The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 405–408

A NOTE ON SOME ERGODIC THEOREMS OF A. PAZ

BY RICHARD W. MADSEN

Iowa State University

In a paper entitled "Ergodic theorems for infinite probabilistic tables," Paz studies the ergodic properties of non-homogeneous Markov chains with countably infinite state space. This note is to draw attention to the fact that Paz's theory is easily extended to arbitrary state spaces.

0. Summary. In [4], Paz studies the ergodic properties of non-homogeneous Markov chains with countably infinite state space. This note is to draw attention to the fact that Paz's theory is easily extended to arbitrary state spaces.

1. Introduction. Let $(S, \mathscr{B}, \mu)$ be a σ-finite measure space. Doob ([3] page 190) defines stochastic transition density functions for the general state space in the homogeneous case. In the non-homogeneous case, we must consider a sequence of transition density functions $\{P_n(x, y)\}$ on $S \times S$ with the following properties:

(a) $P_n(x, y)$ is the density function for $P_n(x, B)$, i.e. $P_n(x, B) = \int_B P_n(x, y)\mu(dy)$ for $B \in \mathscr{B}$.

(b) $P_n(x, \cdot)$ for fixed x determines a probability on $\mathscr{B}$.

(c) $P_n(\cdot, B)$ for fixed B, determines a function of x which is measurable with respect to $\mathscr{B}$.

$P_n(x, y)$ will be referred to as a stochastic kernel and can be thought of as a one step transition kernel from time $n-1$ to n. In view of the properties a, b, and c, the multi-step stochastic kernels $P_{n,n+k}(x, y)$, from time $n-1$ to time $n+k$, are well defined:

$$P_{n,n+k}(x, y) = \int_S \cdots \int_S P_n(x, z_1)P_{n+1}(z_1, z_2) \cdots P_{n+k}(z_k, y)\mu(dz_1)\mu(dz_2) \cdots \mu(dz_k).$$

Following Paz, we define a non-homogeneous Markov chain for a general state space.

DEFINITION 1. A non-homogeneous Markov chain (NMC) is a couple $(S, \{P_n(x, y)\})$ where S is the state space and $\{P_n(x, y)\}$ is an infinite sequence of transition density functions satisfying conditions (a), (b), and (c).

NOTATION. For simplicity we take P_n, K, etc., to mean $P_n(x, y)$, $K(x, y)$, etc.

2. Norms and ergodic coefficients. Dobrushin [2] has defined the ergodic coefficient $\alpha(P)$ of a stochastic kernel as follows:

DEFINITION 2. If P is a stochastic kernel, the ergodic coefficient $\alpha(P)$ is

$$\alpha(P) = 1 - \sup \left| \int_B [P(x, y) - P(z, y)]\mu(dy) \right|$$

where the sup is taken over all $B \in \mathscr{B}$ and all $x, z \in S$.

Received June 8, 1970.

REMARK. Paz [4] defined the ergodic coefficient $\gamma(P)$ as

$$\gamma(P) = \inf_{i_1,i_2} \sum_{j=1}^{\infty} \min(p_{i_1j}, p_{i_2j}).$$

It can be shown that the analog in the continuous case is equivalent to $\alpha(P)$ as defined here.

DEFINITION 3. Let P be a stochastic kernel. The ergodic coefficient $\delta(P)$ is defined as:

$$\delta(P) = 1-\alpha(P).$$

DEFINITION 4. If K is y-integrable, then the norm $||K||$ is defined as:

$$||K|| = \sup_x \int |K(x,y)|\, \mu(dy).$$

If f is a function of a single variable, we take $||f||$ as the $L_1(S)$ norm of $f(||f|| = \int |f(x)|\, \mu(dx))$.

DEFINITION 5. A stochastic kernel is called constant if it is independent of x, i.e. $P(x,y) = f(y)$ for all x.

REMARK. This is equivalent to taking $\alpha(P) = 1$.

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

$$\delta(P) = \sup_{x,z} \int (P(x,y)-P(z,y))^+ \mu(dy) \qquad \textit{and}$$

$$2\delta(P) = \sup_{x,z} \int |P(x,y)-P(z,y)|\, \mu(dy).$$

PROOF. Straightforward and is left to the reader.

LEMMA 2. *Assuming the required integrability, define*

$$KL(x,y) = \int K(x,z)L(z,y)\mu(dz).$$

Then

$$||KL|| \leqq ||K||\, ||L|| \quad \textit{and} \quad ||K+L|| \leqq ||K||+||L||.$$

PROOF. This is a simple application of Fubini's theorem for the first part and the triangle inequality for the second.

LEMMA 3. *Let P be a stochastic kernel and let E_{x_0} be the constant stochastic kernel defined by $E_{x_0}(x,y) = P(x_0,y)$, where x_0 is a fixed point in S then*

(a) $2\delta(P) \geqq ||P-E_{x_0}||$ *and*
(b) *for every $\varepsilon > 0$ there is some $x_0 \in S$ such that*

$$2\delta(P) \leqq ||P-E_{x_0}||+\varepsilon.$$

PROOF. Follows from Lemma 1 and properties of sup.

LEMMA 4. *Every stochastic kernel P can be represented in the form $P = E+Q$ where E is a constant stochastic kernel and* $||Q|| \leqq 2\delta(P)$.

PROOF. Follows from Lemma 3 with $E = E_{x_0}$.

LEMMA 5. *If P and Q are stochastic kernels, then* $\delta(PQ) \leqq \delta(P)\,\delta(Q)$.

PROOF. This is proven by Dobrushin ([2] page 333).

LEMMA 6. *Let P be a stochastic kernel and f an integrable function such that* $0 < \|f\| < \infty$ *and* $\int f(x)\mu(dx) = 0$, *then* $\|fP\| \leqq \|f\|\,\delta(P)$ *where* $fP(y) = \int f(x)P(x, y)\mu(dx)$.

PROOF. Define $g(x) = 2f^+(x)/\|f\|$ and $h(x) = -2f^-(x)/\|f\|$ where $f^+(x) = \max(f(x), 0)$ and $f^-(x) = \min(f(x), 0)$. Since g and h are nonnegative and integrate to 1, a stochastic kernel can be defined as:

$$
\begin{aligned}
Q(x, y) &= g(y) \qquad \text{if} \quad x = x_1 \\
&= h(y) \qquad \text{otherwise}
\end{aligned}
$$

where x_1 is an arbitrary point. With these definitions the reader can complete the proof by referring to Paz [4].

REMARK. Lemma 6 is a consequence of the inequality (1) found in [1].

LEMMA 7. *Suppose that* $\|K\| < \infty$ *and that, for each fixed* x, $0 < \int |K(x, y)|\,\mu(dy)$ *and* $\int K(x, y)\mu(dy) = 0$. *If P is a stochastic kernel, then* $\|KP\| \leqq \|K\|\,\delta(P)$.

PROOF. Straightforward using Lemma 6.

3. Ergodic theorems for non-homogeneous Markov chains. Following Paz, we distinguish two types of long run behavior of an NMC. If $\lim_{n\to\infty} \delta(P_{mn}) = 0$ for all m, then the chain is called weakly ergodic. If there is a constant stochastic kernel Q such that, for any m, $\lim_{n\to\infty} \|P_{mn} - Q\| = 0$, then the chain is called strongly ergodic.

REMARK. Our definition of strongly ergodic is in fact equivalent to Paz's which can be seen as follows. If $K(y)$ is a function satisfying $\lim_{n\to\infty} \|P_{mn}(x, y) - K(y)\| = 0$, then $K(y)$ will be a constant stochastic kernel.

PROOF. $K(y)$ integrates to 1 since $|1 - \int K(y)\mu(dy)| = |\int [P_{mn}(x, y) - K(y)]\mu(dy)| \leqq \int |P_{mn}(x, y) - K(y)|\,\mu(dy) \leqq \sup_x \int |P_{mn}(x, y) - K(y)|\,\mu(dy) = \|P_{mn}(x, y) - K(y)\| \to 0$.
Also $K(y)$ must be nonnegative a.e. If not, let B be the set where $K(y) < 0$, i.e. $\mu(B) > 0$ and $\int_B K(y)\mu(dy) = -\beta$ for $\beta > 0$. Then $\int_B |P_{mn}(x, y) - K(y)|\,\mu(dy) \geqq \int_B - K(y)\mu(dy) = \beta$ for all n, which contradicts $\|P_{mn}(x, y) - K(y)\| \to 0$.
The following theorems and corollaries are stated without proof. The proofs can be supplied by the reader using the above definitions and lemmas and following the earlier work of Paz.

THEOREM 1. *Let* $\{P_n\}$ *be an* NMC. *The following are equivalent*:

(a) $\{P_n\}$ *is weakly ergodic.*
(b) *There exists a subdivision of the chain into blocks of kernels* $\{P_{i_j+1,\, i_{j+1}}\}$ *such that* $\sum_{j=1}^{\infty} \alpha(P_{i_j+1,\, i_{j+1}})$ *diverges.*

(c) *For each m there is a sequence of constant stochastic kernels* $\{E_{mn}\}$ *such that* $\lim_{n\to\infty} ||P_{mn}-E_{mn}|| = 0$.

(d) *If* $P_i = E_i + R_i$ *where* E_i *is a constant stochastic kernel, then* $\lim_{n\to\infty} ||R_{mn}|| = 0$.

THEOREM 2. *An* NMC *is strongly ergodic* iff *for every m there is a sequence of constant stochastic kernels* $\{E_{mn}\}$ *and a constant stochastic kernel* E_m *such that*

(a) $\lim_{n\to\infty} ||P_{mn}-E_{mn}|| = 0$.
(b) $\lim_{n\to\infty} ||E_{mn}-E_m|| = 0$.

COROLLARY 1. *A strongly ergodic chain is also weakly ergodic. A weakly ergodic chain which satisfies* (b) *of Theorem* 2 *is strongly ergodic.*

THEOREM 3. *Let* $\{P_i\}$ *and* $\{Q_i\}$ *be two* NMC's *such that* $\sum_i ||P_i - Q_i|| < \infty$, *then for any* $\varepsilon > 0$ *there is an integer* m_0 *such that* $||P_{mn}-Q_{mn}|| < \varepsilon$ *for all* $m \geqq m_0$ *and all* $n > m$.

COROLLARY 2. *Let* $\{P_i\}$ *and* $\{Q_i\}$ *be two* NMC's *satisfying the conditions of Theorem* 3. *If one of the chains is weakly* (*strongly*) *ergodic then so is the other.*

Acknowledgment. The author would like to thank the referee and Professors H. T. David and D. L. Isaacson for their helpful comments in preparing this note.

REFERENCES

[1] BLUM, J. R. and REICHAW, M. (1971). Two integral inequalities. *Israel J. Math.* (to appear).
[2] DOBRUSHIN, R. L. (1956). Central limit theorem for non-stationary Markov chains I, II. *Theor. Probability Appl.* **1** 65–80 and 329–383 (English translation).
[3] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.
[4] PAZ, A. (1970). Ergodic theorems for infinite probabilistic tables. *Ann. Math. Statist.* **41** 539–550.