The Annals of Probability
1974, Vol. 2, No. 4, 620-628

DEPENDENT CENTRAL LIMIT THEOREMS AND INVARIANCE PRINCIPLES

BY D. L. MCLEISH

University of Chicago

Central limit theorems are proved for martingales and near-martingales without the existence of moments or the full Lindeberg condition. These theorems are extended to invariance principles with a discussion of both random and nonrandom norming.

1. Introduction. In this paper we use a slightly modified classical approach, used by Salem and Zygmund (1947), to generate central limit theorems for martingales and variables which are close to being martingales. These are similar in type, though not in proof, to the results of Dvoretsky (1969), (1972) and Brown (1971), the main difference being that the Lindeberg condition used here (cf. 2.4) is weaker than in these papers, and finite second moments are generally not assumed (in fact many theorems, e.g., 2.7, 3.6, do not even require first moments). These theorems are extended to invariance principles in Section 3, similar to those of Brown (1971), and Drogin (1972), except that again the Lindeberg condition is not required. An invariance principle corresponding to the main theorem (2.2) of Dvoretsky (1972) is included as a corollary.

Let $\{X_{n,i}; i = 1, 2, \cdots, k_n\}$ be an array of random variables on the probability triple $(\Omega, \mathscr{F}, P)$. We denote $EX_{n,i}^2 = \sigma_{n,i}^2 \leqq \infty$, and $S_n = \sum_{i=1}^{k_n} X_{n,i}$. Unless otherwise indicated, summations will be over the range $1 \leqq i \leqq k_n$ (e.g., $S_n = \sum_i X_{n,i}$), and limits will be taken as $n \to \infty$. $I(A)$ or I_A will be used to denote the indicator of the set $A \in \mathscr{F}$. The L_p norms for random variables (i.e., $E^{1/p}|X|^p$) is denoted $\|X\|_p$, and the various kinds of convergence, in L_p, in probability, almost sure, and weak (in distribution) are denoted respectively $\to_{L_p}$, $\to_p$, $\to_{\text{a.s.}}$, and $\to_w$.

We will occasionally require the following conditions;

(1.1) The Lindeberg condition: for all $\varepsilon > 0$, $\sum_i \int_{|X_{n,i}|>\varepsilon} X_{n,i}^2 \, dP \to 0$, $n \to \infty$.

$$\sum_i \sigma_{n,i}^2 \to 1 \qquad \text{as } n \to \infty . \tag{1.2}$$

Now let $\{\mathscr{F}_{n,i}; 0 \leqq i \leqq k_n\}$ be any triangular array of sub-sigma fields of $\mathscr{F}$ such that for each n and $1 \leqq i \leqq k_n$, $X_{n,i}$ is $\mathscr{F}_{n,i}$-measurable and $\mathscr{F}_{n,i-1} \subset \mathscr{F}_{n,i}$. We will abuse notation slightly in the interests of brevity denoting $E(U \mid \mathscr{F}_{n,k})$ where U is some variable (e.g., $= X_{n,i}^2$) taken from the nth row by $E_k U$.

2. Central limit theorems. We begin with a useful basic theorem which has been implicit in a number of papers since Salem and Zygmund (1947). Let

Received March 7, 1973, revised November 21, 1973.

AMS 1970 *subject classifications*. Primary 60F05; Secondary 60G45.

Key words and phrases. Central limit theorem, martingales, dependence, invariance principle.

$\{X_{n,j};\ 1 \leqq j \leqq k_n\}$ and $\{\mathscr{F}_{n,j};\ 0 \leqq j \leqq k_n\}$ be defined as before, let i be the complex number ($i^2 = -1$), and let t be real. For each n, define a random variable by $T_n = \prod_{j=1}^{k_n} (1 + itX_{n,j})$.

(2.1) THEOREM. *Suppose for all real t,*

(a) $ET_n \to 1$,

(b) $\{T_n\}$ *is uniformly integrable,*

(c) $\sum_j X_{n,j}^2 \to_p 1$, *and*

(d) $\max_{j \leqq k_n} |X_{n,j}| \to_p 0$.

Then $S_n \to_w N(0, 1)$.

PROOF. We will use the appoximation $e^{ix} = (1 + ix)\exp\{-x^2/2 + r(x)\}$ where $|r(x)| \leqq |x|^3$ for $|x| < 1$ (verified by computing the series expansion of $\log(1 + z)$). Define a complex valued random variable by $I_n = e^{itS_n}$ and let $U_n = \exp\{-t^2/2 \sum_j X_{n,j}^2 + \sum_j r(X_{n,j}t)\}$. Then by the above approximation, $I_n = T_n e^{-t^2/2} + T_n(U_n - e^{-t^2/2})$ so in order to show $EI_n \to e^{-t^2/2}$, as is required by the method of characteristic function, we need only show, by (a), that

$$T_n(U_n - \exp(-t^2/2)) \to_{L_1} 0 . \tag{2.2}$$

But,

$$|\textstyle\sum_j r(X_{n,j}t)| \leqq |t|^3 \sum_j |X_{n,j}|^3 \leqq |t|^3(\max_{j \leqq k_n} |X_{n,j}|)(\sum_j X_{n,j}^2)$$

which converges in probability to 0 by (c) and (d). Combining this with (b), we have $T_n(U_n - \exp(-t^2/2)) = I_n - T_n \exp(-t^2/2)$ which is uniformly integrable (cf. Meyer Theorem 20; a convex combination of uniformly integrable sets is uniformly integrable; the uniform integrability of I_n follows from the fact that $E|I_n|^2 = 1$). Therefore, in view of the convergence in probability and the uniform integrability, (2.2) holds and the proof is complete.

We will call the array $X_{n,i}$ a martingale difference array (m.d.a.) with respect to $\mathscr{F}_{n,i}$ if each $X_{n,i}$ is $\mathscr{F}_{n,i}$-measurable and $E_{i-1}X_{n,i} = 0$ a.s. for all n and i.

(2.3) THEOREM. *Let $X_{n,i}$ be a martingale difference array satisfying*

(a) $\max_{i \leqq k_n} |X_{n,i}|$ *is uniformly bounded in L_2 norm,*

(b) $\max_{i \leqq k_n} |X_{n,i}| \to_p 0$, *and*

(c) $\sum_i X_{n,i}^2 \to_p 1$.

Then $S_n \to_w N(0, 1)$.

Let us examine briefly the conditions of this theorem. Condition (b) is a consequence of the Lindeberg condition assumed by most authors including Brown (1971), Drogin (1972) and Dvoretsky (1969). In fact, since $P\{\max_i |X_{n,i}| > \varepsilon\} = P[\sum_i X_{n,i}^2 I(|X_{n,i}| > \varepsilon) > \varepsilon^2]$, (b) is equivalent to the weaker version of the Lindeberg condition;

$$\textstyle\sum_i X_{n,i}^2 I(|X_{n,i}| > \varepsilon) \to_p 0 , \qquad \text{for all } \varepsilon > 0 . \tag{2.4}$$

Following 2.6, we furnish an example in which the stronger, L_1 convergence does not hold in (2.4).

Condition (a) is also a consequence of the Lindeberg condition since, for any $\varepsilon > 0$, $\max_{i \le k_n} X_{n,i}^2 \leqq \varepsilon^2 + \sum_{i=1}^{k_n} X_{n,i}^2 I(|X_{n,i}| > \varepsilon)$ and the second term on the majorant side converges in L_1 norm to 0.

Finally condition (c) replaces the convergence of the conditional variances of Dvoretsky, Brown, and others, by convergence of the sum of the squared variables. Such an assumption, similar to that made by Drogin (1972), is perhaps intuitively preferable in that it deals directly with the statistic normally used to approximate the asymptotic variance. In the proof of Corollary 3.8, we show that these two sums are frequently equivalent, and hence that the present theorems contain those of Brown and Dvoretsky.

PROOF OF (2.3). Define an array by $Z_{n,j} = X_{n,j} I(\sum_{k=1}^{j-1} X_{n,k}^2 \leqq 2)$. Observe that $Z_{n,j}$ is also a martingale difference array and that

$$P(Z_{n,j} \neq X_{n,j} \text{ for some } j \leqq k_n) \leqq P(\textstyle\sum_j X_{n,j}^2 > 2) \to 0 \tag{2.5}$$

as $n \to \infty$. Therefore, $Z_{n,j}$ also satisfies the conditions of (2.3). We will verify the conditions of Theorem (2.1) for this equivalent array. Let T_n be defined as in (2.1) but with $X_{n,j}$ replaced by $Z_{n,j}$. Then $ET_n = 1$.

For (2.1 b), define random variables by

$$\begin{aligned} J_n &= \min\{j \leqq k_n;\ \textstyle\sum_{i=1}^{j} X_{n,i}^2 > 2\} \qquad && \text{if} \quad \textstyle\sum_i X_{n,i}^2 > 2\,, \\ &= k_n && \text{otherwise.} \end{aligned}$$

Then

$$\begin{aligned} E|T_n|^2 &= E \textstyle\prod_{j=1}^{k_n} (1 + t^2 Z_{n,j}^2) \leqq E \exp(t^2 \sum_{j=1}^{J_n - 1} X_{n,j}^2)(1 + t^2 X_{n,J_n}^2) \\ &\leqq e^{2t^2}(1 + t^2 E X_{n,J_n}^2) \end{aligned}$$

which is bounded uniformly in n by (a). Therefore, $\{T_n\}$ is uniformly integrable.

The remaining two conditions of (2.1) hold by assumption. Therefore $\sum_j Z_{n,j} \to_w N(0, 1)$ and by (2.5), $S_n \to_w N(0, 1)$. □

(2.6) COROLLARY. *Suppose $X_{n,i}$ is any array satisfying the conditions* (2.3 a, b, c), *and in addition*;

(d) $\sum_i E_{i-1} X_{n,i} \to_p 0$, *and*
(e) $\sum_i E_{i-1}^2 X_{n,i} \to_p 0$.

Then $S_n \to_w N(0, 1)$.

PROOF. Let $Y_{n,i} = X_{n,i} - E_{i-1} X_{n,i}$. Clearly $Y_{n,i}$ is an m.d.a. and we will verify the conditions of Theorem (2.3) for this equivalent array. The demonstration that (2.3 b) holds is routine, and for (2.3 c) observe that

$$\textstyle\sum_i Y_{n,i}^2 = \sum_i X_{n,i}^2 - 2 \sum_i X_{n,i} E_{i-1} X_{n,i} + \sum_i E_{i-1}^2 X_{n,i}\,, \qquad \text{and}$$

so by (c) and (e), it is sufficient to show that the middle term on the right-hand side of this equality converges in probability to 0. But

$$|\textstyle\sum_i X_{n,i} E_{i-1} X_{n,i}| \leqq \sum_i |X_{n,i}||E_{i-1} X_{n,i}| \leqq (\sum_i X_{n,i}^2)^{\frac{1}{2}} (\sum_i E_{i-1}^2 X_{n,i})^{\frac{1}{2}}$$

by Schwartz's inequality, and this converges in probability to 0 by (c) and (e). Finally, to deal with (2.3a), observe that

$$\begin{aligned} E(\max_{i \le k_n} |Y_{n,i}|)^2 &= E(\max_{i \le k_n} (X_{n,i} - E_{i-1} X_{n,i})^2) \\ &\leqq 2E(\max_{i \le k_n} X_{n,i}^2) + 2E(\max_{i \le k_n} E_{i-1}^2 X_{n,i}) \,. \end{aligned}$$

The first term on the majorant side is bounded by assumption and for the second,

$$2 \max_{i \le k_n} (E_{i-1}|X_{n,i}|)^2 \leqq 2 \max_{i \le k_n} (E_{i-1} \max_{j \le k_n} |X_{n,j}|)^2 \,.$$

Now $\{E_{i-1} \max_{j \le k_n} |X_{n,j}|;\, i = 2, 3, \cdots, k_n\}$ is a martingale and therefore by Doob's inequality for submartingales (Doob, Theorem 3.4, page 317),

$$\begin{aligned} &E(\max_{i \le k_n} (E_{i-1} \max_{j \le k_n} |X_{n,j}|)^2) \\ &\qquad \leqq 2^2 E(E_{k_n-1} \max_{j \le k_n} |X_{n,j}|)^2 \\ &\qquad \leqq 2^2 E\{E_{k_n-1}(\max_{j \le k_n} |X_{n,j}|)^2\} \qquad \text{by Jensen's inequality} \\ &\qquad = 2^2 E(\max_{j \le k_n} |X_{n,j}|)^2 \end{aligned}$$

which is also uniformly bounded by assumption. Therefore, by Theorem (2.3), $\sum_i Y_{n,i} \to_w N(0, 1)$ and this, with (d), implies $S_n \to_w N(0, 1)$. □

For a simple example of variables satisfying the conditions of this corollary, but not the usual Lindeberg condition, take $(\Omega, \mathscr{F}, P)$ to be the Lebesgue unit interval with Lebesgue measure, $r_i(\omega)$ to be the Rademacher functions (i.e., $r_i(\omega) = +, -1$ as $[2^i\omega]$ is even, odd), put $X_{n,i}(\omega) = 2^{n/2} I(0 \leqq \omega \leqq 2^{-n}) + n^{-\frac{1}{2}} r_i(\omega)$, for $i \leqq n$, and $\mathscr{F}_{n,i} = \sigma(X_{n,j};\, j \leqq i)$.

(2.7) Corollary. *Suppose there exists a matrix of positive constants* $\{c_{n,i}\}$ *bounded strictly away from* 0 *and* ∞ *such that*:

(a) $\max_i |X_{n,i}| \to_p 0$,
(b) $\sum_i E_{i-1}(X_{n,i} I[|X_{n,i}| \leqq c_{n,i}]) \to_p 0$,
(c) $\sum_i E_{i-1}^2(X_{n,i} I[|X_{n,i}| \leqq c_{n,i}]) \to_p 0$, *and*
(d) $\sum_i X_{n,i}^2 \to_p 1$.

Then $S_n \to_w N(0, 1)$.

Proof. By (2.4d) is equivalent under (a) to the condition $\sum_i X_{n,i}^2 I(|X_{n,i}| \leqq c_{n,i}) \to_p 1$. The proof now follows by setting $Y_{n,i} = X_{n,i} I(|X_{n,i}| \leqq c_{n,i})$, and observing that $Y_{n,i}$ satisfies the conditions of (2.6), therefore converges to the standard normal distribution, and since, by (2.4), $P(\sum_i Y_{n,i} \neq \sum_i X_{n,i}) \to 0$, so too, does $\sum_i X_{n,i}$. □

This result contains most of the known martingale type central limit theorems including 2.1 and 2.2 of Dvoretsky (1972). The next two corollaries show how much weaker the assumptions may be when the array satisfies the norming condition (1.2) on the variances.

(2.8) Corollary. *Let* $X_{n,i}$ *be a martingale difference array satisfying* (1.2) *and* *the two conditions*:

$$\max_{i \leq k_n} |X_{n,i}| \to_p 0 \quad \textit{and} \tag{2.9}$$

$$\textit{for all} \quad \varepsilon > 0 , \qquad P(\textstyle\sum_i X_{n,i}^2 < 1 - \varepsilon) \to 0 . \tag{2.10}$$

Then $S_n \to_w N(0, 1)$.

Before undertaking the proof of Corollary (2.8) we will need an elementary lemma.

(2.11) LEMMA. *Let* X_n *and* X *be positive, integrable random variables such that* $P(X - X_n > \varepsilon) \to 0$, *for all* $\varepsilon > 0$, *and* $EX_n \to EX$. *Then* X_n *converges in* L_1 *to* X.

PROOF OF (2.11). The proof is entirely analogous to that of Scheffé's theorem (cf. Billingsley (1968) page 224) and proceeds by putting $Y_n = X - X_n - (EX - EX_n)$, $Y_n^+ = Y_n$ on the set where Y_n is positive, and zero elsewhere, and noting that, by hypothesis, $Y_n^+ \to_p 0$, for sufficiently large n, Y_n^+ is bounded above by $X + \varepsilon$. Therefore, $Y_n^+ \to_{L_1} 0$, by the Lebesgue dominated convergence theorem, and since $EY_n^- = EY_n^+$, we have $Y_n \to_{L_1} 0$. □

PROOF OF (2.8). By (2.10) and Lemma (2.11),

$$\textstyle\sum_i X_{n,i}^2 \to_{L_1} 1 , \tag{2.12}$$

and so conditions (a) and (c) of (2.3) are satisfied. The result follows from Theorem (2.3). □

(2.13) COROLLARY. *Suppose* $X_{n,i}$ *is a martingale difference array normalized by its variance* (1.2), *satisfying the Lindeberg condition* (1.1) *and the condition*:

$$\limsup_{n\to\infty} \textstyle\sum_{i \neq j} EX_{n,i}^2 X_{n,j}^2 \leqq 1 . \tag{2.14}$$

Then $S_n \to_w N(0, 1)$.

PROOF. By (2.4), (2.9) is implied by the Lindeberg condition. The only remaining condition of (2.8), viz. (2.10), is implied by the following lemma.

(2.15) LEMMA. *If* $\{X_{n,i}\}$ *satisfies* (1.2), *the Lindeberg condition, and* (2.14), *then* $\sum_i X_{n,i}^2 \to_{L_1} 1$.

PROOF. For arbitrary positive ε, put

$$Y_{n,i} = X_{n,i} I_{[|X_{n,i}| \leqq \varepsilon]} . \tag{2.16}$$

The Lindeberg condition implies that

$$\textstyle\sum_i X_{n,i}^2 - \sum_i Y_{n,i}^2 \to_{L_1} 0 \quad \text{and so} \tag{2.17}$$

$$||\textstyle\sum_i X_{n,i}^2 - 1||_1 \leqq ||\sum_i X_{n,i}^2 - \sum_i Y_{n,i}^2||_1 + ||\sum_i Y_{n,i}^2 - 1||_1 .$$

The first term on the right-hand side converges to 0 by (2.17) and the second is less than $(\sum_i EY_{n,i}^4 + \sum_{i \neq j} EY_{n,i}^2 Y_{n,j}^2 + 1 - 2 \sum_i EY_{n,i}^2)^{\frac{1}{2}}$. But since $EY_{n,i}^4 \leqq \varepsilon^2 EX_{n,i}^2$, $EY_{n,i}^2 Y_{n,j}^2 \leqq EX_{n,i}^2 X_{n,j}^2$, and $\sum_i EY_{n,i}^2 \to 1$ by (2.17), the limit of this expression is less than ε. Again, since ε was arbitrary, the limit of $||\sum_i X_{n,i}^2 - 1||_1$ must be 0. □

3. Invariance principles. There are a number of invariance principles for martingales; those of Brown (1971), and Drogin (1972), are outstanding examples. Both of these assume the Lindeberg condition, and in this section we prove some corresponding results when the Lindeberg condition is weakened to (2.4). As a corollary, we obtain the central limit theorem (2.2) of Dvoretsky (1972) and its corresponding invariance principle.

Let J be an interval either of the form $[0, T]$ for some $T < \infty$ or the half line $[0, \infty)$. Let $D(J)$ be the space of right continuous real valued functions on J, endowed with the Skorohod J_1 topology. We wish to discuss weak convergence (denoted $\rightarrow_w$) of random functions with values in this space (for compact J, cf. Billingsley, Chapter 3). For semi-infinite J, T will denote an arbitrary real, $0 < T < \infty$, and observe that by Stone (1963) page 695, it is sufficient to show weak convergence of the functions restricted to each finite interval $[0, T]$.

We now consider a doubly infinite array of variables $\{X_{n,i}\}$ and a sequence of integer valued, non-decreasing, right continuous functions defined on J, $k_n(t)$ such that $k_n(0) = 0$ for all n. We form a random function

$$W_n(t) = \sum_{i=1}^{k_n(t)} X_{n,i} \qquad \text{for } t \in J \tag{3.1}$$

(of course the sum over an empty index set is 0). Observe that W_n is a right continuous step function, a measurable random element of $D(J)$, and $W_n(0) = 0$. The object of the invariance principle is to show that W_n converges weakly to the standard Brownian motion process W on $D(J)$.

(3.2) THEOREM. *Suppose $X_{n,i}$ is a martingale difference array satisfying;*

(a) $\max_{i \leq k_n(t)} |X_{n,i}| \rightarrow_{L_2} 0$, *and*
(b) $\sum_{i=1}^{k_n(t)} X_{n,i}^2 \rightarrow_p t$, *for each* $t \in J$.

Then $W_n \rightarrow_w W$ on $D(J)$.

PROOF. Put $\bar{X}_{n,i} = X_{n,i} I(\sum_{j=1}^{i-1} X_{n,j}^2 \leqq T + 1)$. Then $P(X_{n,i} \neq \bar{X}_{n,i}$ for some $i \leqq k_n(T)) \rightarrow 0$, and $\sum_i \bar{X}_{n,i}^2 \leqq T + 1 + \max_{i \leq k_n(T)} X_{n,i}^2$ which is uniformly integrable by (a). Therefore $\bar{X}_{n,i}$ is an equivalent m.d.a. satisfying;

$$\max_{i \leq k_n(T)} |\bar{X}_{n,i}| \rightarrow_p 0\,, \tag{3.3}$$

and

$$\sum_{i=1}^{k_n(t)} \bar{X}_{n,i}^2 \rightarrow_{L_1} t \qquad \text{for each } t\,. \tag{3.4}$$

We will show that $\bar{W}_n(t) = \sum_{i=1}^{k_n(t)} X_{n,i} \rightarrow_w W$ on $D[0, T]$, from which it will follow, since $P(W_n \neq \bar{W}_n) \rightarrow 0$, that $W_n \rightarrow_w W$ on $D[0, T]$.

Such an invariance principle, under these conditions (3.3) and (3.4), is similar to Theorem 3 of Brown (1971), for by (2.4), (3.3) is equivalent under (3.4) to the Lindeberg condition. For this reason we only sketch the proof. The finite dimensional distributions of $\bar{W}_n$ are shown to converge to the corresponding finite dimensional distributions of W by the Cramér–Wold technique; if $u_1, u_2, \cdots, u_m$ are arbitrary real numbers and $0 = t_0 < t_1 \cdots < t_m$ elements of

J, we put

$$Y_{n,i} = u_j \bar{X}_{n,i}$$
$$= 0 \qquad \text{if} \quad k_n(t_m) < i\,, \qquad \text{where } j = \inf(1;\, k_n(t_l) \geqq i)\,.$$

The conditions of Corollary (2.8) are easily verified for the array $Y_{n,i}$ (with the unit variance replaced by $\sigma^2 = \sum_{i=1}^m u_i^2(t_i - t_{i-1})$) and thus $\sum_i Y_{n,i} = \sum_{i=1}^m u_i[\bar{W}_n(t_i) - \bar{W}_n(t_{i-1})]$ converges to the normal law with mean 0 and variance σ^2.

For tightness (cf. Stone (1963) page 695) it is sufficient to prove for each positive ε,

(3.5) $\qquad \lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{|s-t|<\delta;\, s,t\leqq T} |\bar{W}_n(s) - \bar{W}_n(t)| > \varepsilon) = 0\,.$

This is done with only minor changes in the proof of Theorem 3, Brown (1971). □

The following invariance principle corresponds to Theorem (2.7).

(3.6) THEOREM. *Suppose* $c_{n,i}$ *is an array of positive constants bounded away from* 0 *and* ∞ *and*:

(a) $\max_{i\leqq k_n(t)} |X_{n,i}| \to_p 0$,
(b) $\sum_{i=1}^{k_n(t)} |E_{i-1}X_{n,i}I(|X_{n,i}| \leqq c_{n,i})| \to_p 0$, *and*
(c) $\sum_{i=1}^{k_n(t)} X_{n,i}^2 \to_p t$, *for each* $t \in J$.

Then $W_n \to_w W$ *on* $D(J)$.

PROOF. Put $Y_{n,i} = X_{n,i}I(|X_{n,i}| \leqq c_{n,i})$, and observe that, by (2.4) and (3.6c),

(3.7) $\qquad P(Y_{n,i} \neq X_{n,i}$ for some $i \leqq k_n(T)) \to 0 \qquad$ and $\qquad \sum_{i=1}^{k_n(t)} Y_{n,i}^2 \to_p t\,.$

Therefore, by essentially the same argument as that used in the proof of (2.6), the array $Y_{n,i} - E_{i-1}Y_{n,i}$ satisfies the conditions of Theorem (3.2), hence $\sum_{i=1}^{k_n(t)} (Y_{n,i} - E_{i-1}Y_{n,i}) \to_w W(t)$. But since (b) implies $\sup_{s\leqq T} \sum_{i=1}^{k_n(s)} E_{i-1}Y_{n,i} \to_p 0$, it follows (cf. Billingsley, Theorems 4.1 and 5.1) that $\sum_{i=1}^{k_n(t)} Y_{n,i} \to_w W$. The convergence of W_n to W now follows from this and (3.7). □

We now give an invariance principle corresponding to a central limit theorem of Dvoretsky [(1972) Theorem 2.2].

(3.8) COROLLARY. *Let* $X_{n,i}$ *be an array satisfying*;

(3.9) *The conditional Lindeberg*: $\sum_{i=1}^{k_n(t)} E_{i-1}X_{n,i}^2 I(|X_{n,i}| > \varepsilon) \to_p 0$ *for all* $\varepsilon > 0$,

(3.10) $\qquad \sum_{i=1}^{k_n(t)} E_{i-1}X_{n,i}^2 \to_p t\,,$ *and*

(3.11) $\qquad \sum_{i=1}^{k_n(t)} |E_{i-1}X_{n,i}| \to_p 0\,,$ *for all* $t \in J$.

Then $W_n \to_w W$ *on* $D(J)$.

PROOF. We verify the conditions of (3.6).

(a) From Lemma 3.5 of Dvoretsky (1972), for all $\varepsilon, \eta > 0$,

(3.12) $\qquad P(\max_i |X_{n,i}| > \varepsilon) \leqq \eta + P\{\sum_i P(|X_{n,i}| > \varepsilon \mid \mathscr{F}_{n,i-1}) > \eta\}\,.$

Applying the conditional form of Chebyschev's inequality, and then (3.9), the limit of this as $n \to \infty$ is less than η, and since η is arbitrary, this limit must be 0.

(b) Pick any $0 < c < \infty$. It is clearly sufficient by (3.11) to show $\sum_{i=1}^{k_n(t)} |E_{i-1}X_{n,i}I(|X_{n,i}| > c)| \to_p 0$. Again this follows from (3.9) and the inequality $|E_{i-1}X_{n,i}I(|X_{n,i}| > c)| \leqq c^{-1}E_{i-1}X_{n,i}^2 I(|X_{n,i}| > c)$.

(c) For arbitrary $\varepsilon > 0$ and $t \in J$, put

$$W_{n,i} = X_{n,i}I(|X_{n,i}| \leqq \varepsilon,\ \textstyle\sum_{j=1}^{i} E_{j-1}X_{n,j}^2 \leqq t + 1)$$

and observe that both

$$\text{(3.13)} \qquad P(X_{n,i} \neq W_{n,i} \text{ for some } i \leqq k_n(t)) \to 0\,, \qquad \text{and}$$
$$\textstyle\sum_{i=1}^{k_n(t)} E_{i-1}(X_{n,i}^2 - W_{n,i}^2) \to_p 0 \qquad \text{as } n \to \infty\,,$$

by (3.9), (3.10) and (3.12). Moreover,

$$\text{(3.14)} \qquad E(\textstyle\sum_{i=1}^{k_n(t)} W_{n,i}^2 - E_{i-1}W_{n,i}^2)^2 = \sum_{i=1}^{k_n(t)} (EW_{n,i}^4 - EE_{i-1}^2 W_{n,i}^2)$$
$$\leqq \varepsilon^2 E(\textstyle\sum_{i=1}^{k_n(t)} E_{i-1}W_{n,i}^2) \leqq \varepsilon^2(t+1)\,.$$

Consequently for $\eta > 0$, by (3.13) and (3.14), $\limsup_{n\to\infty} P\{\sum_{i=1}^{k_n(t)} (X_{n,i}^2 - E_{i-1}X_{n,i}^2) > \eta\} \leqq \varepsilon^2(t+1)/\eta^2$, which limit, since ε is arbitrary, must be 0. Therefore (3.6c) follows from (3.10). $\square$

The last part of the proof of Corollary (3.8) is of independent interest. With minor modifications we can show that if $\{X_{n,i}\}$ satisfies the conditional Lindeberg (3.9), and if $\limsup_{n\to\infty} P(\sum_i E_{i-1}X_{n,i}^2 > K)$ is $o(1)$ as $K \to \infty$, then

$$\text{(3.15)} \qquad \max\nolimits_{j \leq k_n(t)} \textstyle\sum_{i=1}^{j} (X_{n,i}^2 - E_{i-1}X_{n,i}^2) \to_p 0\,.$$

This result was obtained by Drogin [(1972) Theorem 1] under the Lindeberg condition. In a number of papers (e.g., Drogin, Lévy), random norming is used—that is, $k_n(t)$ permitted to be a random function. For example, we may define k_n by

$$\text{(3.16)} \qquad k_n(t) = \inf\{j;\ \textstyle\sum_{i=1}^{j-1} X_{n,i}^2 > t\} \qquad \text{or}$$

$$\text{(3.17)} \qquad k_n(t) = \inf\{j;\ \textstyle\sum_{i=1}^{j} E_{i-1}X_{n,i}^2 > t\}\,.$$

The effect of (3.15) is to demonstrate that these two normings are equivalent under fairly weak conditions.

The theorems of this section remain valid also when $k_n(t)$ is, for each n and t, a stopping time with respect to $\{\mathscr{F}_{n,i};\ i = 1, 2, \cdots\}$. Certainly the theorems and proof of Section 2 remain essentially unchanged if we allow $k_n = \infty$. Note, then that the theorems can be verified when k_n is a stopping time by replacing $X_{n,i}$ by $X_{n,i}I(i \leqq k_n)$, for which array the conditions (e.g., of (2.7)) still hold. The theorems of Section 3 can be similarly extended; for example, the finite dimensional convergence in (3.2) carries over unchanged. Finally observe that if k_n is defined as in (3.16), (3.6c) is implied by (3.6a) and similarly if it is defined as in (3.17), (3.10) is implied by (3.9). Thus if (3.6a, b) hold for $k_n(t)$ defined by (3.16), or if (3.9) and (3.11) hold for $k_n(t)$ defined by (3.17) then

$W_n \to_w W$. These results are similar to Drogin (1972), except that the Lindeberg condition is weakened to convergence in probability (2.4).

It is perhaps worthy of note that when the triangular array $X_{n,i}$ arises in the usual way from a sequence of random variables, then $\mathscr{F}_{n,i} \subset \mathscr{F}_{n-1,i}$ for all n, $i \geqq 2$. Under this condition it is easy to show that all of the above invariance principles are mixing in the sense of Rényi (1958), and hence that we can pass directly from nonrandom to random sum central limit theorems and invariance principles as is done in Billingsley (1968), Theorem 17.2.

Acknowledgment. I am deeply indebted to the referee for a number of valuable contributions, particularly for pointing out and proving the equivalence of the martingale and non-martingale case (Corollary (2.6)).

REFERENCES

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

[2] BROWN, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.

[3] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] DROGIN, R. (1972). An invariance principle for martingales. *Ann. Math. Statist.* **43** 602–620.

[5] DVORETSKY, A. (1969). Central limit theorems for dependent random variables and some applications. Abstract, *Ann. Math. Statist.* **40** 1971.

[6] DVORETSKY, A. (1972). Central limit theorems for dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press, 513–535.

[7] LÉVY, P. (1954). *Le Mouvement Brownien*. Gauthier Villars, Paris.

[8] MEYER, P. A. (1966). *Probability and Potentials*. Blaisdell, Waltham.

[9] RÉNYI, A. (1958). On mixing sequences of random variables. *Acta. Math. Acad. Sci. Hungar.* **9** 389–393.

[10] SALEM, R. and ZYGMUND, A. (1947). On lacunary trigonometric series I. *Proc. Nat. Acad. Sci. U.S.A.* **33** 333–338.

[11] STONE, C. (1963). Weak convergence of stochastic processes defined on semi-infinite time intervals. *Proc. Amer. Math. Soc.* **14** 694–696.

DEPARTMENT OF MATHEMATICS
YORK UNIVERSITY
DOWNSVIEW, ONTARIO
CANADA