows easily by induction. If we note that

$$\int f_n(y)\mu(dy) = \int\int f_{n-1}(x)M_n(x, y)\mu(dx)\mu(dy) = \int f_{n-1}(x)[\int M_n(x, y)\mu(dy)]\mu(dx) \geqq v \int f_{n-1}(x)\mu(dx) \tag{3.13}$$

and

$$f_n(y) = \int f_{n-1}(x)M_n(x, y)\mu(dx) \leqq \Delta \int f_{n-1}(x)\mu(dx)\,, \tag{3.14}$$

then (b) and (c) can easily be proven.

Using (3.11), (d) and (e) also can be easily proven. □

A sequence of kernels satisfying (3.1), (3.2a), (3.3), (3.6) and either (3.2b) or (3.4) is said to satisfy *Condition* W.

THEOREM 3.1. *Let M_n be a sequence of kernels satisfying Condition* W. *If*

$$\int |\Phi_n(x) - \Phi_{n+1}(x)|\mu(dx) \to_n 0 \tag{3.15}$$

then there exist a sequence of functions $\{q_n\}$, independent of f_0, such that

$$|f_n{}^*(x) - q_n(x)| \to_n 0 \tag{3.16}$$

uniformly in x and f_0.

LEMMA 3.2. *Under the conditions of Theorem* 3.1, *for all k*,

$$\int \left| \frac{M_{n+1,n+k}(x, y)\Phi_{n+k}(y)}{\Lambda(n, k)\Phi_{n+k}(x)} - P_{n+1,n+k}(x, y)\right| \mu(dx) \to_n 0$$

where we write $\Lambda(n, k) = \prod_{i=n+1}^{n+k} \lambda_i$.

PROOF. When $k = 1$ the result holds since by (3.5) the integrand is zero. We proceed by induction. Assume the result is true for k and consider

$$\begin{aligned} &\int \left| \frac{M_{n+1,n+k+1}(x, y)\Phi_{n+k+1}(y)}{\Lambda(n, k+1)\Phi_{n+k+1}(x)} - P_{n+1,n+k+1}(x, y)\right| \mu(dx) \\ &= \int \left| \int \left\{ \frac{M_{n+1,n+k}(x, z)}{\Lambda(n, k)} \left[\frac{\Phi_{n+k+1}(z)}{\Phi_{n+k+1}(x)} - \frac{\Phi_{n+k}(z)}{\Phi_{n+k}(x)} \right] \right.\right. \\ &\qquad \left.\left. + \left[\frac{M_{n+1,n+k}(x, z)\Phi_{n+k}(z)}{\Lambda(n, k)\Phi_{n+k}(x)} - P_{n+1,n+k}(x, z) \right] \right\} P_{n+k+1}(z, y)\mu(dz) \right| \mu(dx) \\ &\leqq \frac{\Delta'\Delta^k\mu(S)^{k-1}}{v^k} \int\int \left| \frac{\Phi_{n+k+1}(z)}{\Phi_{n+k+1}(x)} - \frac{\Phi_{n+k}(z)}{\Phi_{n+k}(x)} \right| \mu(dz)\mu(dx) \\ &\qquad + \Delta' \int\int \left| \frac{M_{n+1,n+k}(x, z)\Phi_{n+k}(z)}{\Lambda(n, k)\Phi_{n+k}(x)} - P_{n+1,n+k}(x, z) \right| \mu(dz)\mu(dx)\,, \end{aligned} \tag{3.17}$$

where $\Delta' = \Delta B/vb$.

Considering the first term of (3.17) and using the bounds of (3.3), it can be shown that

$$\left| \frac{\Phi_{n+k+1}(z)}{\Phi_{n+k+1}(x)} - \frac{\Phi_{n+k}(z)}{\Phi_{n+k}(x)} \right| \leqq \frac{1}{b} |\Phi_{n+k+1}(z) - \Phi_{n+k}(z)| + \frac{B}{b^2} |\Phi_{n+k}(x) - \Phi_{n+k+1}(x)|\,.$$

Since the space is of finite measure, it follows from (3.15) that the first term goes to zero.

In the second term of (3.17) the integrand is bounded independently of n. Since $\mu(S) < \infty$, the bounded convergence theorem and the induction hypothesis imply that the second term goes to zero. □

LEMMA 3.3. *Under the conditions of Theorem* 3.1, *given* $\varepsilon > 0$, *there exists a sequence of functions* $\{t_n\}$ *such that for all* $n \geqq N(\varepsilon)$, $|f_n^* - t_n| < \varepsilon$, *independently of the choice of* f_0.

PROOF. Define $s_n^k(y) = \sup_x P_{n+1,n+k}(x, y)$. Since P_n is bounded for all n by Δ' and since $\delta(P_n) \leqq \delta < 1$, Corollary 2.1 implies that given $\gamma > 0$, if $k > N(\gamma)$,

$$|s_n^k(y) - P_{n+1,n+k}(x, y)| \leqq \Delta\delta^{k-1} < \gamma \,. \tag{3.18}$$

Define $r_n^k(y) = s_n^k(y)/\Phi_{n+k}(y)$ and $R_n^k = \int r_n^k(y)\mu(dy)$. We will show that for suitable k_0, the appropriate sequence of functions $\{t_n(y)\}$ can be defined by

$$t_{n+k_0}(y) = r_n^{k_0}(y)/R_n^{k_0} \,.$$

Let $(\Phi, f) = \int \Phi(x)f(x)\mu(dx)$. For any k,

$$\begin{aligned} |f_{n+k}^*(y) - t_{n+k}(y)| &\leqq \left| f_{n+k}^*(y) - \frac{f_{n+k}^*(y) \int f_{n+k}(y)\mu(dy)}{\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n)} \right| \\ &\quad + \left| \frac{f_{n+k}^*(y) \int f_{n+k}(y)\mu(dy)}{\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n)} - t_{n+k}(y) \right| \\ &\leqq \sup_y f_{n+k}^*(y) \left| 1 - \frac{\int f_{n+k}(y)\mu(dy)}{\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n)} \right| \\ &\quad + \left| \frac{f_{n+k}(y)}{\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n)} - t_{n+k}(y) \right| \\ &= \sup_y f_{n+k}^*(y) \left| \int A_n^k(y)\mu(dy) \right| + |A_n^k(y)| \,, \end{aligned}$$

where $A_n^k(y) = \{f_{n+k}(y)/\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n)\} - t_{n+k}(y)$. In view of (3.10) and since $\mu(S) < \infty$, it suffices to show that $|A_n^k(y)|$ can be made uniformly small for some fixed k and n sufficiently large. We begin by finding lower bounds for R_n^k and (Φ_{n+k}, f_n^*):

$$\begin{aligned} R_n^k &= \int [s_n^k(y)/\Phi_{n+k}(y)]\mu(dy) \geqq (1/B) \int \sup_x P_{n+1,n+k}(x, y)\mu(dy) \\ &\geqq (1/B) \int P_{n+1,n+k}(x_0, y)\mu(dy) = 1/B \end{aligned}$$

where x_0 is any point in S.

$$(\Phi_{n+k}, f_n^*) = \int \Phi_{n+k}(x)f_n^*(x)\mu(dx) \geqq b \int f_n^*(x)\mu(dx) = b \,.$$

$$\begin{aligned} |A_n^k(y)| &= \left| \frac{f_{n+k}(y)}{\Lambda(n, k)R_n^k(\Phi_{n+k}, f_n^*) \int f_n(y)\mu(dy)} - \frac{r_n^k(y)(\Phi_{n+k}, f_n^*)}{R_n^k(\Phi_{n+k}, f_n^*)} \right| \\ &\leqq \frac{B}{b} \left| \frac{f_{n+k}(y)}{\Lambda(n, k) \int f_n(y)\mu(dy)} - r_n^k(y)(\Phi_{n+k}, f_n^*) \right| \end{aligned}$$

$$= \frac{B}{b}\left|\frac{\int f_n(x)M_{n+1,n+k}(x,y)\mu(dx)}{\Lambda(n,k)\int f_n(y)\mu(dy)} - \int r_n{}^k(y)\Phi_{n+k}(x)f_n{}^*(x)\mu(dx)\right|$$

$$\leqq \frac{B}{b}\int f_n{}^*(x)\left|\frac{M_{n+1,n+k}(x,y)}{\Lambda(n,k)} - \frac{s_n{}^k(y)\Phi_{n+k}(x)}{\Phi_{n+k}(y)}\right|\mu(dx)$$

$$\leqq \frac{B}{b}\sup_x f_n{}^*(x)\int \frac{\Phi_{n+k}(x)}{\Phi_{n+k}(y)}\left\{\left|\frac{M_{n+1,n+k}(x,y)\Phi_{n+k}(y)}{\Lambda(n,k)\Phi_{n+k}(x)} - P_{n+1,n+k}(x,y)\right|\right.$$

$$\left. + |P_{n+1,n+k}(x,y) - s_n{}^k(y)|\right\}\mu(dx)$$

$$\leqq \frac{B^2\Delta}{b^2v}\left\{\int\left|\frac{M_{n+1,n+k}(x,y)\Phi_{n+k}(y)}{\Lambda(n,k)\Phi_{n+k}(x)} - P_{n+1,n+k}(x,y)\right|\mu(dx)\right.$$

$$\left. + \int |P_{n+1,n+k}(x,y) - s_n{}^k(y)|\mu(dx)\right\},$$

where the last inequality holds from (3.3) and (3.10). From (3.18) it is clear that given $\gamma > 0$, there exist a k_0 such that $\Delta'\delta^{k_0-1}\mu(S) < \gamma$, hence

$$\int |P_{n+1,n+k_0}(x,y) - s_n{}^{k_0}(y)|\mu(dx) < \Delta'\delta^{k_0-1}\mu(S) < \gamma .$$

Further, from Lemma 3.2, given k_0, there is some $N = N(k_0, \gamma)$ such that for $n \geqq N(k_0, \gamma)$,

$$\int\left|\frac{M_{n+1,n+k_0}(x,y)\Phi_{n+k_0}(y)}{\Lambda(n,k_0)\Phi_{n+k_0}(x)} - P_{n+1,n+k_0}(x,y)\right|\mu(dx) < \gamma .$$

Hence

$$|A_n{}^{k_0}| \leqq 2B^2\Delta\gamma/b^2v \equiv c\gamma . \tag{3.19}$$

Since γ can be chosen arbitrarily small and since (3.19) holds for all y, the proof follows. □

Proof of Theorem 3.1. Let $\{\varepsilon_j\}$ be a sequence of numbers decreasing to zero. From Lemma 3.3, there are (increasing) sequences $\{k_j\}$ and $\{N(k_j, \varepsilon_j)\}$ and a sequence of functions $\{t_n{}^{(j)}\}$ such that for $n \geqq N_j = N(k_j, \varepsilon_j)$,

$$|f^*_{n+k_j}(y) - t^{(j)}_{n+k_j}(y)| < \varepsilon_j .$$

By construction, the following sequence satisfies the conclusion of Theorem 3.1.

$$\begin{aligned} q_n(y) &= 1/\mu(S) && n = 1, 2, \cdots, N_1 + k_1 \\ &= t_n{}^{(j)}(y) && n = N_j + k_j + 1, \cdots, N_{j+1} + k_{j+1}, j = 1, 2, \cdots . \end{aligned}$$ □

Remark. It is clear that the existence of functions $\{q_n\}$ satisfying (3.16) is equivalent to the condition in Definition 1.1.

Note that if the sequence of kernels under consideration is stochastic, then $\Phi_n(x) = 1$ is a positive bounded right eigenfunction for all n, hence (3.15) holds trivially. Also $\int M_n(x,y)\mu(dy) = 1$ for all n implies that (3.2a) is satisfied. Hence to apply Theorem 3.1, only conditions (3.1) and (3.6) need be verified.

4. Strong ergodicity. We first prove some results relating to PWE behavior and from these results we obtain sufficient conditions for PSE behavior.

If $m > n$, the notation $M_{m,n}$ will mean the "reverse order" superposition of kernels, $M_m, M_{m-1}, \cdots, M_n$. When positive left eigenfunctions exist for a nonnegative kernel M, define

$$R(x, y) = \psi(x)M(x, y)/\lambda\psi(y) .$$

Thus, $Q(x, y) = R(y, x)$, and

$$R_{m,n}(x, y) = Q_{n,m}(y, x) . \tag{4.1}$$

A sequence of kernels satisfying (3.1), (3.2a), (3.4), and (3.8) is said to satisfy *Condition* S.

Since ψ is assumed to be bounded (condition (3.4)) it follows that it is integrable. Hence, we can take ψ to integrate to 1.

The following lemma is needed in proving Theorem 4.1.

LEMMA 4.1. *Let* $\{M_n\}$ *be a sequence of kernels satisfying Condition* S.

(a) *If* $\{\psi_n - \psi_{n+1}\}$ *converges to zero in* $L_1(\mu)$ *then so does*

$$\frac{M_{n+1,n+k}(\cdot, y)\psi_{n+k}}{\Lambda(n, k)\psi_{n+k}(y)} - R_{n+1,n+k}(\cdot, y)$$

for every k *and* y.

(b) *If*

$$s_n^k(x) = \sup_y Q_{n+k,n+1}(y, x) \tag{4.2}$$

then

$$s_n^k(x) \geqq \left[\frac{vd}{D\Delta\mu(S)}\right]^k \frac{1}{\mu(S)} > 0 .$$

PROOF. (a) is proved similarly to Lemma 3.2. To prove (b) use inequalities (3.4) and (3.11) to get

$$\begin{aligned} s_n^k(x) &= \sup_y \int \cdots \int Q_{n+k}(y, z_k) \cdots Q_{n+1}(z_2, x)\mu(dz_k) \cdots \mu(dz_2) \\ &= \sup_y \int \cdots \int \frac{M_{n+k}(z_k, y) \cdots M_{n+1}(x, z_2)}{\Lambda(n, k)} \\ &\qquad \times \left[\frac{\psi_{n+k}(z_k)}{\psi_{n+k}(y)} \cdots \frac{\psi_{n+1}(x)}{\psi_{n+1}(z_2)}\right]\mu(dz_k) \cdots \mu(dz_2) \\ &\geqq \left(\frac{d}{D\Delta\mu(S)}\right)^k \sup_y M_{n+1,n+k}(x, y) . \end{aligned}$$

It therefore suffices to show that $\sup_y M_{n+1,n+k}(x, y) \geqq v^k/\mu(S)$. But this follows as a consequence of $M_{n+1,n+k}(x, S) \geqq v^k$ which in turn follows from (3.2) and Fubini's theorem. □

THEOREM 4.1. *Let* $\{M_n\}$ *be a sequence of kernels satisfying Condition* S. *If* $\int |\psi_n(x) - \psi_{n+1}(x)|\mu(dx) \to_n 0$, *then* $|f_n^*(x) - \psi_n(x)| \to_n 0$, *uniformly in* x *and* $f_0 \in \mathscr{S}$.

PROOF. Define $t_n^k(x) = s_n^k(x)f_n(x)/\psi_{n+k}(x)$ and $T_n^k = \int t_n^k(x)\mu(dx)$. Then

$$\begin{aligned} &|f_{n+k}^*(x) - \psi_{n+k}(x)| \\ &\qquad \leqq \left|f_{n+k}^*(x) - \frac{f_{n+k}^*(x)\int f_{n+k}(x)\mu(dx)}{\Lambda(n, k)T_n^k}\right| + \left|\frac{f_{n+k}(x)}{\Lambda(n, k)T_n^k} - \psi_{n+k}(x)\right| . \end{aligned} \tag{4.3}$$

Since $\phi_{n+k}(x)$ integrates to 1, it follows that (4.3) is

$$\leqq \sup_x f^*_{n+k}(x)|\int A_n^k(x)\mu(dx)| + |A_n^k(x)|$$

where $A_n^k(x) = f_{n+k}(x)/\Lambda(n, k)T_n^k - \phi_{n+k}(x)$. It suffices to show that for some k_0, $|A_n^{k_0}(x)|$ will be less than some preassigned ε if n is sufficiently large. By (b) of Lemma 4.1, we have

$$\int \frac{s_n^k(y)f_n^*(y)}{\phi_{n+k}(y)}\mu(dy) \geqq \frac{1}{D}\int s_k^n(y)f_n^*(y)\mu(dy) \geqq \left[\frac{dv}{D\Delta\mu(S)}\right]^k \frac{1}{D\mu(S)} = c_1 .$$

Thus,

$$|A_n^k(x)| = \left|\frac{f_{n+k}(x)\int f_n(y)\mu(dy)}{\Lambda(n, k)T_n^k\int f_n(y)\mu(dy)} - \frac{T_n^k\int f_n(y)\mu(dy)\phi_{n+k}(x)}{T_n^k\int f_n(y)\mu(dy)}\right|$$

$$\left[\int \frac{s_n^k(y)f_n^*(y)}{\phi_{n+k}(y)}\mu(dy)\right]^{-1}\left|\frac{f_{n+k}(x)}{\Lambda(n, k)\int f_n(y)\mu(dy)}\right.$$

$$\left. - \int \frac{s_n^k(y)f_n^*(y)\phi_{n+k}(x)}{\phi_{n+k}(y)}\mu(dy)\right| \tag{4.4}$$

$$\leqq \frac{1}{c_1}\left|\int \frac{f_n^*(y)M_{n+1,n+k}(y, x)}{\Lambda(n, k)}\mu(dy) - \int f_n^*(y)s_n^k(y)\frac{\phi_{n+k}(x)}{\phi_{n+k}(y)}\mu(dy)\right|$$

$$= \frac{1}{c_1}\left|\int f_n^*(y)\frac{\phi_{n+k}(x)}{\phi_{n+k}(y)}\left[\frac{M_{n+1,n+k}(y, x)\phi_{n+k}(y)}{\Lambda(n, k)\phi_{n+k}(x)} - s_n^k(y)\right]\mu(dy)\right| .$$

Using inequalities (3.4) and (3.10), it follows that (4.4) is

$$\leqq \frac{\Delta D}{dvc_1}\left\{\int \left|\frac{M_{n+1,n+k}(y, x)\phi_{n+k}(y)}{\Lambda(n, k)\phi_{n+k}(x)} - R_{n+1,n+k}(y, x)\right|\mu(dy)\right.$$

$$\left. + \int |R_{n+1,n+k}(y, x) - s_n^k(y)|\mu(dy)\right\} . \tag{4.5}$$

It follows from (4.1) and (4.2) that the second term of (4.5) is

$$\int |Q_{n+k,n+1}(x, y) - \sup_x Q_{n+k,n+1}(x, y)|\mu(dy) .$$

Using the bound $\Delta' = \Delta D/vd$ from (3.12) and using Corollary 2.1, it follows that given $\gamma > 0$, there exists k_0 such that

$$|Q_{n+k_0,n+1}(x, y) - \sup_x Q_{n+k_0,n+1}(x, y)| \leqq \Delta'\delta^{k_0-1} < \gamma/\mu(S) .$$

Then for every x,

$$\int |Q_{n+k_0,n+1}(x, y) - \sup_x Q_{n+k_0,n+1}(x, y)|\mu(dy) < \Delta'\delta^{k_0-1}\mu(S) < \gamma . \tag{4.6}$$

By Lemma 4.1, given k_0 and γ, there exists an $N = N(k_0, \gamma)$ such that for $n \geqq N$,

$$\int \left|\frac{M_{n+1,n+k_0}(y, x)\phi_{n+k_0}(y)}{\Lambda(n, k_0)\phi_{n+k_0}(x)} - R_{n+1,n+k_0}(y, x)\right|\mu(dy) < \gamma . \tag{4.7}$$

Combining (4.6) and (4.7), we have that for $n \geqq N(k_0, \gamma)$, $|A_n^{k_0}(x)| \leqq (2\Delta D/dc_1v)\gamma$ as was to be shown. Note that in this proof, k_0 and $N(k_0, \varepsilon)$ are chosen independently of f_0 and x. □

COROLLARY 4.1. *Let $\{M_n\}$ be a sequence of kernels satisfying Condition* S. *If for each x, $|\psi_n(x) - \psi_{n+1}(x)| \to_n 0$, then for every starting function $f_0(x)$, $|f_n^*(x) - \psi_n(x)| \to_n 0$* uniformly.

PROOF. Follows from the bounded convergence theorem. □

THEOREM 4.2. *Let $\{M_n\}$ be a sequence of kernels satisfying Condition* S. *If the sequence of left eigenfunctions $\{\psi_n\}$ converges uniformly to ψ, then $\{M_n\}$ is pointwise strongly ergodic.*

PROOF. Using

$$|f_n^*(x) - \psi(x)| \leqq |f_n^*(x) - \psi_n(x)| + |\psi_n(x) - \psi(x)| ,$$

and Corollary 4.1, the proof is straight forward. □

REMARK. In view of Corollary 4.1, under Condition S,

$$(4.8) \qquad |\psi_n(x) - \psi_{n+1}(x)| \to_n 0 \qquad \text{implies} \quad |f_n^*(x) - f_{n+1}^*(x)| \to_n 0 .$$

Under some additional conditions, it is possible to show that the converse of (4.8) holds.

A sequence of kernels satisfying Condition S and such that the eigenvalue λ_n (corresponding to eigenfunctions $\psi_n(x)$ and $\Phi_n(y)$) is simple and such that $\Phi_n(y) \geqq b > 0$ is integrable is said to satisfy Condition S′.

Note that all the conditions for Condition S′ are satisfied for kernels which satisfy (3.9). Further the conditions that λ_n be simple and that $\Phi_n(x)$ be integrable are satisfied for primitive kernels according to Harris [7].

LEMMA 4.2. *Let $\{M_n\}$ be a sequence of kernels satisfying Condition* S′. *For a given starting function f_0, define*

$$\rho_n = \int f_{n-1}(y)\mu(dy) / \int f_n(y)\mu(dy) .$$

If $\int |f_n^*(y) - f_{n+1}^*(y)|\mu(dy) \to_n 0$, then for all k,

$$|f_n^*(y) - \rho_n^k \int f_n^*(x) M_n^k(x, y)\mu(dx)| \to_n 0 ,$$

uniformly in y.

PROOF. Using (3.13) and (3.14), it is easy to see that ρ_n satisfies

$$(4.9) \qquad 1/\Delta\mu(S) \leqq \rho_n \leqq 1/v .$$

Now proceed by induction. When $k = 1$

$$\begin{aligned} &|f_n^*(y) - \rho_n \int f_n^*(x) M_n(x, y)\mu(dx)| \\ &\qquad = \left| \int \frac{f_{n-1}(x) M_n(x, y)}{\int f_n(y)\mu(dy)} \mu(dx) - \rho_n \int f_n^*(x) M_n(x, y)\mu(dx) \right| \\ &\qquad \leqq \rho_n \Delta \int |f_{n-1}^*(x) - f_n^*(x)|\mu(dx) \\ &\qquad \leqq (\Delta/v) \int |f_{n-1}^*(x) - f_n^*(x)|\mu(dx) . \end{aligned}$$

This last expression tends to zero by the hypothesis, further it gets small independently of the choice of y.

Now assume the result holds for k and consider

$$
\begin{aligned}
&|\rho_n^{k+1} \textstyle\int f_n^*(x) M_n^{k+1}(x, y)\mu(dx) - f_n^*(y)| \\
&\quad = |\rho_n \textstyle\int \rho_n^k f_n^*(x)[\int M_n^k(x, z) M_n(z, y)\mu(dz)]\mu(dx) \\
&\qquad - \rho_n \textstyle\int f_{n-1}^*(z) M_n(z, y)\mu(dz)| \\
&\quad = \rho_n | \textstyle\int [\int \rho_n^k f_n^*(x) M_n^k(x, z)\mu(dx) - f_n^*(z) \\
&\qquad + f_n^*(z) - f_{n-1}^*(z)] M_n(z, y)\mu(dz)| \\
&\quad \leqq \frac{\Delta}{v} \{ \textstyle\int | \int \rho_n^k f_n^*(x) M_n^k(x, z)\mu(dx) - f_n^*(z)|\mu(dz) \\
&\qquad + \textstyle\int |f_n^*(z) - f_{n-1}^*(z)|\mu(dz)\} .
\end{aligned}
\tag{4.10}
$$

By the induction hypothesis, the integrand of the first term of (4.10) goes to zero uniformly in z. Since $\mu(S) < \infty$, the first term goes to zero. The second term of (4.10) goes to zero by the hypothesis of the lemma. □

The following lemma considers iterates of single kernel, i.e., the stationary case.

LEMMA 4.3. *Let M be a kernel which satisfies Condition* S′. *Then*

$$\left|\frac{M^k(x, y)}{\lambda^k} - \Phi(x)\phi(y)\right| \leqq \frac{D}{d}\,\Delta'\delta^{k-1}$$

where $\Delta' = D\Delta/dv$ and $0 \leqq \delta < 1$.

PROOF. Since $\Phi(x)$ is assumed integrable and $\phi(x)$ is bounded, these functions can be chosen in such a way that $\int \Phi(x)\phi(x)\mu(dx) = 1$. From (3.12), $Q(x, y)$ is bounded by $\Delta' = D\Delta/dv$, and by assumption (3.8), $\delta(Q) = \delta < 1$. Hence Corollaries 2.1 and 2.2 can be applied, so we know that

$$\sup_x Q^k(x, y) - \inf_x Q^k(x, y) \leqq \Delta'\delta^{k-1} \tag{4.11}$$

and that there exists $q(y) = \lim_{k\to\infty} \sup_x Q^k(x, y) = \lim_{k\to\infty} \inf_x Q^k(x, y)$ which is a left eigenfunction of $Q(x, y)$ corresponding to the eigenvalue 1. In view of our remark after Corollary 2.2, we know that $q(y)$ integrates to 1.

Now $\Phi(x)$ is a right eigenfunction corresponding to the simple eigenvalue λ, and it is easy to show that $q(x)/\phi(x)$ is also a right eigenfunction for the same eigenvalue. However, since λ is simple, it must be that $\Phi(x) = q(x)/\phi(x)$, since both $q(x)$ and $\Phi(x)\phi(x)$ integrate to 1.

It is easy to see that $Q^k(y, x) = \phi(x)M^k(x, y)/\lambda^k\phi(y)$. Also from (4.11), we know that $|Q^k(y, x) - q(x)| \leqq \Delta'\delta^{k-1}$, hence

$$\left|\frac{\phi(x)M^k(x, y)}{\lambda^k\phi(y)} - \Phi(x)\phi(x)\right| = \frac{\phi(x)}{\phi(y)}\left|\frac{M^k(x, y)}{\lambda^k} - \Phi(x)\phi(y)\right| \leqq \Delta'\delta^{k-1}$$

and so

$$\left|\frac{M^k(x, y)}{\lambda^k} - \Phi(x)\phi(y)\right| \leqq \frac{\phi(y)}{\phi(x)}\,\Delta'\delta^{k-1} \leqq \frac{D\Delta'}{d}\,\delta^{k-1} .$$

□

REMARK. Harris [7] proved a result similar to this one for primitive kernels.

THEOREM 4.3. *Let* $\{M_n\}$ *be a sequence of kernels satisfying Condition* S′. *For a given starting function* f_0, *if* $\int |f_n^*(y) - f_{n+1}^*(y)|\mu(dy) \to_n 0$ *then* $|f_n^*(y) - \psi_n(y)| \to_n 0$, *uniformly in* y.

PROOF.

$$\begin{aligned}|f_n^*(y) - \psi_n(y)| &\leqq \left|f_n^*(y) - \frac{f_n^*(y)}{\rho_n^k\lambda_n^k(\Phi_n, f_n^*)}\right| + \left|\frac{f_n^*(y)}{\rho_n^k\lambda_n^k(\Phi_n, f_n^*)} - \psi_n(y)\right| \\ &\leqq \sup_y f_n^*(y)\left|1 - \frac{1}{\rho_n^k\lambda_n^k(\Phi_n, f_n^*)}\right| + \left|\frac{f_n^*(y)}{\rho_n^k\lambda_n^k(\Phi_n, f_n^*)} - \psi_n(y)\right| \\ &= \sup_y f_n^*(y)\left|\int A_n^k(y)\mu(dy)\right| + |A_n^k(y)| ,\end{aligned}$$

where $A_n^k(y) = [f_n^*(y)/\rho_n^k\lambda_n^k(\Phi_n, f_n^*)] - \psi_n(y)$. It suffices to show that for k_0 and n sufficiently large, $|A_n^{k_0}(y)|$ can be made small.

$$|A_n^k(y)| \leqq \frac{1}{(\Phi_n, f_n^*)}\left\{\left|\frac{f_n^*(y)}{\rho_n^k\lambda_n^k} - \frac{\int f_n^*(x)M_n^k(x, y)\mu(dx)}{\lambda_n^k}\right| + \left|\frac{\int f_n^*(x)M_n^k(x, y)\mu(dx)}{\lambda_n^k} - \int \psi_n(y)\Phi_n(x)f_n^*(x)\mu(dx)\right|\right\} . \tag{4.12}$$

The second term of (4.12) is less than or equal to

$$\sup_x f_n^*(x) \int \left|\frac{M_n^k(x, y)}{\lambda_n^k} - \psi_n(y)\Phi_n(x)\right| \mu(dx) . \tag{4.13}$$

From Lemma 4.3, we know that given $\gamma > 0$, there is a k_0 such that for all n,

$$\left|\frac{M_n^{k_0}(x, y)}{\lambda_n^{k_0}} - \psi_n(y)\Phi_n(x)\right| \leqq \frac{D}{d}\,\Delta'\delta^{k_0-1} < \frac{v\gamma}{\Delta\mu(S)} .$$

Hence (4.13) is less than or equal to γ.

Using the bounds in (4.9) for ρ_n, the bounds for λ_n, and Lemma 4.2 we know that for k_0 fixed and $\gamma > 0$ given, there exists an $N = N(k_0, \gamma)$ such that for $n \geqq N$,

$$|f_n^*(y) - \rho_n^{k_0} \int f_n^*(x)M_n^{k_0}(x, y)\mu(dx)| \leqq [v\Delta\mu(S)]^{k_0}\gamma .$$

Hence

$$\begin{aligned}&\left|\frac{f_n^*(y)}{\rho_n^{k_0}\lambda_n^{k_0}} - \frac{\int f_n^*(x)M_n^{k_0}(x, y)\mu(dx)}{\lambda_n^{k_0}}\right| \\ &\quad = \frac{1}{\rho_n^{k_0}\lambda_n^{k_0}}|f_n^*(y) - \rho_n^{k_0}\int f_n^*(x)M_n^{k_0}(x, y)\mu(dx)| \leqq \left[\frac{v\Delta\mu(S)}{\rho_n\lambda_n}\right]^{k_0}\gamma \leqq \gamma .\end{aligned}$$

Using these inequalities with (4.12), we see that for $n \geqq N(k_0, \gamma)$, $|A_n^{k_0}(y)| < 2\gamma/b$. □

COROLLARY 4.2. *Let* $\{M_n\}$ *be a sequence of kernels satisfying Condition* S′. *If for some starting function* f_0, $|f_n^*(y) - f_{n+1}^*(y)| \to_n 0$, *then* $|\psi_n(y) - \psi_{n+1}(y)| \to_n 0$.

PROOF. This follows from the bounded convergence theorem and the inequality

$$|\psi_n(y) - \psi_{n+1}(y)| \leqq |\psi_n(y) - f_n^*(y)| + |f_n^*(y) - f_{n+1}^*(y)| + |f_{n+1}^*(y) - \psi_{n+1}(y)| . \qquad \square$$

REMARK. The hypotheses of Corollary 4.2 require that for *some* starting function $f_0(x)$, $|f_n^*(y) - f_{n+1}^*(y)| \to_n 0$. It follows from Corollary 4.2 and (4.8) that if this condition does hold for one starting function, then it will hold for *all* starting functions.

We can summarize the results of this section by noting that under Condition S′, since $\mu(S) < \infty$

$$|\psi_n(y) - \psi_{n+1}(y)| \to_n 0 \qquad \text{if and only if} \quad |f_n^*(y) - f_{n+1}^*(y)| \to_n 0 ,$$

for all starting functions $f_0(x)$. Further, either condition implies $|f_n^*(y) - \psi_n(y)| \to 0$. Hence the sequence $\{f_n^*(y)\}$ converges if and only if $\{\psi_n(y)\}$ converges, and $\{M_n(x, y)\}$ is PSE if and only if $\{\psi_n(y)\}$ converges uniformly.

REMARK. When the kernels $\{M_n\}$ are stochastic, the right eigenfunction corresponding to $\lambda = 1$ is $\Phi_n(x) = 1$, which is bounded and integrable. Condition (3.2) is also satisfied. Hence one need only check the remaining conditions of S′ to apply the above results.

5. Convergence theorems for eigenfunctions. In general, it is not easy to find the eigenfunctions of a kernel, making it difficult to verify the sufficient conditions of Theorems 4.1 and 4.2.

In this section we give sufficient conditions for the uniform convergence to zero of $\psi_n - \psi_{n+1}$, which may be easier to verify than the convergence itself.

A sequence of primitive kernels satisfying Condition S′ is said to satisfy *Condition* C.

REMARKS. The condition that $0 < \Delta_l \leqq M_n(x, y) \leqq \Delta_u \leqq \infty$ for all n implies Condition C. Also, Harris [7] has shown that for primitive kernels, the eigenvalues corresponding to positive bounded eigenfunctions is in fact the dominant eigenvalue or dominant root.

THEOREM 5.1. *Let $\{M_n\}$ be a sequence of kernels satisfying Condition* C. *If $\int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) \to_n 0$ uniformly in y, then $|\lambda_n - \lambda_{n+1}| \to_n 0$.*

PROOF. This proof is based on a characterization of the dominant root of primitive kernels given by Harris [7]. According to Harris, if

$$S_n = \{\lambda_n' > 0 : \text{there exists a bounded nonnegative function } f(x) \text{ such that } \int f(x)M_n(x, y)\mu(dx) \geqq \lambda_n' f(y), \text{ with strict inequality for some } y\}$$

then $\lambda_n = \sup \lambda_n' \in S_n$ is the dominant root.

Let $\varepsilon > 0$ be given. It suffices to show that for all n sufficiently large, $(\lambda_{n+1} - \varepsilon) \in S_n$ and $(\lambda_n - \varepsilon) \in S_{n+1}$, since this implies $|\lambda_n - \lambda_{n+1}| \leqq \varepsilon$. Choose $\gamma < \varepsilon d$. Then for all n, $\varepsilon\psi(y) \geqq \varepsilon d > \gamma$. It follows from the hypothesis that there exists an $N = N(\gamma)$ such that for $n \geqq N(\gamma)$, $\int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) < \gamma/D$.

Then

$$|\int \phi_{n+1}(x)M_n(x, y)\mu(dx) - \int \phi_{n+1}(x)M_{n+1}(x, y)\mu(dx)| \leqq D \int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) < \gamma .$$

Hence

$$\int \phi_{n+1}(x)M_n(x, y)\mu(dx) \geqq \int \phi_{n+1}(x)M_{n+1}(x, y)\mu(dx) - \gamma = \lambda_{n+1}\phi_{n+1}(y) - \gamma > (\lambda_{n+1} - \varepsilon)\phi_{n+1}(y) .$$

Therefore $(\lambda_{n+1} - \varepsilon) \in S_n$. A similar argument shows that $(\lambda_n - \varepsilon) \in S_{n+1}$. □

LEMMA 5.1. *Let $\{M_n\}$ be a sequence of kernels satisfying Condition* C *and such that $\int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) \to_n 0$ uniformly in y. Then for all k,*

$$\int \left|\frac{M_n^k(x, y)}{\lambda_n^k} - \frac{M_{n+1}^k(x, y)}{\lambda_{n+1}^k}\right| \mu(dx) \to_n 0 .$$

PROOF. By induction. For notational convenience, define $K_n^k(x, y) = M_n^k(x, y)/\lambda_n^k$. When $k = 1$,

$$\begin{aligned} &\int |K_n(x, y) - K_{n+1}(x, y)|\mu(dx) \\ (5.1)\qquad &\leqq \int \left|\frac{M_n(x, y)}{\lambda_n} - \frac{M_n(x, y)}{\lambda_{n+1}}\right| \mu(dx) + \int \left|\frac{M_n(x, y)}{\lambda_{n+1}} - \frac{M_{n+1}(x, y)}{\lambda_{n+1}}\right| \mu(dx) \\ &\leqq \Delta\mu(S)\left|\frac{\lambda_{n+1} - \lambda_n}{\lambda_n \lambda_{n+1}}\right| + \frac{1}{\lambda_{n+1}} \int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) \\ &\leqq \frac{\Delta\mu(S)}{v^2} |\lambda_{n+1} - \lambda_n| + \frac{1}{v} \int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) . \end{aligned}$$

The first term of (5.1) goes to zero by Theorem 5.1 and the second term by hypothesis. Next consider

$$\begin{aligned} &\int |K_n^{k+1}(x, y) - K_{n+1}^{k+1}(x, y)|\mu(dx) \\ &= \int |\int [K_n^k(x, z)K_n(z, y) - K_n^k(x, z)K_{n+1}(z, y) \\ (5.2)\qquad &\quad + K_n^k(x, z)K_{n+1}(z, y) - K_{n+1}^k(x, z)K_{n+1}(z, y)]\mu(dz)|\mu(dx) \\ &\leqq \iint K_n^k(x, z)|K_n(z, y) - K_{n+1}(z, y)|\mu(dz)\mu(dx) \\ &\quad + \iint |K_n^k(x, z) - K_{n+1}^k(x, z)|K_{n+1}(z, y)\mu(dz)\mu(dx) . \end{aligned}$$

Since $K_n^k(x, z)$ is bounded, it follows from the first part of this proof that the first term of (5.2) tends to zero. Applying the bounded convergence theorem and the induction hypothesis, the second term of (5.2) also tends to zero. □

LEMMA 5.2. *Under the conditions of Lemma* 5.1,

$$|\phi_n(y) - \int \phi_n(x)K_{n+1}^k(x, y)\mu(dx)| \to_n 0 \qquad \textit{for every } k .$$

PROOF. Since $\phi_n(y) = \int \phi_n(x)[M_n(x, y)/\lambda_n]\mu(dx)$, it follows that $\phi_n(y) = \int \phi_n(x)M_n^k(x, y)/\lambda_n^k\mu(dx) = \int \phi_n(x)K_n^k(x, y)\mu(dx)$ for every k. Hence

$$|\phi_n(y) - \int \phi_n(x)K_{n+1}^k(x, y)\mu(dx)| = |\int \phi_n(x)[K_n^k(x, y) - K_{n+1}^k(x, y)]\mu(dx)| \leqq D \int |K_n^k(x, y) - K_{n+1}^k(x, y)|\mu(dx)$$

which tends to zero by Lemma 5.1. □

The main result of this section is

THEOREM 5.2. *If* $\{M_n\}$ *is a sequence of kernels satisfying Condition* C *and if* $\int |M_n(x, y) - M_{n+1}(x, y)|\mu(dx) \to_n 0$ *uniformly in* y, *then* $|\psi_n(y) - \psi_{n+1}(y)| \to_n 0$ *uniformly in* y.

PROOF.

$$\begin{aligned}|\psi_n(y) - \psi_{n+1}(y)| &\leqq |\psi_n(y) - (\psi_n, \Phi_{n+1})\psi_{n+1}(y)| + \psi_{n+1}(y)|(\psi_n, \Phi_{n+1}) - 1| \\ &= |A_n(y)| + \psi_{n+1}(y)|\int A_n(y)\mu(dy)|\end{aligned}$$

where $A_n(y) = \psi_n(y) - (\psi_n, \Phi_{n+1})\psi_{n+1}(y)$. Since $\psi_n(y)$ and $A_n(y)$ are bounded, it suffices to show that $|A_n(y)| \to_n 0$.

$$\begin{aligned}|A_n(y)| &\leqq |\psi_n(y) - \int \psi_n(x)K^k_{n+1}(x, y)\mu(dx)| \\ &\quad + |\int \psi_n(x)K^k_{n+1}(x, y)\mu(dx) - \int \psi_n(x)\Phi_{n+1}(x)\psi_{n+1}(y)\mu(dx)| \\ &\leqq |\psi_n(y) - \int \psi_n(x)K^k_{n+1}(x, y)\mu(dx)| \\ &\quad + D \int \left|\frac{M^k_{n+1}(x, y)}{\lambda^k_{n+1}} - \Phi_{n+1}(x)\psi_{n+1}(y)\right| \mu(dx) \,.\end{aligned} \tag{5.3}$$

Since Condition C is stronger than Condition S′, Lemma 4.4 can be applied. That is, it is possible to choose k_0 large enough so that the second term of (5.3) is less than $\varepsilon/2$. Further, Lemma 5.2 can be applied so that given k_0, there exists an $N = N(k_0, \varepsilon)$ such that for $n \geqq N$, the first term of (5.3) is less than $\varepsilon/2$. Hence for $n \geqq N(k_0, \varepsilon)$, $|A_n(y)| < \varepsilon$, independently of y. □

It is possible to show by arguments like those given in this section that if $\{M_n\}$ and M_0 are kernels satisfying Condition C, and if

$$\int |M_n(x, y) - M_0(x, y)|\mu(dx) \to_n 0 \tag{5.4}$$

uniformly in y, then $\lambda_n \to \lambda_0$ and $\psi_n(y) \to \psi_0(y)$ uniformly in y. Hence, in view of Theorem 4.2, (5.4) is sufficient for PSE behavior.

REMARK. Conditions analogous to those given in Theorem 5.2 can be given which guarantee that $|\Phi_n(x) - \Phi_{n+1}(x)| \to_n 0$ uniformly.

6. Applications. A sequential probability ratio test (SPRT) of $H_0: f = f_0$ against $H_1: f = f_1$ reduces to consideration of $\sum_{i=1}^n z_i$ where

$$z_i = \ln [f_1(x_i)/f_0(x_i)] \,. \tag{6.1}$$

The test requires continued sampling if

$$a \equiv \ln B < \textstyle\sum_{i=1}^n z_i < \ln A \equiv b \,.$$

One accepts or rejects H_0 depending on whether $\sum_{i=1}^n z_i \leqq a$ or $\sum_{i=1}^n z_i \geqq b$. A and B are determined by the probability of type I and type II errors. (See Wald [10]).

REMARK. David and Mendigo [3] considered the relationship between certain finite Markov chains and binomial SPRT's. We here consider a more general situation.

If the random variables z_i, have densities h_i and are independent, then $S_n = \sum_{i=1}^n z_i$ forms a Markov chain with transition kernels given by $P_i(x, y) = h_i(y - x)$. If the random variables are identically distributed, the corresponding Markov chain is homogeneous. The numbers a and b defined above act as absorbing barriers for the Markov chain.

To study the probability distribution over the non-absorbing states, conditioned on the event that absorption has not yet taken place, define $M = P$ on $[a, b] \times [a, b]$ and set

$$f_n^*(y) = f_n(y)/\int_a^b f_n(y)\,dy$$

where

$$f_n(y) = \int_a^b f_{n-1}(x)M(x, y)\,dx = \int_a^b f_0(x)M^n(x, y)\,dx\,.$$

By checking the conditions of the theorems concerning PSE or PWE behavior, one might ascertain the asymptotic behavior of $f_n^*(y)$.

REMARK. If a density, f_0, does not exist for the starting distribution, then set $f_1(y) = \int_a^b M(x, y)\,dF_0(x)$. In the case where $S_0 = 0$, $dF_0(0) = 1$ and $f_1(y)$ is $M(0, y)$.

EXAMPLE. For the negative exponential with β known, test $H_0\colon \alpha = \alpha_0$ against $H_1\colon \alpha = \alpha_1$. Without loss of generality, assume $\beta = 0$ and $\alpha_0 > \alpha_1$. Then $\ln[f_1(x)/f_0(x)] = \ln(\alpha_1/\alpha_0) + (\alpha_0 - \alpha_1)x$. Defining

$$a = \ln B/(\alpha_0 - \alpha_1)\,, \qquad b = \ln A/(\alpha_0 - \alpha_1) \tag{6.2}$$

and

$$\gamma = (\ln \alpha_0 - \ln \alpha_1)/(\alpha_0 - \alpha_1)\,,$$

the corresponding SPRT requires continued sampling if $a < \sum_{i=1}^n (x_i - \gamma) < b$, etc. If $z_i = x_i - \gamma$, the density for z_i is given by

$$\begin{aligned} h_i(t) &= \alpha e^{-\alpha(t+\gamma)} && \text{if } t \geqq -\gamma \\ &= 0 && \text{if } t < -\gamma\,. \end{aligned}$$

With a and b as defined in (6.2), the appropriate kernel for this problem, for $(x, y) \in (a, b) \times (a, b)$, is

$$\begin{aligned} M(x, y) &= \alpha e^{-\alpha(y-x+\gamma)} && \text{if } y - x \geqq -\gamma \\ &= 0 && \text{if } y - x < -\gamma\,. \end{aligned}$$

It is not hard to show that this kernel is primitive and bounded, hence it has a positive left eigenfunction $\psi(y)$. In fact Condition S is satisfied and it follows from Theorem 4.2 that $|f_n^*(y) - \psi(y)| \to_n 0$. Further we can exhibit the eigenfunction $\psi(y)$ for this kernel.

Define $K = \alpha e^{-\alpha y}$ and $c_n = \int_a^{b-n\gamma} \psi(x)e^{\alpha x}dx$. Then for $y \in [b - n\gamma, b - (n-1)\gamma)$,

$$\psi(y) = \sum_{r=1}^n \{c_{n-r}(K/\lambda)^r[y - (b - n\gamma)]^{r-1}/(r-1)!\}e^{-\alpha y}\,. \tag{6.3}$$

Equation (6.3) holds for $n = 1, 2, \cdots, N$ where $N = (b - a)/\gamma$ if this is an integer or $N = [(b - a)/\gamma] + 1$, where $[\cdot]$ represents the greatest integer function.

Weiss [11] defined a generalized sequential probability ratio test (GSPRT) for testing a simple hypothesis against a simple alternative. With a GSPRT, one considers sequences of constants $\{A_n\}$ and $\{B_n\}$ rather than fixed constants A and B to determine whether to continue sampling, accept H_0, or reject H_0.

If $a_n = \ln B_n$, $b_n = \ln A_n$, and $S_n = \sum_{i=1}^n z_i$ for z_i as defined in (6.1), then the GSPRT can be thought of as a random walk with changing absorbing barriers. To study the probability distribution over the non-absorbing states, one considers the sequence $\{f_n(y)\}$, where

$$\begin{aligned} f_n(y) &= \int_{a_n}^{b_n} f_{n-1}(x)M(x, y)\,dx \\ &= \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f_0(z_1)M(z_1, z_2) \cdots M(z_n, y)\,dz_1 \cdots dz_n . \end{aligned}$$

In this case, $M(z_k, z_{k+1})$ is defined on $(a_k, b_k) \times (a_{k+1}, b_{k+1})$. If $a = \inf_n a_n$ and $b = \sup_n b_n$ are both finite, then the following transformation will change this problem to one with non-homogeneous kernels defined on $(a, b) \times (a, b)$. For $z \in (a_k, b_k)$, make the transformation $z = c_k + r_k w$, where $c_k = (a_k b - ab_k)/(b - a)$ and $r_k = (b_k - a_k)/(b - a)$. Further, define $g_0(w) = f_0(c_1 + r_1 w)$ and, for $(v, w) \in (a, b) \times (a, b)$, define

$$M_k(v, w) = r_k M(c_k + r_k v, c_{k+1} + r_{k+1} w) . \tag{6.4}$$

Finally, for $y \in (a_{n+1}, b_{n+1})$, let $w = (y - c_{n+1})/r_{n+1}$. Then

$$\begin{aligned} f_n(y) &= f_n(c_{n+1} + r_{n+1} w) = g_n(w) \\ &= \int_a^b \cdots \int_a^b g_0(w_1)M_1(w_1, w_2) \cdots M_n(w_n, w)\,dw_1 \cdots dw_n . \end{aligned} \tag{6.5}$$

One can check the theorems of Sections 3 and 4 to investigate the asymptotic behavior of $g_n^*(w)$.

In order to apply Theorem 5.2 to this problem, one needs to know that $\int_a^b |M_n(v, w) - M_{n+1}(v, w)|\,dv \to_n 0$ uniformly in w. The following lemma gives sufficient conditions for this to hold.

LEMMA 6.1. *Let $\{M_n\}$ be a sequence of kernels defined by* (6.4) *with $M(x, y)$, the original kernel, nonnegative, and uniformly continuous in both arguments for $(x, y) \in (a, b) \times (a, b)$. If $|b_n - b_{n+1}| \to 0$ and $|a_n - a_{n+1}| \to 0$, then $\int_a^b |M_n(v, w) - M_{n+1}(v, w)|\,dv \to_n 0$ uniformly in w.*

PROOF. It is easy to see that both $|r_n - r_{n+1}|$ and $|c_n - c_{n+1}|$ tend to zero and the hypotheses imply that $M(x, y)$ is bounded. Consider

$$\begin{aligned} \int_a^b |M_n(v, w) - M_{n+1}(v, w)|\,dv &= \int_a^b |r_n M(c_n + r_n v, c_{n+1} + r_{n+1} w) \\ &\quad - r_{n+1} M(c_{n+1} + r_{n+1} v, c_{n+2} + r_{n+2} w)\,dv| \\ &\leqq \int_a^b |r_n - r_{n+1}| M(c_n + r_n v, c_{n+1} + r_{n+1} w)\,dv \\ &\quad + \int_a^b r_{n+1} |M(c_n + r_n v, c_{n+1} + r_{n+1} w) \\ &\quad - M(c_{n+1} + r_{n+1} v, c_{n+2} + r_{n+2} w)|\,dv . \end{aligned} \tag{6.6}$$

The first term of (6.6) goes to zero since $M(x, y)$ is bounded and since $|r_n - r_{n+1}| \to 0$. Since $r_{n+1} \leqq 1$, $|r_n - r_{n+1}| \to 0$, $|c_n - c_{n+1}| \to 0$, and since $M(x, y)$ is uniformly continuous, it follows that the second term of (6.6) also goes to zero. □

COROLLARY 6.1. *Under the conditions of Lemma* 6.1, *if the kernels* $\{M_n\}$ *satisfy Condition* C, *then* $|g_n^*(w) - \psi_n(w)| \to_n 0$, *where* $g_n(w)$ *is defined by* (6.5).

PROOF. The proof follows immediately from Lemma 6.1, Theorem 5.2 and Corollary 4.1. □

Acknowledgments. Both authors would like to express their sincere appreciation to Professor Herb T. David for his guidance in their dissertation research. We thank Professors David and Dean Isaacson for their helpful comments in the preparation of this paper. We also thank the referee for his comments.

REFERENCES

[1] BLUM, J. R. and REICHAW, M. (1971). Two integral inequalities. *Israel J. Math.* **9** 20-26.

[2] CONN, PATRICIA S. (1969). Asymptotic properties of sequences of positive kernels. Ph. D. dissertation, Iowa State Univ.

[3] DAVID, HERBERT T. and MENGIDO, RODOLFO M. (1963). Asymptotically exact truncation in binomial sequence analysis (preliminary report). Abstract, *Ann. Math. Statist.* **34** 1617.

[4] DOBRUSHIN, R. L. (1956). Central limit theorem for non-stationary Markov chains I, II. *Theor. Probability Appl.* **1** 65-80.

[5] DOOB, J. L. (1953). *Stochastic Process.* Wiley, New York.

[6] FELLER, WILLIAM (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

[7] HARRIS, THEODORE E. (1963). *The Theory of Branching Processes.* Springer-Verlag, Berlin.

[8] MADSEN, RICHARD W. (1971). A note on some ergodic theorems of A. Paz. *Ann. Math. Statist.* **42** 405-408.

[9] PAZ, A. (1970). Ergodic theorems for infinite probabilistic tables. *Ann. Math. Statist.* **41** 539-550.

[10] WALD, ABRAHAM (1947). *Sequential Analysis.* Wiley, New York.

[11] WEISS, LIONEL (1953). Testing of one simple hypothesis against another. *Ann. Math. Statist.* **24** 273-281.

RICHARD W. MADSEN
DEPARTMENT OF STATISTICS
UNIVERSITY OF MISSOURI
222 MATH. SCIENCES BUILDING
COLUMBIA, MISSOURI 65201

PATRICIA S. CONN
DIVISION OF MATHEMATICAL SCIENCES
PURDUE UNIVERSITY
FORT WAYNE, INDIANA 46805