The Annals of Probability
1981, Vol. 9, No. 4, 671–675

AN INVARIANCE PRINCIPLE FOR CERTAIN DEPENDENT SEQUENCES

BY C. M. NEWMAN[1] AND A. L. WRIGHT

University of Arizona

Let $X_1, X_2, \cdots$ be a strictly stationary second order sequence which is "associated"; i.e., is such that any two coordinatewise nondecreasing functions of the X_i's (of finite variance) are nonnegatively correlated. If $\sum_j \operatorname{Cov}(X_1, X_j) < \infty$, then the partial sum processes, $W_n(t)$, defined in the usual way so that $W_n(m/n) = (X_1 + \cdots + X_m - mE(X_1))/\sqrt{n}$ for $m = 1, 2, \cdots$, converge in distribution on $C[0, T]$ to a Wiener process. This result is based on two general theorems concerning associated random variables which are of independent interest.

Introduction. Following the suggestion of an invariance principle by Erdös and Kac (1946) and its proof by Donsker (1951) for sequences of independent random variables, there have been various extensions to the case of dependent variables. The necessary assumption of asymptotic independence has generally been given in terms of mixing conditions (see, e.g., Ibragimov (1962) or Theorem 20.1 of Billingsley (1968)) which can be difficult to verify in practice. It is the purpose of this paper to replace these mixing conditions for second order sequences by a simple and natural summability condition on the covariance; this is accomplished at the expense of restricting the nature of the dependence to that of "associated" random variables.

A finite collection of random variables, $X_1, \cdots, X_m$, is said to be associated if for any two coordinatewise nondecreasing functions f_1, f_2 on $\mathbb{R}^m$ such that $\tilde{f}_j \equiv f_j(X_1, \cdots, X_m)$ has finite variance for $j = 1, 2$, $\operatorname{Cov}(\tilde{f}_1, \tilde{f}_2) \geq 0$; an infinite collection is said to be associated if every finite subcollection is associated. This definition was introduced in Esary, Proschan and Walkup (1967) as an extension of the bivariate notion of positive quadrant dependence of Lehmann (1966). The basic concept however actually originated in Harris (1960) in the context of percolation models and was subsequently generalized in Fortuin, Kasteleyn and Ginibre (1971) and shown to apply to the Ising models of statistical mechanics; in the statistical mechanics literature which developed subsequently, associated random variables are said to satisfy the FKG inequalities. Much of the motivation for our present work is statistical mechanical in nature and some of the results appear in that context in Newman (1980). On a technical level, the relevance of the FKG inequalities to asymptotic independence (but not to a central limit theorem) was first pointed out in Lebowitz (1972) and the subsequent related work of Simon (1973).

In this paper, Theorems 1 and 2 are general results concerning associated random variables which are presented because of their use in the proof of our invariance principle, Theorem 3. In Newman and Wright (1980), martingale type inequalities related to Theorem 2 are derived and then used to obtain an invariance principle for two-parameter sequences of associated random variables. The problem of determining when random variables are associated is an interesting one which does not properly belong in the present paper; we refer the reader to Kemperman (1977) for criteria in terms of the joint distribution and note that independent variables are always associated.

Received March 10, 1980.

[1] Alfred P. Sloan Research Fellow, on leave from Indiana University. Research supported in part by NSF Grant MSC 77-20683.

AMS 1970 *subject classifications*. 60B10, 60F05.

Key words and phrases. Invariance principle, central limit theorem, associated random variables, FKG inequalities.

2. Some results about associated random variables. Before stating our first theorem, we note that if X, Y are associated, then

$$H(x, y) \equiv P(X > x, Y > y) - P(X > x)P(Y > y) \geq 0, \tag{1}$$

so that if X, Y also have finite variance, we can generalize an argument of Lehmann (1966) to obtain:

$$\begin{aligned} |\operatorname{Cov}(\exp(irX), \exp(isY))| &= \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ir)(is)\exp(irx + isy)H(x, y)\, dx\, dy \right| \\ &\leq |rs| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y)\, dx\, dy = |r|\,|s| \operatorname{Cov}(X, Y). \end{aligned} \tag{2}$$

We also point out that nondecreasing functions (such as positive linear combinations) of associated variables are associated.

THEOREM 1. *Suppose $X_1, \cdots, X_m$ are associated finite variance random variables with joint and marginal characteristic functions, $\phi(r_1, \cdots, r_m)$ and $\phi_j(r)$; then*

$$|\phi(r_1, \cdots, r_m) - \textstyle\prod_{j=1}^m \phi_j(r_j)| \leq \frac{1}{2} \sum\sum_{1 \leq j \neq k \leq m} |r_j|\,|r_k| \operatorname{Cov}(X_j, X_k). \tag{3}$$

PROOF. We proceed by induction on m. The result is true for $m = 1$ trivially and for $m = 2$ by (2); we suppose it is true for $m \leq M$. For $m = M + 1$, we may (by relabelling indices if necessary) assume that for some $\epsilon = \pm 1$, $\delta = \pm 1$ and $m' \in \{1, \cdots, M\}$, $\epsilon r_j \geq 0$ for $1 \leq j \leq m'$ while $\delta r_l \geq 0$ for $m' + 1 \leq l \leq M + 1$. We then define

$$Y_1 = \textstyle\sum_{j=1}^{m'} \epsilon r_j X_j, \qquad Y_2 = \sum_{j=m'+1}^{M+1} \delta r_j X_j \tag{4}$$

and note that Y_1 and Y_2 are associated. Denoting the joint characteristic function of Y_1 and Y_2 by ψ, the marginal characteristic functions by ψ_l ($l = 1, 2$), $\prod_{j=1}^{m'} \phi_j(r_j)$ by η_1, and $\prod_{j=m'+1}^{M+1} \phi_j(r_j)$ by η_2, we have that the left-hand side of (3) is bounded by

$$|\psi(\epsilon, \delta) - \psi_1(\epsilon)\psi_2(\delta)| + |\psi_1(\epsilon)| \cdot |\psi_2(\delta) - \eta_2| + |\psi_1(\epsilon) - \eta_1| \cdot |\eta_2|$$

which by the induction hypothesis is in turn bounded by

$$\begin{aligned} &|\epsilon|\,|\delta| \operatorname{Cov}(Y_1, Y_2) + \tfrac{1}{2} \textstyle\sum \sum_{m'+1 \leq j \neq k \leq M+1} |r_j|\,|r_k| \operatorname{Cov}(X_j, X_k) \\ &\qquad + \tfrac{1}{2} \textstyle\sum \sum_{1 \leq j \neq k \leq m'} |r_j|\,|r_k| \operatorname{Cov}(X_j, X_k) \end{aligned}$$

which equals the right hand side of (3). This completes the proof.

REMARK. Theorem 1 implies that associated random variables which are uncorrelated are jointly independent; a less elegant proof of this fact based on the techniques of Lebowitz (1972) and Simon (1973) is contained in Section III-0 of Wells (1977).

The next theorem will be used to provide the tightness needed for our invariance principle. Although its proof is not difficult, we have not been able to find this result stated in the literature even for the case of independent random variables.

THEOREM 2. *Suppose $X_1, \cdots, X_m$ are associated, mean zero, finite variance, random variables and $M_m = \max(S_1, S_2, \cdots, S_m)$ where $S_n = X_1 + \cdots + X_n$; then*

$$E(M_m^2) \leq \operatorname{Var}(S_m). \tag{5}$$

PROOF. We define $K_m = \min(X_2 + \cdots + X_m, X_3 + \cdots + X_m, \cdots, X_m, 0)$, $L_m = \max(X_2, X_2 + X_3, \cdots, X_2 + \cdots + X_m)$, $J_m = \max(0, L_m)$, and note that $K_m = X_2 + \cdots + X_m - J_m$ is a nondecreasing function of the X_i's so that $\operatorname{Cov}(X_1, K_m) \geq 0$, that $J_m^2 \leq L_m^2$ pointwise, and that $M_m = X_1 + J_m$; thus

$$
\begin{aligned}
E(M_m^2) = E((X_1 + J_m)^2) &= \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, J_m) + E(J_m^2) \\
&= \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, X_2 + \cdots + X_m) \\
&\quad - 2\,\mathrm{Cov}(X_1, K_m) + E(J_m^2) \\
&\le \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, X_2 + \cdots + X_m) + E(L_m^2).
\end{aligned} \tag{6}
$$

The proof is completed by induction on m since the induction hypothesis implies $E(L_m^2) \le \mathrm{Var}(X_2 + \cdots + X_m)$ which together with (6) yields (5).

REMARK. A slight variation of the above proof shows that (5) remains valid with M_m replaced by $S_{(j)}$, the jth order statistic of $(S_1, \cdots, S_m)$.

3. An invariance principle. The next theorem gives our invariance principle; the portion of the proof based on Theorem 1 can be applied to obtain a central limit theorem for certain associated random fields, $(X_i, i \in \mathbb{Z}^d\}$ (see Newman (1980)).

THEOREM 3. *Suppose $X_1, X_2, \cdots$ is a nondegenerate, strictly stationary, finite variance sequence which is associated and such that*

$$\sigma^2 \equiv \mathrm{Cov}(X_1, X_1) + 2 \textstyle\sum_{j=2}^{\infty} \mathrm{Cov}(X_1, X_j) < \infty. \tag{7}$$

For each $n = 1, 2, \cdots$, define the stochastic process

$$W_n(t) = [X_1 + \cdots + X_m + (nt - m)X_{m+1} - ntE(X_1)]/(\sigma\sqrt{n}), \qquad m/n \le t < (m+1)/n, \tag{8}$$

for $0 \le t \le T$; then the sequence of processes W_n converges in distribution (on $C[0, T]$) to the standard Wiener process.

PROOF. Without loss of generality, we may assume that $E(X_1) = 0$. We claim that it suffices to prove that

$$S_n/\sqrt{n} \to_{\mathscr{D}} N(0, \sigma^2) \tag{9}$$

where $S_n = X_1 + \cdots + X_n$. To see this, first note that it would follow from (9) by simple estimates that for $0 \le t_0 \le t_1 \le \cdots \le t_N \le T$,

$$U_{n,i} \equiv W_n(t_{i+1}) - W_n(t_i) \to_{\mathscr{D}} N(0, t_{i+1} - t_i) \quad \forall i, \tag{10}$$

and then by simple estimates based on (7) that

$$\mathrm{Cov}(U_{n,i}, U_{n,j}) \to 0 \quad \forall i \ne j.$$

Thus if $(U_1, \cdots, U_N)$ is a limit in distribution of any subsequence of $(U_{n,1}, \cdots, U_{n,N})$, the U_i's would be associated and uncorrelated random variables and hence independent by the remark following Theorem 1. This, together with (10), shows that the finite dimensional distributions of W_n converge to those of the standard Wiener process. We next define S_m^* as

$$S_m^* = \max(0, S_1, \cdots, S_m)$$

and note that for $\lambda_1 < \lambda_2$,

$$
\begin{aligned}
P(S_m^* \ge \lambda_2) &\le P(S_m \ge \lambda_1) + P(S_{m-1}^* \ge \lambda_2, S_{m-1}^* - S_m > \lambda_2 - \lambda_1) \\
&\le P(S_m \ge \lambda_1) + P(S_{m-1}^* \ge \lambda_2)P(S_{m-1}^* - S_m > \lambda_2 - \lambda_1) \\
&\le P(S_m \ge \lambda_1) + P(S_m^* \ge \lambda_2)E((S_{m-1}^* - S_m)^2)/(\lambda_2 - \lambda_1)^2,
\end{aligned}
$$

where the second inequality (compare equation (1) above) follows from the fact that S_{m-1}^*

and $S_m - S^*_{m-1}$ are associated since they are both non-decreasing functions of the X_i's. Now Theorem 2 with X_i replaced by $Y_i = -X_{m-i+1}$ yields that

$$E([S^*_{m-1} - S_m]^2) = E([\max(Y_1, Y_1 + Y_2, \cdots, Y_1 + \cdots + Y_m)]^2) \le E(S^2_m)$$

and thus we have, for $(\lambda_2 - \lambda_1)^2 \ge E(S^2_m) \equiv s^2_m$, that

$$P(S^*_m \ge \lambda_2) \le (1 - s^2_m/(\lambda_2 - \lambda_1)^2)^{-1} P(S_m \ge \lambda_1). \tag{11}$$

By adding to (11) the analogous inequality with each X_i replaced by $-X_i$, and by choosing $\lambda_2 = \lambda s_m$, $\lambda_1 = (\lambda - \sqrt{2})s_m$, we obtain

$$P(\max(|S_1|, \cdots, |S_m|) \ge \lambda s_m) \le 2P(|S_m| \ge (\lambda - \sqrt{2})s_m) \tag{12}$$

which yields by standard argument the needed tightness of the distributions of the W_n's to obtain the desired convergence in distribution (see the proof of Theorem 10.1 in Billingsley (1968)).

It remains to prove (9); i.e., to show that

$$\psi_n(r) \equiv E(\exp(ir\, S_n/\sqrt{n})) \to \exp(-\sigma^2 r^2/2). \tag{13}$$

For any *fixed* $l = 1, 2, \cdots$, we let $m = [n/l]$ (where $[\cdot]$ denotes the usual greatest integer function) and note that as $n \to \infty$,

$$|\psi_n(r) - \psi_{ml}(r)| \le |r| [\mathrm{Var}(S_n/\sqrt{n} - S_{ml}/\sqrt{ml})]^{1/2} \to 0. \tag{14}$$

We next define $Y^l_j = (S_{jl} - S_{(j-1)l})/\sqrt{l}$ for $j = 1, \cdots, m$ (with $S_0 = 0$) so that $S_{ml}/\sqrt{ml} = (Y^l_1 + \cdots + Y^l_m)/\sqrt{m}$; it then follows from Theorem 1 with $X_j = Y^l_j$ and $r_j = r/\sqrt{m}$, that

$$\begin{aligned} |\psi_{ml}(r) - (\psi_l(r/\sqrt{m}))^m| &\le \tfrac{1}{2} \textstyle\sum\sum_{1 \le j \ne k \le m} (r^2/m)\mathrm{Cov}(Y^l_j, Y^l_k) \\ &= (r^2/2)[\mathrm{Var}(S_{ml}/\sqrt{ml}) - \mathrm{Var}(S_l/\sqrt{l})] \to (r^2/2)(\sigma^2 - \sigma^2_l) \end{aligned} \tag{15}$$

where $\sigma^2_l \equiv Var(S_l/\sqrt{l})$ and we have used in the right hand convergence the fact that $\sigma^2_n \to \sigma^2$ which follows from (7). As in the proof of the standard central limit theorem, we have that

$$|(\psi_l(r/\sqrt{m}))^m - \exp(-\sigma^2_l r^2/2)| \to 0, \tag{16}$$

so that by combining (14), (15), and (16), we have for *any fixed* l,

$$\begin{aligned} &\limsup_{n\to\infty} |\psi_n(r) - \exp(-\sigma^2 r^2/2)| \\ &\qquad \le (r^2/2)(\sigma^2 - \sigma^2_l) + |\exp(-\sigma^2_l r^2/2) - \exp(-\sigma^2 r^2/2)|. \end{aligned} \tag{17}$$

Since $\sigma^2_l \to \sigma^2$ as $l \to \infty$, this yields (13) as desired and completes the proof of Theorem 3.

REMARK. Assume all the hypotheses of Theorem 3 except (7). Now let $K(R) = \mathrm{Cov}(X_1, X_1) + 2 \sum_{2 \le j \le R} \mathrm{Cov}(X_1, X_j)$ and suppose that $K(R) \to \infty$ as $R \to \infty$ but that K is slowly varying as $R \to \infty$ (see e.g. Section VIII-8 of Feller (1971)). We conjecture that in this case the conclusions of Theorem 3 remain valid providing the $\sigma\sqrt{n}$ normalization in (8) is replaced by $\sqrt{K(n) \cdot n}$.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

DONSKER, M. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6**.

ERDÖS, P., and KAC, M. (1946). On central limit theorems in the theory of probability. *Bull. Amer. Math. Soc.* **52** 292–302.

ESARY, J., PROSCHAN, F., and WALKUP, D. (1967). Association of random variables with applications. *Ann. Math. Statist.* **38** 1466–1474.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*. 2, 2nd Ed. Wiley, New York.

FORTUIN, C., KASTELEYN, P., and GINIBRE, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89–103.

HARRIS, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Camb. Phil. Soc.* **59** 13–20.

IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.* **7** 349–382.

KEMPERMAN, J. H. B. (1977). On the FKG-inequalities for measures on a partially ordered space. *Indag. Math.* **39** 313–331.

LEBOWITZ, J. (1972). Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems. *Comm. Math. Phys.* **28** 313–321.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

NEWMAN, C. M. (1980). Normal fluctuations and the FKG inequalities. To appear in *Comm. Math. Phys.* **74** 119–128.

NEWMAN, C. M. and WRIGHT, A. L. (1980). Associated random variables and martingale inequalities. Univ. of Arizona preprint.

SIMON, B. (1973). Correlation inequalities and the mass gap in $P(\phi)_2$, I. Domination by the two point function. *Comm. Math. Phys.* **31** 127–136.

WELLS, D. R. (1977). Some moment inequalities and a result on multivariate unimodality. Indiana Univ. Ph.D. Thesis.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721