Cowles Foundation Paper 604

J Appl Prob **21**, 930–934 (1984)
Printed in Israel
© *Applied Probability Trust* 1984

NON-STRONG MIXING AUTOREGRESSIVE PROCESSES

DONALD W K ANDREWS,* *Yale University*

Abstract

Certain first-order autoregressive processes are shown not to be strong mixing A direct proof is given This proof gives considerably more insight into the nature of the result than do proofs by contradiction The result and proof help to clarify the relation between the autoregressive and strong mixing conditions.

STRONG MIXING

1. Introduction

Let $\langle X_t \rangle \equiv (\cdots X_{t-1}, X_t, X_{t+1}, \cdots)$ be a sequence of random variables, and $\mathscr{X}_a^b$ be the σ-algebra generated by $(X_a, X_{a+1}, \cdots, X_b)$ for $-\infty \leqq a \leqq b \leqq \infty$. $\langle X_t \rangle$ is said to be *strong mixing* if

$$(1) \qquad \sup_{-\infty<t<\infty} \sup_{A \in \mathscr{X}_{-\infty}^t, B \in \mathscr{X}_{t+s}^{\infty}} |P(A \cap B) - P(A)P(B)| \equiv \alpha(s) \downarrow 0 \quad \text{as } s \to \infty.$$

The strong mixing condition was introduced by Rosenblatt in 1956 to prove the central limit theorem for 'weakly dependent' random variables. Since then it has assumed a position of considerable importance in probability theory. This is due to its tractability in the derivation of asymptotic properties of various functions of sequences of dependent random variables. Its areas of application are wide and include central limit theorems, strong laws of large numbers, laws of the iterated logarithm, empirical processes, order statistics, and robust estimators. Although the strong mixing condition has been widely adopted in the literature, a clear understanding of the condition itself is lacking. The categorization of well-known processes as strong mixing or non-strong mixing is still far from complete. However, Chanda [1] has shown that members of the important class of linear stochastic processes are strong mixing, provided they are based on innovation random variables which have Lebesgue-integrable characteristic

Received 9 November 1983; revision received 20 December 1983.
* Postal address: Department of Economics, Cowles Foundation for Research in Economics, Yale University, Box 2125, Yale Station, New Haven, CT 06520, USA.

functions. The latter condition is not superfluous. Ibragimov and Linnik [4] and Chernick [2] give examples of first-order autoregressive (AR(1)) processes based on discrete-valued innovation random variables which are not strong mixing. Unfortunately, their proofs are by contradiction, and do not give much insight into the reason why the strong mixing condition fails. In this brief note we show that certain AR(1) processes are *not* strong mixing by explicitly constructing sequences of sets which violate the strong mixing condition. The examples include non-strong mixing AR(1) processes with arbitrarily small autoregressive (AR) parameters. By direct proof of the non-strong mixing character of these processes it becomes clear why the strong mixing condition fails. Hopefully these results help in clarifying the nature of the strong mixing condition.

Note that the results given here relate to a comment made by Rosenblatt [5], p. 47, when first introducing the strong mixing condition:

> It would be of very great interest to see how much stronger the notion of a strong mixing condition is than that of an ordinary mixing condition [as defined, e.g., in Hannan [3]] in the case of a strictly stationary process.

All L^2 AR(1) processes are ordinary mixing (see, e.g., Hannan [3], Chapter IV, Theorem 3), but as shown here, not all L^2 AR(1) processes are strong mixing.

A brief description of the direct proof that certain AR(1) processes are non-strong mixing may be helpful. Suppose $\langle X_t \rangle$ is an AR(1) process based on Bernoulli (q) innovation random variables, and $X_{t,s}$ is equal to X_{t+s} minus its component which depends on $X_t, X_{t-1}, \cdots$. If we know X_t is small, then we know that with probability 1 X_{t+s} must fall in a set which is a small neighborhood of the support of $X_{t,s}$. A sequence of such small neighborhoods can be constructed for $s = 1, 2, \cdots$ which have unconditional probability bounded away from 1. Hence, knowledge that X_t is small increases the probability of certain sets which are determined by the 'future' of the process, no matter how far in the future, by a non-negligible amount. This implies $\langle X_t \rangle$ is non-strong mixing.

2. Results

Let $\langle \varepsilon_t \rangle$ be a doubly infinite sequence of independent Bernoulli (q) random variables. The AR(1) process $\langle X_t \rangle$ with innovation random variables $\langle \varepsilon_t \rangle$ and AR parameter ρ (in $(0, 1/2]$) is defined by

$$X_t = \sum_{l=0}^{\infty} \rho^l \varepsilon_{t-l}. \tag{2}$$

X_t is well defined since the series is absolutely convergent for all realizations of $\langle \varepsilon_t \rangle$. We shall exhibit sets A and B_s, $s = 1, 2, \cdots$, such that the conditions of the following lemma hold.

Lemma 1. If there exists a set A in $\mathscr{X}_{-\infty}^{t}$ with $P(A)>0$ and sets B_s in $\mathscr{X}_{t+s}^{\infty}$ with $P(B_s) \leqq k$ for all $s=1,2,\cdots$, for some constant $k<1$, such that

$$P(B_s \mid A)=1 \quad \text{for all } s,$$

then $\langle X_t \rangle$ is *not* strong mixing.

Comment. Clearly the notion that $\langle X_t \rangle$ is 'weakly dependent' is untenable if sets A and B_s exist. The occurrence of A alters the probability of B_s by at least $1-k$, for arbitrarily large s.

Proof of Lemma 1. We have

$$\alpha(s) \geqq P(A \cap B_s)-P(A)P(B_s)=P(A)[P(B_s \mid A)-P(B_s)] \geqq P(A)[1-k].$$

The latter expression is bounded away from 0 independently of s, so $\alpha(s) \not\to 0$ as $s \to \infty$. This completes the proof.

To find sets A and B_s as in Lemma 1 write $X_{t+s}=X_{t,s}+\rho^s X_t$, where

$$X_{t,s} \equiv \sum_{l=0}^{s-1} \rho^l \varepsilon_{t+s-l}.$$

$X_{t,s}$ is independent of X_t. Let W_s be the support of $X_{t,s}$, and w_j, $j=1,\cdots,J$, be the elements of W_s, where $w_j<w_{j+1}$ and $J \leqq 2^s$. If X_t takes a value x_t in $(0,\rho)$, then X_{t+s} must fall in the set $\bigcup_{j=1}^{J}(w_j, w_j+\rho^{s+1})$, since $0<\rho^s x_t<\rho^{s+1}$. Define

$$A=\{X_t \in (0,\rho)\}, \quad \text{and} \quad B_s=\left\{X_{t+s} \in \bigcup_{j=1}^{J}(w_j, w_j+\rho^{s+1})\right\}. \tag{3}$$

Note, if $\varepsilon_t=\varepsilon_{t-1}=\varepsilon_{t-2}=0$ and $\varepsilon_{t-3}=1$, then $X_t \in (0,\rho)$. Thus, $P(A)>0$. We have proved the following lemma.

Lemma 2. Let A and B_s be defined as in (3). If A occurs, then B_s occurs, $\forall s$. That is, $P(B_s \mid A)=1$, $\forall s$. Further, $P(A)>0$.

To get the desired non-strong mixing result it remains to show $P(B_s) \leqq k<1$, for all s. The upper bound, $w_j+\rho^{s+1}$, of the intervals in B_s has been chosen sufficiently small so that this result holds.

Theorem. For $A \in \mathscr{X}_{-\infty}^{t}$ and $B_s \in \mathscr{X}_{t+s}^{\infty}$, $s=1,2,\cdots$, as defined in (3),

$$|P(A \cap B_s)-P(A)P(B_s)|>P(A)[1-k]>0 \quad \forall s,$$

where $k \in (0,1)$ is independent of s. That is, the sets A and B_s violate the strong mixing condition (1), and the AR(1) process $\langle X_t \rangle$ based on Bernoulli (q) innovation random variables and AR parameter $\rho \in (0,1/2]$ is not strong mixing.

Comment. The condition $\rho \leqq 1/2$ facilitates the proof. One might conjecture that it is not a necessary condition for the theorem.

Proof of theorem. We introduce a set D (independent of s) such that $P(D) \equiv 1-k>0$ and $P(D \cap B_s)=0$, $\forall s$. Then,

$$P(B_s) \leqq P(D \cap B_s) + P(D^c) \leqq k < 1$$

(where D^c is the complement of D), as desired. The result of the theorem then follows by Lemmas 1 and 2.

D is defined as $\{X_t \in [\rho, 1]\}$. If $\varepsilon_t = 0$ and $\varepsilon_{t-1} = 1$, then $X_t \in [\rho, 1]$. Thus $P(D) > 0$.

To show $P(D \cap B_s) = 0$, $\forall s$, we *claim* the elements w_j, $j = 1, \cdots, J$, of the support W_s of $X_{t,s}$ are at least of distance ρ^{s-1} apart, provided $\rho \in (0, 1/2]$. Then, $X_t = x_t \in [\rho, 1]$ and $X_{t,s} = w_j$ imply $X_{t+s} = w_j + \rho^s x_t > w_j + \rho^{s+1}$ and $X_{t+s} = w_j + \rho^s x_t < w_{j+1}$. This holds for all j, so

$$P(D \cap B_s) \equiv P\left(X_t \in [\rho, 1],\ X_{t+s} \in \bigcup_{j=1}^{J} (w_j, w_j + \rho^{s+1})\right) = 0.$$

The claim is proved by induction. For sets G, G_1, and G_2 in R, let $d(G) = \inf\{|g_0 - g_1| : g_0, g_1 \in G\}$ and

$$d(G_1, G_2) = \inf\{|g_1 - g_2| : g_1 \in G_1, g_2 \in G_2\}.$$

For $s = 1$, $W_s = \{0, 1\}$, and $d(W_s) = 1 \geqq \rho^{s-1}$. Hence, the claim holds for $s = 1$. Suppose, for $s = v$,

$$d(W_s) \geqq \rho^{s-1}. \tag{4}$$

We now show (4) holds for $s = v + 1$,

$$W_{v+1} = W_v \cup (W_v + \rho^v)$$

where $W_v + \rho^v \equiv \{w + \rho^v : w \in W_v\}$, and

$$\begin{aligned} d(W_{v+1}) &= d(W_v) \wedge d(W_v + \rho^v) \wedge d(W_v, W_v + \rho^v) \\ &\geqq \rho^{v-1} \wedge \rho^{v-1} \wedge [\rho^v \wedge (\rho^{v-1} - \rho^v)] \text{ since (4) holds for } s = v, \\ &\geqq \rho^v \end{aligned}$$

since $\rho^{v-1} - \rho^v \geqq \rho^v$ for $\rho \in (0, 1/2]$. Hence, the claim holds, and the theorem is proved.

Acknowledgments

I would like to thank David B. Pollard, John J. Beggs, and the referee for their comments and suggestions.

References

[1] Chanda, K. C (1974) Strong mixing properties of linear stochastic processes *J Appl Prob* **11**, 401–408.

[2] Chernick, M R. (1981) A limit theorem for the maximum of autoregressive processes with uniform marginal distributions *Ann Prob* **9**, 145–149

[3] Hannan, E J (1970) *Multiple Time Series* Wiley, New York

[4] Ibragimov, I A. and Linnik, Yu V (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Nordhoff, Groningen

[5] Rosenblatt, M. (1956) A central limit theorem and a strong mixing condition *Proc Nat Acad Sci USA* **42**, 43–47.