Z. Wahrscheinlichkeitstheorie 2, 1—11 (1963)

# On the Strong Law of Large Numbers for a Class of Stochastic Processes

By

J. R. Blum,* D. L. Hanson,** and L. H. Koopmans**

## 1. Introduction and Summary

Let $(\Omega, \mathfrak{F}, P)$ be a probability space and let $\{X_n, n = 1, 2, \ldots\}$ be a sequence of real-valued random variables defined on $(\Omega, \mathfrak{F}, P)$. For each positive integer $n$ let $\sum_n$ be the smallest $\sigma$-algebra with respect to which $X_n$ is measurable and for $n \leqq m$ let $\sum_n^m$ be the smallest $\sigma$-algebra with respect to which $X_n, \ldots, X_m$ are jointly measurable.

**Definition.** The sequence $\{X_n\}$ will be called *-mixing if there exists a positive integer $N$ and a real-valued function $f$ defined for the integers $n \geqq N$ such that

i) $f$ is non-increasing with $\lim_{n\to\infty} f(n) = 0$, and

ii) if $n \geqq N$, $A \in \sum_1^m$, $B \in \sum_{m+n}$ then

$$|P(AB) - P(A)\,P(B)| \leqq f(n)\,P(A)\,P(B).$$

In section 2 we prove several versions of the strong law of large numbers for *-mixing sequences of random variables. Section 3 is devoted to a discussion of the *-mixing condition and examples of such sequences.

## 2. Strong laws of large numbers

For convenience we shall assume that $\Omega = \prod_{i=1}^{\infty} E_i$ where each $E_i$ is a copy of the real line and that $X_n$ is the $n$th coordinate projection. In that case $\sum_n^m$ is the $\sigma$-algebra of Borel sets on the coordinates $n$ through $m$. For each

$$n = 1, 2, \ldots \quad \text{let} \quad S_n = \sum_{i=1}^{n} X_i .$$

**Theorem 1.** *Let $\{X_n, n \geqq 1\}$ be a *-mixing stochastic process such that*

i) *the marginal distribution of $X_n$ is independent of $n$.*

ii) $EX_n = 0$.

iii) *the common marginal moment generating function of $X_n$ exists in a neighborhood of the origin.*

---

* Research supported by the National Science Foundation, Grant NSF — G 21219.
** Research supported by the Atomic Energy Commission.

*Let $\delta > 0$. There exists numbers $A$ and $\alpha$ with $A > 0$, $0 < \alpha < 1$, such that*

$$P\{\sup_{m \geqq n} |S_m/m| > \delta\} \leqq A\,\alpha^n \quad \textit{for} \quad n = 1, 2, \ldots.$$

*Proof.* We shall prove the existence of numbers $A, \alpha$ with $A > 0$ and $0 < \alpha < 1$ such that $P\{|S_n/n| > \delta\} \leqq A\alpha^n$ for all $n$. From this the conclusion of the theorem follows in straightforward fashion.

Let $P^*$ be the product measure induced on the process by the marginal distributions. Then it is known (see e.g. CHERNOFF [*1*], CRAMÉR [*2*]) that there exist numbers $B, \beta$ with $B > 0$ and $0 < \beta < 1$ such that $P^*\{|S_n/n| > \delta\} \leqq B\beta^n$ for all $n$. Choose an integer $k$ such that $k \geqq N$, where $N$ is the number occurring in the definition of *-mixing processes, and such that $[1 + f(k)]\beta < 1$. Let $j$ be an integer with $0 < j \leqq k$. Suppose $S$ is a set measurable with respect to the random variables $X_{ik+j}$, $i = 0, \ldots, n - 1$. Then we claim that

$$P(S) \leqq [1 + f(k)]^n P^*(S).$$

This follows at once from the definition of *-mixing processes when $S$ is a product of one-dimensional Borel sets, and consequently also for denumerable unions of disjoint sets of this kind. The inequality then follows from a familiar approximation argument.

Now write $n = km + r$ where $1 \leqq r \leqq k$. Then

$$S_n = \sum_{j=1}^{r} \sum_{i=0}^{m} X_{ik+j} + \sum_{j=r+1}^{k} \sum_{i=0}^{m-1} X_{ik+j}$$

and we have

$$\begin{aligned} P\{|S_n| > n\delta\} &\leqq \sum_{j=1}^{r} P\{|\sum_{i=0}^{m} X_{ik+j}| > (m+1)\delta\} + \sum_{j=r+1}^{k} P\{|\sum_{i=0}^{m-1} X_{ik+j}| > m\delta\} \\ &\leqq \sum_{j=1}^{r} [1 + f(k)]^{m+1} P^*\{|\sum_{i=0}^{m} X_{ik+j}| > (m+1)\delta\} \\ &+ \sum_{j=r+1}^{k} [1 + f(k)]^m P^*\{|\sum_{i=0}^{m-1} X_{ik+j}| > m\delta\} \\ &\leqq r[1 + f(k)]^{m+1} B\beta^{m+1} + (k - r)[1 + f(k)]^m B\beta^m \leqq kB\alpha^{km}, \end{aligned}$$

where $\alpha^k = [1 + f(k)]\beta$. Setting $A = kB/\alpha^k$ completes the proof.

Just as in the case of independent random variables the requirement of identical marginal distribution may be somewhat relaxed. We omit the details.

The remainder of the section is concerned with two standard forms of the strong law of large numbers. Before stating these we shall need some preliminary results. For $n = 1, 2, \ldots$ let $\mathfrak{S}_n$ be the family of sets each of which is a finite union of sets of the form $\{X_i > a_i, i = 1, \ldots, n\}$ where $a_1, \ldots, a_n$ are numbers with $-\infty \leqq a_i \leqq \infty$.

**Lemma 1.** *Let $A \in \mathfrak{S}_{n+1}$ There exists finitely many sets $A_0$ (= empty set) $\subset A_1 \subset \cdots \subset A_m$ with each $A_j \in \mathfrak{S}_n$ and sets $B_1, \ldots, B_m$ where $B_j = \{X_{n+1} > a_j\}$, $a_1 \leqq \cdots \leqq a_m$, such that*

$$A = \bigcup_{j=1}^{m} (A_j - A_{j-1}) \bigcap B_j.$$

*Proof.* Let

$$A = \bigcup_{i=1}^{m} \{X_j > a_{i,j}; j = 1, \ldots, n+1\} \quad \text{where} \quad a_{1,n+1} \leqq \cdots \leqq a_{m,n+1}.$$

Choose

$$B_i = \{X_{n+1} > a_{i,n+1}\} \quad \text{and} \quad A_i = \bigcup_{k=1}^{i} \{X_j > a_{k,j}; j = 1, \ldots, n\};$$
$$i = 1, \ldots, m.$$

**Lemma 2.** *Let $P$ and $Q$ be probability measures on $(\Omega, \mathfrak{F})$ satisfying*

i) $P\{X_i > a\} \leqq Q(X_i > a), \ -\infty < a, \ i = 1, 2, \ldots,$

ii) *$Q$ is a product measure,*

iii) *if $B \in \sum_1^{j-1}$ and $C = \{X_j > a\}$ then $P\{B \cap C\} \leqq P(B)\,Q(C)$* for $j = 1, 2, \ldots$.

*Then for every $n$, if $A \in \mathfrak{S}_n$ we have $P(A) \leqq Q(A)$.*

*Proof.* For $n = 1$ the lemma follows from i). Assume then that the conclusion of the lemma holds for $n < N$ and suppose $A \in \mathfrak{S}_N$. Choose $A_0, \ldots, A_m$; $B_1, \ldots, B_m$ in accordance with lemma 1. Then

$$P(A) = \sum_{j=1}^{m} P\{(A_j - A_{j-1}) \cap B_j\}.$$

Now
$$A_j - A_{j-1} \in \sum_1^{N-1} \quad \text{and} \quad B_j = \{X_N > a_j\}$$

for all $j$ and hence from iii) we have

$$P(A) \leqq \sum_{j=1}^{m} P(A_j - A_{j-1})\,Q(B_j) = P(A_m)\,Q(B_m) + \sum_{j=1}^{m-1} P(A_j)\,Q(B_j - B_{j+1}).$$

From the induction hypothesis we conclude

$$P(A) \leqq Q(A_m)\,Q(B_m) + \sum_{j=1}^{m-1} Q(A_j)\,Q(B_j - B_{j+1}) = Q(A),$$

since $Q$ is a product measure.

Let $\mathfrak{S}_n'$ be the collection of sets $A$ in $\sum_1^n$ satisfying: if $(X_1, \ldots, X_n) \in A$, and $a_i \geqq 0$, $i = 1, \ldots, n$ then $(X_1 + a_1, \ldots, X_n + a_n) \in A$.

**Lemma 3.** *Let $P$ and $Q$ be probability measures on $(\Omega, \mathfrak{F})$ satisfying the hypotheses of lemma 2. Then $P(A) \leqq Q(A)$ for every $A \in \mathfrak{S}_n'$ which is an open subset of Euclidean $n$-space.*

*Proof.* If $A \in \mathfrak{S}_n'$ and is open then

$$A = \bigcup_{j=1}^{\infty} \{X_i > a_{i,j}; i = 1, \ldots, n\}.$$

Let

$$A_N = \bigcup_{j=1}^{N} \{X_i > a_{i,j}; i = 1, \ldots, n\}.$$

Then $A_N \in \mathfrak{S}_N$ and hence $P(A_N) \leqq Q(A_N)$ for every positive integer $N$. Thus, by continuity, $P(A) \leqq Q(A)$.

Let $\{X_n, n \geqq 1\}$ be a *-mixing process and for each $n$ let $F_n$ be the distribution function of $X_n$. For $0 < \beta < 1$ let $\alpha_n$ be a "$\beta$-percentile" of $(1+\beta) F_n$, i.e. $\alpha_n$ is a number such that $(1+\beta) F_n(\alpha_n) \geqq \beta$ and $(1+\beta) F_n(\alpha_n - 0) \leqq \beta$. We define the distribution function $G_{n,\beta}$ by

$$G_{n,\beta}(x) = \begin{cases} 0 & \text{if } x < \alpha_n \\ (1+\beta) F_n(x) - \beta & \text{if } x \geqq \alpha_n . \end{cases}$$

Let $P_\beta$ be the infinite product measure generated by the sequence $\{G_{n,\beta}\}$.

**Lemma 4.**

i) *$P(X_n > a) \leqq P_\beta(X_n > a)$ for every number $a$ and every integer $n$.*

ii) *$P\{A \bigcap (X_{n+k} > a)\} \leqq P(A) P_\beta(X_{n+k} > a)$ for every $A \in \sum_1^n$, every number $a$, and every integer $k$ such that $f(k) < \beta$, where $f$ is the function occurring in the definition of *-mixing processes.*

The lemma follows easily from the definition of $P_\beta$.

**Theorem 2.** *Let $\{X_n, n \geqq 1\}$ be a *-mixing process such that $E X_n = 0$, $E X_n^2 < \infty$ for every $n$. Suppose*

i) *the random variables of the process are uniformly integrable and*

ii) $\sum_{n=1}^{\infty} E X_n^2/n^2 < \infty$.

*Then* $P\{\lim_{n\to\infty} S_n/n = 0\} = 1$.

*Proof.* Let $\beta$ be a positive number and choose $k$ so that $f(k) < \beta$. Let $P_\beta$ be defined as above and let $E_\beta\{\cdot\}$ be expectation with respect to the measure $P_\beta$. Now $E_\beta X_n^2 \leqq (1+\beta) E X_n^2$ so that $\sum_n E_\beta X_n^2/n^2 < \infty$. It follows from the known result for independent random variables that

$$P_\beta\{\lim_{n\to\infty} \sum_{i=1}^{n} [X_{ik+j} - E_\beta X_{ik+j}]/n = 0\} = 1$$

for every $j$ with $0 \leqq j < k$. Since for every $\delta > 0$ the set

$$\{\sup_{i_1 \leqq n \leqq i_2} \sum_{i=1}^{n} [X_{ik+j} - E_\beta X_{ik+j}]/n \leqq \delta\}$$

is the complement of an open set it follows from lemma 3 and lemma 4 that

$$P\{\limsup_{n\to\infty} \sum_{i=1}^{n} [X_{ik+j} - E_\beta X_{ik+j}]/n \leqq 0\} = 1$$

and consequently that

$$P\{\limsup_{n\to\infty} \sum_{i=1}^{n} [X_i - E_\beta X_i]/n \leqq 0\} = 1 .$$

If we can show that $\lim_{\beta\to 0} \sum_{i=1}^{n} E_\beta X_i/n = 0$ uniformly in $n$, then we will have $P\{\limsup_{n\to\infty} \frac{S_n}{n} \leqq 0\} = 1$. Choose positive numbers $\alpha$ and $\delta$ such that

$$\int_{-\alpha}^{\alpha} dF_i \geqq 1/2 \quad \text{and} \quad \int_{\{|x|>\alpha\}} |x| \, dF_i < \delta/2$$

uniformly in $i$, where $F_i$ is the distribution of $X_i$. Choose $\beta$ such that $0 < \beta < 1/2$ and $\alpha\beta < \delta/2$. Let $\tau_i$ be a "$\beta$-percentile" of $(1+\beta)F_i$ and for the sake of brevity assume that each $F_i$ is continuous at $\tau_i$. Then

$$E_\beta X_i = (1+\beta)\int_{\tau_i}^{\infty} x \, dF_i = -(1+\beta)\int_{-\infty}^{\tau_i} x \, dF_i$$

since $EX_i = 0$. Hence

$$|E_\beta X_i| \leqq (1+\beta)\Big[\int_{\left\{\substack{x<\tau_i \\ |x|>\alpha}\right\}} |x| \, dF_i + \int_{\left\{\substack{x<\tau_i \\ |x|\leqq\alpha}\right\}} |x| \, dF_i\Big] \leqq 3/2\,[\delta/2 + \alpha\beta] \leqq 3\,\delta/2 .$$

By considering the sequence $\{-X_n\}$ we get in exactly the same way

$$P\{\liminf_{n\to\infty} \frac{S_n}{n} \geqq 0\} = 1 \quad \text{so} \quad P\{\lim_{n\to\infty} \frac{S_n}{n} = 0\} = 1$$

completing the proof of the theorem.

**Lemma 5.** *Let $\{p_n, n \geqq 0\}$ and $\{q_n, n \geqq 0\}$ be sequences of nonnegative numbers. Then $\sum_n f_n p_n \leqq \sum_n f_n q_n$ for every nonnegative, nondecreasing sequence of numbers $\{f_n\}$ if and only if $\sum_{i\geqq k} p_i \leqq \sum_{i\geqq k} q_i$ for every integer $k \geqq 0$.*

*Proof.* Let $f_{n,k} = 0$ for $n = 0, \ldots, k-1$; $f_{n,k} = 1$ for $n \geqq k$. By choosing nonnegative linear combinations of such sequences we see that $\sum_n f_n p_n \leqq \sum_n f_n q_n$ under the hypothesis of the lemma for every nonnegative nondecreasing sequence $\{f_n\}$ with finitely many jumps. The sufficiency of the condition follows from a limiting argument. The necessity is obvious.

If $A$ is a set we shall denote its set characteristic function by $I_A$.

**Lemma 6.** *Let $\{A_n, n \geqq 1\}$ be a sequence of elements of $\mathfrak{F}$ such that the sequence $\{I_{A_n}\}$ is *-mixing, and let $A = \limsup_{n\to\infty} A_n$. Then $P(A) = 0$ unless $\sum_n P(A_n) = \infty$, and in that case $P(A) = 1$.*

*Proof.* If $\sum_n P(A_n) < \infty$ then $P(A) = 0$ in any case, hence assume $\sum_n P(A_n) = \infty$. Choose $\delta$ with $0 < \delta < 1$, a positive integer $k$, and an integer $j$ with $1 \leqq j \leqq k$ such that the *-mixing function $f(k) < \delta$ and such that $\sum_n P(A_{nk+j}) = \infty$. Let $B = \limsup_{n\to\infty} A_{nk+j}$. It is clearly sufficient to show that $P(B) = 1$. If this is not the case choose $m$ so that $P(\bigcup_{i=m}^{\infty} A_{ik+j}) < 1$. Then

$$\begin{aligned} P(\bigcup_{i=m}^{\infty} A_{ik+j}) &= P(A_{mk+j}) + P(A_{m+1,\,k+j} \cap A^c_{mk+j}) + \cdots \\ &\geqq (1-\delta)\{P(A_{mk+j}) + P(A_{(m+1)k+j})\,P(A^c_{mk+j}) + \cdots \\ &\geqq (1-\delta)\,P\}\,[\bigcup_{i=m}^{\infty} A_{ik+j}]^c\}\sum_{i=m}^{\infty} P(A_{ik+j}) \\ &= M\sum_{i=m}^{\infty} P(A_{ik+j}) \text{ where } M > 0. \text{ But this is impossible.} \end{aligned}$$

Let $\{X_n, n \geqq 1\}$ be a stochastic process. We define the sequences

$$\{p_{i,n}; i \geqq 0\} \quad \text{and} \quad \{s_{i,n}; i \geqq 0\} \quad \text{by} \quad p_{i,n} = P\{i \leqq |X_n| < i+1\}$$

and $s_{i,n} = P\{|X_n| \geqq i\} = \sum_{k \geqq i} p_{k,n}$.

**Theorem 3.** *Let $\{X_n, n \geqq 1\}$ be a *-mixing process such that $\sum_i \sup_n s_{i,n} < \infty$. Then $P\{\lim_{n\to\infty} \sum_{i=1}^{n} (X_{k_i} - E X_{k_i})/n = 0\} = 1$, for every increasing sequence of positive integers $\{k_n\}$.*

*Proof.* Since a subsequence of a *-mixing sequence is again *-mixing we shall assume that $k_n = n$ for all $n$. Define $p_i = \sup_n s_{i,n} - \sup_n s_{i+1,n}$. Then $p_i \geqq 0$ for all $i$, $\sum_i p_i = 1$, $\sum_i i p_i < \infty$, and $\sum_{i \geqq k} p_{i,n} \leqq \sum_{i \geqq k} p_i$ for every $n$ and $k$. Now define the *-mixing sequence $\{Y_n, n \geqq 1\}$ by

$$Y_n = \begin{cases} X_n & \text{if } |X_n| < n \\ n & \text{if } |X_n| \geqq n. \end{cases}$$

Then

$$E Y_n^2 \leqq \sum_{i=0}^{n-1} (i+1)^2 p_{i,n} + n^2 \sum_{i \geqq n} p_{i,n} \leqq \sum_{i=0}^{n-1} (i+1)^2 p_i + n^2 \sum_{i \geqq n} p_i$$

by lemma 5. From the properties of the sequence $\{p_i\}$ it is easily verified that $\sum_n E Y_n^2/n^2 < \infty$ and it follows from the hypothesis that the random variables $Y_n$ are uniformly integrable. Applying theorem 2 we obtain

$$P\{\lim_{n\to\infty} \sum_{i=1}^{n} [Y_i - E Y_i]/n = 0\} = 1 . \quad \text{Now} \quad E|X_n - Y_n|$$

$$\leqq \sum_{i \geqq n} (i+1) p_{i,n} + n \sum_{i \geqq n} p_{i,n} \leqq \sum_{i \geqq n} (i+1) p_i + n \sum_{i \geqq n} p_i$$

which approaches zero, and

$$\sum_n P\{X_n \neq Y_n\} = \sum_n s_{n,n} < \infty .$$

Thus $P\{X_n \neq Y_n \text{ infinitely often}\} = 0$ and the theorem is proved.

**Corollary.** *Let $\{X_n, n \geqq 1\}$ be a *-mixing process such that the distribution of $X_n$ is independent of $n$. Then $\sum_{i=1}^{n} X_{k_i}/n$ converges with probability one for every increasing sequence of positive integers $\{k_n\}$ if and only if $E|X_n| < \infty$. If this is the case then*

$$P\{\lim_{n\to\infty} \sum_{i=1}^{n} X_{k_i}/n = E X_n\} = 1 .$$

We omit the proof. Note that for independent random variables the corollary reduces to the well known Kolmogorov strong law of large numbers. We note that theorem 3 appears to be new even in the independent case.

## 3. Examples

We have yet to exhibit an example of a $*$-mixing proces. Clearly independent and $m$-dependent processes are $*$-mixing, and it is obvious that the condition imposes a strong form of asymptotic independence. Just how strong this form may be is demonstrated by Gaussian processes. As we shall show below, a Gaussian $*$-mixing process is always $m$-dependent for some nonnegative integer $m$.

**Lemma 7.** *Let $\{X_i, i = 1, 2\}$ be normal random variables with means $\mu_i$, variances $\sigma_i^2$, and covariance $\beta$. Then there exists a positive number c such that for every pair of Borel sets $A$, $B$*

$$|P\{X_1 \in A, X_2 \in B\} - P\{X_1 \in A\} P\{X_2 \in B\}| \leqq c P\{X_1 \in A\} P\{X_2 \in B\}$$

*if and only if $\beta = 0$.*

*Proof.* If $\beta = 0$ the inequality clearly holds for every $c \geqq 0$. Suppose then that the inequality holds for some $c > 0$ and assume for brevity that $\mu_i = 0$, $\sigma_i^2 = 1$ for $i = 1, 2$. From the inequality we easily deduce that

$$\left|\int f(X_1) g(X_2)\, dP - \int f(X_1)\, dP \int g(X_2)\, dP\right| \leqq c \int |f(X_1)|\, dP \int |g(X_2)|\, dP$$

whenever the integrals exist. Now let $t_1, t_2$ be real numbers. Then we have

$$\begin{aligned}\exp\{\tfrac{1}{2}[t_1^2 + t_2^2 + 2\beta t_1 t_2]\} &= E(e^{t_1 X_1 + t_2 X_2}) \leqq (1 + c) E(e^{t_1 X_1}) E(e^{t_2 X_2}) \\ &= (1 + c) \exp\{\tfrac{1}{2}[t_1^2 + t_2^2]\}.\end{aligned}$$

Choosing $t_1 = t_2 \operatorname{sign} \beta = t$ we have $|\beta| \leqq (1/t^2) \log(1 + c)$. Since $t$ is arbitrary we have $\beta = 0$.

As an immediate consequence of the lemma we have

**Theorem 4.** *Let $\{X_n, n \geqq 1\}$ be a Gaussian process. Then the process is $*$-mixing if and only if it is m-dependent for some nonnegative integer m.*

This result for stationary Gaussian processes but under a somewhat weaker form of asymptotic independence was stated by Ibragimov [*4*].

The remainder of the paper is devoted to exhibiting classes of processes which may be $*$-mixing without being $m$-dependent. To this end let $\Omega$ be a Borel subset of the real line and let $\mathfrak{S}$ be the $\sigma$-algebra of Borel sets relative to $\Omega$. Let $\{X_n, n \geqq 1\}$ be a stationary ergodic Markov process with state space $(\Omega, \mathfrak{S})$. Then the transition probabilities $P(x, A)$ may be taken to be regular, i.e. $P(x, A)$ is Borel measurable for fixed $A \in \mathfrak{S}$ and is a probability measure on $\mathfrak{S}$ for fixed $x \in \Omega$. The higher order transition probabilities are given by

$$\left.\begin{aligned} P^{(1)}(x, A) &= P(x, A) \\ P^{(m+n)}(x, A) &= \int P^{(m)}(y, A) P^{(n)}(x, dy) \end{aligned}\right| \begin{aligned} &x \in \Omega,\ A \in \mathfrak{S} \\ &m, n = 1, 2, \ldots, \end{aligned}$$

and the stationary probability satisfies

$$\prod(A) = \int P(y, A) \prod(dy), \quad A \in \mathfrak{S}$$

where we have changed to the notation that is commonly used in the theory of Markov processes (see e.g. Doob [*3*], p. 90). The probability measure on the process is then generated in the usual way by the relations

$$\begin{aligned} P\{X_{m+n} \varepsilon A \mid X_m = x, X_r = y_r, r < m\} &= P\{X_{m+n} \in A \mid X_m = x\} \\ &= P^{(n)}(x, A) \quad \text{and} \quad P\{X_n \in A\} = \prod(A). \end{aligned}$$

The *-mixing condition for Markov processes reduces to a condition on one-dimensional sets as is shown in

**Lemma 8.** *Let $m < n$ and $\varepsilon > 0$. If $|P(A \cap B) - P(A)P(B)| \leqq \varepsilon P(A)P(B)$ for $A \in \sum_m$, $B \in \sum_n$ then the same inequality holds for $A \in \sum_1^m$, $B \in \sum_n^{n+j}$ for an arbitrary positive integer $j$.*

*Proof.* Let $A \in \sum_m$, $B \in \sum_n$. Then

$$P(A \cap B) = \int_A p^{(n-m)}(y, B) \prod (dy) \leqq (1+\varepsilon) \int_A P(B) \prod (dy).$$

Since this inequality holds for all $A \in \sum_m$ we have $p^{(n-m)}(y, B) \leqq (1+\varepsilon) P(B) = (1+\varepsilon) \prod (B)$ for $B \in \sum_n$ and $y$ in a set whose complement has $\prod$ measure zero (a.e. $\prod$). Now choose sets $A_{m-k+1}, \ldots, A_m$; $B_n, \ldots, B_{n+j-1}$ with $A_i \in \sum_i$, $B_i \in \sum_i$ and let $A = \bigcap A_i$; $B = \bigcap B_i$. Then $P(A \cap B)$

$$\begin{aligned} &= \int_{A_{m-k+1}} \prod (dx_1) \int_{A_{m-k+2}} P(x_1, dx_2) \cdots \int_{A_m} P(x_{k-1}, dx_k) \int_{B_n} P^{(n-m)}(x_k, dy_1) \cdots \int_{B_{n+j-1}} P(y_{j-1}, dy_j) \\ &\leqq \int_{A_{m-k+1}} \prod (dx_1) \cdots \int_{A_m} P(x_{k-1}, dx_k) \int_{B_n} (1+\varepsilon) \prod (dy_1) \cdots \int_{B_{n+j-1}} P(y_{j-1}, dy_j) \\ &= (1+\varepsilon) P(A) P(B). \end{aligned}$$

By the usual measure extension arguments we find

$$P(A \cap B) \leqq (1+\varepsilon) P(A) P(B) \quad \text{for all} \quad A \in \sum_1^m, \quad B \in \sum_n^{n+k}$$

and a similar argument gives the inequality $P(A \cap B) \geqq (1-\varepsilon) P(A) P(B)$.

Lemma 8 enables us to state the *-mixing condition in terms of the stationary probability and the transition probabilities of the process as follows: There exists a positive integer $N$ and *-mixing function $f(k)$ defined for $k \geqq N$ such that

$$\left| \int_A P^{(k)}(x, B) \prod (dx) - \prod (A) \prod (B) \right| \leqq f(k) \prod (A) \prod (B)$$

for all $A, B \in \mathfrak{S}$ and $k \geqq N$. Thus we have

$$(1 - f(k)) \prod (B) \leqq P^{(k)}(x, B) \leqq (1 + f(k)) \prod (B)$$

a.e. $\prod$, and in particular we see that for $k \geqq N$ almost all transition probabilities are absolutely continuous with respect to $\prod$, i.e. there exist nonnegative functions $g^{(k)}(x, y)$ such that

$$P^{(k)}(x, B) = \int_B g^{(k)}(x, y) \prod (dy) \ \text{ a.e. } \prod .$$

Let $a^{(k)}(x, y) = g^{(k)}(x, y) - 1$, and let $\prod \times \prod$ be the product measure on $(\Omega \times \Omega, \mathfrak{S} \times \mathfrak{S})$ induced by $\prod$. With this notation we can state our basic result concerning *-mixing Markov processes.

**Theorem 5.** *Let $\{X_n, n = 0, \pm 1, \ldots\}$ be a stationary ergodic Markov process. Then the process is *-mixing if and only if there exists a positive integer $M$ and a number $\beta$ with $0 < \beta < 1$ such that*

i) $P^{(M)}(x, A)$ *is absolutely continuous with respect to* $\prod$ *a.e.* $\prod$ *and*

ii) $\prod \times \prod \{|a^{(M)}(x, y)| > \beta\} = 0$.

*If these conditions are satisfied then the process is exponentially *-mixing, i.e. there exist positive numbers* $C$ *and* $\gamma$ *with* $\gamma < 1$ *such that* $C\gamma^k$ *is a *-mixing function for the process for* $k \geqq M$.

*Proof.* Suppose the process is *-mixing. Choose an integer $M \geqq N$ and $\beta > 0$ such that $f(M) < \beta < 1$. The necessity of condition i) for $M \geqq N$ has already been discussed. As for ii) let $A, B \in \mathfrak{S}$. From the inequality $| P^{(M)}(x, B) - \prod(B)| \leqq f(M) \prod(B)$ it follows that

$$\Big|\int_A \int_B a^{(M)}(x, y) \prod(dx) \prod(dy\Big| \leqq f(M) \prod \times \prod(A \times B)$$

and by the previously mentioned extension argument we obtain

$$\Big|\int_U a^{(M)}(x, y) \prod(dx) \prod(dy)\Big| \leqq f(M) \prod \times \prod(U) \quad \text{for all} \quad U \in \mathfrak{S} \times \mathfrak{S},$$

and ii) follows from the fact that $f(M) < \beta$. To prove sufficiency let $M$ and $\beta$ be such that i) and ii) hold. If $k$ is a positive integer we have

$$\begin{aligned} P^{(M+k)}(x, B) &= \int_\Omega P^{(k)}(x, dy)\, P^{(M)}(y, B) \\ &= \int_B \Big[\int_\Omega P^{(k)}(x, dy)\, g^{(M)}(y, z)\Big] \prod(dz) \end{aligned}$$

and thus $P^{(n)}(x, B)$ is absolutely continuous with respect to $\prod$ a.e. $\prod$ for all $n \geqq M$. From the relations between the transition probabilities and the stationary probability we have for $m, n \geqq M$

$$\int a^{(n)}(x, y) \prod(dx) = \int a^{(n)}(x, y) \prod(dy) = 0,$$

each a.e. $\prod$ and consequently

$$a^{(m+n)}(x, y) = \int a^{(m)}(x, z)\, a^{(n)}(z, y) \prod(dz).$$

Thus if $|a^{(M)}(x, y)| \leqq \beta$ a.e. $\prod \times \prod$ then

$$|a^{(2M)}(x, y)| \leqq \int |a^{(M)}(x, z)|\, |a^{(M)}(z, y)| \prod(dz) \leqq \beta^2 \quad \text{a.e.} \quad \prod \times \prod$$

and more generally

$$|a^{(kM)}(x, y)| \leqq \beta^k \quad \text{a.e.} \quad \prod \times \prod \quad \text{for} \quad k = 1, 2, \ldots.$$

Now suppose $p^{(n)}(x, B)$ is absolutely continuous with respect to $\prod$ with density $g^{(n)}(x, y)$ and $k$ is a positive integer. Then

$$p^{(n+k)}(x, B) = \int P^{(n)}(y, B)\, P^{(k)}(x, dy) = \int_B \int g^{(n)}(y, z)\, P^{(k)}(x, dy) \prod(dz)$$

so that the density of $P^{(n+k)}(x, B)$ has the same bounds as $g^{(n)}(x, y)$. Consequently we have $|a^{(n)}(x, y)| \leqq \beta^{(n/M)-1}$ a.e. $\prod \times \prod$ for all $n \geqq M$. Now

$$\begin{aligned} &\Big|\int_A P^{(n)}(x, B) \prod(dx) - \prod(A) \prod(B)\Big| \\ &\leqq \int\int_{A \times B} |a^{(n)}(x, y)| \prod(dx) \prod(dy) \leqq \beta^{(n/M-1)} \prod(A) \prod(B) \end{aligned}$$

for $n \geqq M$, so that $C\gamma^n$ with $C = 1/\beta$ and $\gamma = \beta^{1/M}$ is a *-mixing function for the process. The theorem is proved.

We now turn to countable state space Markov processes. Let $\Omega$ be a subset of the positive integers,

$$\prod_j = \prod(\{j\}) \quad \text{and} \quad p^{(n)}(i,j) = P^{(n)}(i, \{j\}).$$

The *-mixing condition in this case becomes

$$\sup_{i,j \in \Omega} \left| \frac{P_{i,j}^{(M)} - \Pi_j}{\Pi_j} \right| \leqq \beta$$

for some positive integer $M$ and some number $\beta$ with $0 < \beta < 1$. Just as in the proof of theorem 5 it follows that this inequality must hold for all $p_{i,j}^{(n)}$ with $n \geqq M$ and consequently a necessary condition for a countable state space process to be *-mixing is that $p_{i,j}^{(n)} > 0$ for $n$ sufficiently large. Thus every such process is irreducible and aperiodic. It is also easily verified that in this case the condition of Doeblin (see e.g. Doob [3], p. 192) holds.

If $\Omega$ is finite, i.e. if the process is finite-state, then the process is *-mixing whenever it is irreducible and aperiodic. For in that case it is known that there are positive numbers $C$, $\gamma$ with $\gamma < 1$ such that for all $(i,j)$ with $\prod_i > 0$, $\prod_j > 0$ we have $|p_{i,j}^{(n)} - \prod_j| \leqq C\gamma^n$ for $n$ sufficiently large and the *-mixing inequality follows. One might conjecture that every countable state space process which is irreducible and aperiodic and which satisfies Doeblin's condition is also *-mixing. The following example shows that this is not the case. Let $\Omega = \{1, 2, \ldots\}$ and set

$$p_{i,j} = \begin{cases} 1/2 & \text{if } j = 1 \text{ or } j = i+1 \\ 0 & \text{otherwise}. \end{cases}$$

Then $\prod_j = 1/2^j > 0$ for all $j$. Yet for every positive integer $n$ there exist positive integers $i$ and $j$ such that $p_{i,j}^{(n)} = 0$. Consequently the process is not *-mixing yet obviously satisfies Doeblin's condition.

We conclude this section by exhibiting a fairly wide class of countable state space Markov processes which are *-mixing. To this end we note that the identity $\prod_j = \sum_i \prod_i p_{ij}$ implies $\inf_i p_{ij} \leqq \prod_j \leqq \sup_i p_{ij}$. From this we easily verify that the following condition is sufficient for the process to be *-mixing. There exists $\beta$ with $0 < \beta < 1$ such that for every $j = 1, 2, \ldots$ we have $\sup_i p_{i,j} \leqq (1 + \beta) \inf_i p_{i,j}$. Using this condition we may construct *-mixing processes by setting $p_{i,j} = p_j + (\delta_{i,j} - \delta_{i+1,j})\,\varepsilon_i$, $i, j$ positive integers where $\{p_j, j \geqq 1\}$ is a sequence of positive numbers with $\sum_j p_j = 1$, $\delta_{i,j}$ is the Kronecker delta, and for each positive integer $i$ we have $0 \leqq \varepsilon_i \leqq \min(1 - p_i, p_{i+1})$. Then $p_{i,j} \geqq 0$ for all $i$ and $j$ and $\sum_j p_{i,j} = 1$ for all $i$ so that the $p_{i,j}$'s form a set of transition probabilities. If in addition there exists $\beta$ with $0 < \beta < 1$ such that

$$[\varepsilon_{i+1} + \varepsilon_i(1 + \beta)]/\beta \leqq p_{i+1} \quad \text{for} \quad i = 0, 1, \ldots,$$

where we set $\varepsilon_0 = 0$, then the *-mixing condition holds.

As an example, let $\beta = 1/2$ and set $p_{i,j} = (1/2)^j + (\delta_{i,j} - \delta_{i+1,j})(1/2)^{n+i}$ where $n \geqq 3$ but otherwise arbitrary.

We conclude by briefly mentioning two other types of stochastic processes which may be *-mixing:

i) If $\{X_n\}$ is a *-mixing process and $F$ is a function of $k$ variables then the process $\{Y_n\}$ defined by $Y_n = F(X_{n+1}, \ldots, X_{n+k})$ is also *-mixing. Note that while $\{X_n\}$ may be a Markov process, $\{Y_n\}$ will in general not be Markovian.

ii) Lamperti and Suppes [5] have discussed a class of processes which they call "chains of infinite order". Under certain conditions such processes are also *-mixing.

## References


[1] Chernoff, H.: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. math. Statistics **23**, 493–507 (1952).

[2] Cramér, H.: Sur un nouveau théoreme-limite de la théorie des probabilités. Actual. sci. industr., No. 736, Paris (1938).

[3] Doob, J. L.: Stochastic Processes. New York: John Wiley and Sons 1953.

[4] Ibragimov, I. A.: Spectral functions of certain classes of stationary Gaussian processes. Doklady. Akad. Nauk. SSSR n. Ser. **137**, 1046–1048 (1961). Translated in Soviet Mathematics **2**, 403–405 (1961).

[5] Lamperti, J., and P. Suppes: Chains of infinite order and their application to learning theory. Technical Report No. 18, Applied Math. and Stat. Lab., Stanford Univ. (1958).



University of New Mexico
Dept. of Mathematics
Albuquerque, New Mexico (USA)
and
Sandia Corporation


*(Received September 6, 1962)*