The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1732-1736

ON TWO RECENT PAPERS ON ERGODICITY IN NONHOMOGENEOUS MARKOV CHAINS[1]

BY MARIUS IOSIFESCU

Academy of Socialist Republic of Romania and University of Cambridge

In [7] ergodic properties of nonhomogeneous denumerable state space Markov chains were studied. It was noticed in [6] that the results obtained in [7] were easily extended to arbitrary state spaces. However, it was admitted in [6] that the transition mechanism of the chain was defined by means of transition density functions, thus restricting the generality of the approach and, moreover, introducing elements irrelevant to the problem. The aim of this note is to draw attention to the fact that ergodic properties of the most general nonhomogeneous Markov chains are easily obtained by using a theory developed by Dobrušin [2] in the middle fifties.

1. The ergodicity coefficient. Let $(X, \mathscr{X})$ and $(Y, \mathscr{Y})$ be two measurable spaces and let P denote a transition probability function (t.p.f.) from the first measurable space to the second one. In other words, P is a real valued function defined on $X \times \mathscr{Y}$ such that $P(x, \cdot)$ is a probability on $\mathscr{Y}$ for any $x \in X$ and $P(\cdot, B)$ is a $\mathscr{X}$-measurable function for any $B \in \mathscr{Y}$.

DEFINITION. The real number

$$\alpha(P) = 1 - \sup |P(x', B) - P(x'', B)| ,$$

where the sup is taken over all $x', x'' \in X$ and all $B \in \mathscr{Y}$, is called the *ergodicity coefficient* of the t.p.f. P.

Clearly, $0 \leqq \alpha(P) \leqq 1$. In case $\alpha(P) = 1$, $P(\cdot, B)$ is a constant function for any $B \in \mathscr{Y}$ and P is said to be a *constant* t.p.f. (In fact, a constant t.p.f. is a probability on Y.)

The basic properties of the ergodicity coefficient have been established by Dobrušin [2] (for a compact treatment see [5] pages 38–45). We now list those properties we shall need. Given a measurable space $(V, \mathscr{V})$, let $L(V, \mathscr{V})$ denote the linear space of all finite completely additive signed measures λ defined on $\mathscr{V}$ such that $\lambda(V) = 0$. It is easy to see that under the norm

$$||\lambda||_{\mathscr{V}} = \sup_{A \in \mathscr{V}} |\lambda(A)|$$

$L(V, \mathscr{V})$ is a Banach space. Obviously, we can write

$$(1) \qquad \delta(P) =_{\text{def}} 1 - \alpha(P) = \sup_{x', x'' \in X} ||P(x', \cdot) - P(x'', \cdot)||_{\mathscr{Y}} .$$

Received December 1, 1971.

[1] This paper was written while the author was an Overseas Fellow of Churchill College, Cambridge, England.

AMS 1970 *subject classification.* Primary 60J05.

Key words and phrases. Nonhomogeneous Markov chains, ergodicity coefficient, weak ergodicity, strong ergodicity, trivial tail σ-algebra.

It follows from (1) that

$$\sup_{x' \in X} ||P(x', \cdot) - P^x(\cdot)||_{\mathscr{Y}} = \delta(P)$$

for any $x \in X$, where P^x is the constant t.p.f. defined by $P^x(\cdot) = P(x, \cdot)$. This means that any t.p.f. P can be represented as $P = E + R$, where E is a constant t.p.f. and $\sup_{x \in X} ||R(x, \cdot)||_{\mathscr{Y}} \leqq \delta(P)$.

Further, an operator $\mathfrak{P}$ mapping $L(X, \mathscr{X})$ into $L(Y, \mathscr{Y})$ is associated with the t.p.f. P by setting

$$(\lambda\mathfrak{P})(B) = \int_X \lambda(dx)P(x, B)\,, \qquad B \in \mathscr{Y}\,.$$

Then it can be proved that the norm of $\mathfrak{P}$,

$$||\mathfrak{P}|| = \sup_{\lambda \in L(X, \mathscr{X})} \frac{||\lambda\mathfrak{P}||_{\mathscr{Y}}}{||\lambda||_{\mathscr{X}}}$$

and the ergodicity coefficient of P are connected by

(2) $$||\mathfrak{P}|| = 1 - \alpha(P)(=\delta(P))\,.$$

In particular, it follows that

(3) $$||\lambda\mathfrak{P}||_{\mathscr{Y}} \leqq \delta(P)||\lambda||_{\mathscr{X}}$$

for all $\lambda \in L(X, \mathscr{X})$.

Finally, let $(X, \mathscr{X})$, $(Y, \mathscr{Y})$, $(Z, \mathscr{Z})$ be three measurable spaces. Consider a t.p.f. P' from $(X, \mathscr{X})$ to $(Y, \mathscr{Y})$ and a t.p.f. P'' from $(Y, \mathscr{Y})$ to $(Z, \mathscr{Z})$. Let P be the t.p.f. from $(X, \mathscr{X})$ to $(Z, \mathscr{Z})$ defined by

$$P(x, C) = \int_Y P'(x, dy)P''(y; C)$$

for $x \in X$, $C \in Z$. Due to the fact that the operators $\mathfrak{P}$, $\mathfrak{P}'$ and $\mathfrak{P}''$ associated with P, P' and P'' respectively satisfy $\mathfrak{P} = \mathfrak{P}', \mathfrak{P}''$, which in turn implies

$$||\mathfrak{P}|| \leqq ||\mathfrak{P}'||\,||\mathfrak{P}''||\,;$$

equation (2) leads to

(4) $$\delta(P) \leqq \delta(P')\delta(P'')\,.$$

Notice that in case $X = Y = I$ (at most a denumerable set), when the t.p.f. P is defined by means of a stochastic matrix $(p_{ij})_{i,j \in I}$ by

$$P(i, A) = \sum_{j \in A} p_{ij}\,, \qquad i \in I,\ A \subset I\,,$$

Dobrušin [2] pages 333–335 (see also [5] pages 1–2) has shown that

(5) $$\alpha(P) = \inf_{i,j \in I} \sum_{k \in I} \min(p_{ik}, p_{jk})\,.$$

Taking into account that

$$\min(a, b) = \tfrac{1}{2}(a + b - |a - b|)\,,$$

we get

$$\alpha(P) = 1 - \tfrac{1}{2} \sup_{i,j \in I} \sum_{k \in I} |p_{ik} - p_{jk}|\,,$$

that is

(6) $$\delta(P) = \tfrac{1}{2} \sup_{i,j \in I} \textstyle\sum_{k \in I} |p_{ik} - p_{jk}| \,.$$

Therefore, the fact that $\delta(P)$ and $\alpha(P)$, defined by (5) and (6), are connected by $\delta(P) = 1 - \alpha(P)$ is not at all surprising, as was suggested by Paz ([7] page 542).

2. Nonhomogeneous Markov chains. Following Dobrušin [2] pages 67–68 (see also [5] pages 43–44) a nonhomogeneous Markov chain (NMC) can be considered as a sequence of measurable state spaces $(X_j, \mathscr{X}_j)$ and t.p.f.'s ${}^jP$ from $(X_j, \mathscr{X}_j)$ to $(X_{j+1}, \mathscr{X}_{j+1})$, $j = 0, 1, 2 \cdots$. Clearly, ${}^jP(x_j, A_{j+1})$ is to be thought of as the probability of being in the set $A_{j+1} \in \mathscr{X}_{j+1}$ at time $j + 1$ conditional on being in $x_j \in X_j$ at time j. Let us define as usual the n-step t.p.f.'s ${}^jP^n$ from $(X_j, \mathscr{X}_j)$ to $(X_{j+n}, \mathscr{X}_{j+n})$, $j \geqq 0$, $n \geqq 1$ by

$$\begin{aligned} {}^jP^n(x_j, A_{j+n}) = \int_{X_{j+1}} {}^jP(x_j, dx_{j+1}) \cdots \int_{X_{j+n-1}}^{n+j-2} P(x_{n+j-2}, dx_{n+j-1}) \\ \times \int_{A_{j+n}}^{n+j-1} P(x_{n+j-1}, dx_{n+j}) \,, \end{aligned}$$

where $x_j \in X_j$, $A_{j+n} \in \mathscr{X}_{j+n}$. An immediate consequence of (4) is the inequality

(7) $$\delta({}^jP^n) \leqq \textstyle\prod_{i=j}^{n+j-1} \delta({}^iP)$$

for any $j \geqq 0$, $n \geqq 1$.

3. Weak ergodicity. An NMC is said to be *weakly ergodic* if $\lim_{n\to\infty} \delta({}^jP^n) = 0$ for all $j \geqq 0$.

By making use of (1) and (7) one can first prove

THEOREM 1. *An* NMC *is weakly ergodic* iff (= *if and only if*) *either one of the following conditions is fulfilled.*

(a) *There exists a strictly increasing sequence* $(j_k)_{k\geqq 1}$ *of natural numbers such that* $\sum_{k\geqq 1} \alpha({}^{j_k}P^{j_{k+1}-j_k})$ *diverges.*

(b) *For an arbitrarily fixed* $0 < \varepsilon < 1$ *there is a function f mapping the set of the natural numbers into itself such that* $\liminf_{j\to\infty} \alpha({}^jP^{f(j)}) \geqq \varepsilon$.

By making use of (1) and (3) (cf. [7] pages 544–545) one obtains

THEOREM 2. *An* NMC *is weakly ergodic* iff *either one of the following conditions is fulfilled.*

(c) *For any* $j \geqq 0$ *there is a sequence of constant* t.p.f.'*s* $({}^jE_n)_{n\geqq 1}$ *such that*

$$\lim_{n\to\infty} \sup_{x_j \in X_j} \|{}^jP^n(x_j, \cdot) - {}^jE_n(\cdot)\|_{\mathscr{X}_{n+j}} = 0 \,.$$

(d) *If* ${}^jP = {}^jE + {}^jR$, *where the* ${}^jE$ *are constant* t.p.f.'*s then*

$$\lim_{n\to\infty} \sup_{x_j \in X_j} \|{}^jR^n(x_j, \cdot)\|_{\mathscr{X}_{n+j}} = 0$$

for all $j \geqq 0$.

Clearly, the ${}^jR^n$ above are constructed from the ${}^jR$ in the same way as the ${}^jP^n$ were constructed from the ${}^jP$.

REMARK 1. Theorem 1 is an easy extension of the corresponding theorem for finite NMCs (i.e. such that any $X_j, j \geqq 0$ is a finite set), proved by Hajnal ([3] page 239).

REMARK 2. Condition (b) and Theorem 1.2.13 in [5] imply that any weakly ergodic NMC has a trivial tail σ-algebra under any initial distribution. The converse is true at least for finite NMC, it being a consequence of Theorem 2 in [1].

REMARK 3. A more narrow concept of weak ergodicity, namely *uniform* weak ergodicity, is obtained by requiring that $\lim_{n\to\infty} \delta(^jP^n) = 0$ uniformly with respect to $j \geqq 0$. An NMC is uniformly weakly ergodic iff there exist an $\varepsilon > 0$ and a natural number n_0 such that $\alpha(^jP^{n_0}) \geqq \varepsilon$ for all $j \geqq 0$. If this condition is fulfilled, we have $\delta(^jP^n) \leqq (1-\varepsilon)^{n/n_0-1}$ for all $j \geqq 0$, $n \geqq 1$. For details see [4] or [5] pages 89–94.

4. Strong ergodicity. Suppose all the state spaces $(X_j, \mathscr{X}_j)$, $j \geqq 0$, are copies of a given measurable space $(X, \mathscr{X})$. An NMC with state space $(X, \mathscr{X})$ is said to be *strongly ergodic* if there exists a probability Q on $\mathscr{X}$ such that

$$\lim_{n\to\infty} \sup_{x\in X} \|^jP^n(x, \cdot) - Q(\cdot)\|_{\mathscr{X}} = 0 \tag{8}$$

for all $j \geqq 0$.

By making use of an elementary argument (cf. [7] pages 545–546) one can prove

THEOREM 3. *An* NMC *is strongly ergodic* iff *for any $j \geqq 0$ there is a sequence of constant* t.p.f.'*s* $(^jE_n)_{n\geqq 1}$ *and a constant* t.p.f. ^jE *such that*

$$\lim_{n\to\infty} \sup_{x\in X} \|^jP^n(x, \cdot) - {}^jE_n(\cdot)\|_{\mathscr{X}} = 0$$

and

$$\lim_{n\to\infty} \|^jE_n - {}^jE\|_{\mathscr{X}} = 0\,.$$

Note that Theorem 3 in [6] (corresponding to Theorem 5 in [7]) can also be transcribed.

REMARK 4. It seems that unlike weak ergodicity, strong ergodicity is not a "natural" concept for NMCs. In any case, strongly ergodic NMCs are a very restricted class of NMCs. On the other hand, in the homogeneous case weak ergodicity and strong ergodicity coincide.

REMARK 5. A more narrow concept of strong ergodicity, namely *uniform* strong ergodicity, is obtained by requiring that (8) holds uniformly with respect to $j0$. For a sufficient condition for uniform strong ergodicity in terms of the ergodicity coefficient see [4] or [5] pages 89–94.

REFERENCES

[1] BLACKWELL, D. and FREEDMAN, D. (1964). The tail σ-field of a Markov chain and a theorem of Orey. *Ann. Math. Statist.* **35** 1291–1295.

[2] DOBRUŠIN, R. L. (1956). Central limit theorem for nonstationary Markov chains I, II. *Theor. Probability Appl.* **1** 65–80 and 329–383 (English translation).

[3] HAJNAL, J. (1958). Weak ergodicity in nonhomogeneous Markov chains. *Proc. Cambridge Philos. Soc.* **54** 233–246.

[4] IOSIFESCU, M. (1966). Conditions nécessaires et suffisantes pour l'ergodicite uniforme des chaînes de Markoff variables et multiples. *Rev. Roumaine Math. Pures Appl.* **11** 325–330.

[5] IOSIFESCU, M. and THEODORESCU, R. (1969). *Random Processes and Learning.* Springer-Verlag, Berlin.

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

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

STATISTICAL LABORATORY
CAMBRIDGE CB2 ISB ENGLAND