The Annals of Probability
1978, Vol. 6, No. 1, 144–150

AN EXTENDED MARTINGALE INVARIANCE PRINCIPLE

BY D. L. MCLEISH
University of Alberta

In this note the conditions on an invariance principle for triangular arrays of random variables contained in an earlier paper are weakened. Random norming by functions which are not stopping times is permitted, the L^2-boundedness conditions on the maximum of the summands relaxed, and joint convergence with an arbitrary sequence of random elements of some other metric space proved.

1. In an earlier (1974) paper, the author proves some invariance principles for martingales and near martingales. Specifically random variables X_{ni}; $i = 1, 2, \cdots, k_n(t)$ are considered on some arbitrary probability spaec $(\Omega, \mathscr{F}, P)$ (which we there and here assume without loss of generality is sufficiently rich to define all the random elements we require) and $k_n(t)$ is a sequence of nondecreasing right-continuous nonnegative integer-valued functions with $k_n(0) = 0$. The purpose of the invariance principle is to show the sequence of functions $\sum_{i=1}^{k_n(t)} X_{ni}$ converges (as $n \to \infty$) weakly to a standard Brownian motion process. This is done under the conditions that X_{ni}; $i = 1, 2, \cdots$ is a martingale difference sequence with respect to sigma-algebras $\mathscr{F}_{ni}$; $i = 1, 2, \cdots$ adapted to it, that the sample variances $\sum_{i=1}^{k_n(t)} X_{ni}^2$ converge in probability to the appropriate value t and that $E[\max_{i \leqq k_n(t)} X_{ni}^2] \to 0$. This last condition or any such condition requiring second moments seems unnecessary for a theorem proving weak covergence and is dispensed with in the present paper. It is observed also in the earlier paper that $k_n(t)$ may be replaced by a sequence of random stopping times without changing the proofs. In the present paper we drop the conditions that they be stopping times and prove joint convergence of the above random function with some sequence of weakly convergent random elements of any other separable metric space. The present theorems also do not require that the X_{ni} are martingale differences or that any moments exist for them, only that variables suitably truncated are reasonably close to being martingale differences.

Throughout this paper, $I(A)$ will denote the indicator random variable of the set $A \subset \mathscr{F}$, and convergence in probability, almost surely, and weakly (in distribution) are denoted $\to_p$, $\to_{\text{a.s.}}$, and $\to_{\mathscr{D}}$ respectively, and the sigma-algebra generated by a random vector U is denoted $\sigma(U)$.

For the required theory of weak convergence, see Billingsley (1968) and Stone (1963).

Let $D[0, \infty)$ be the space of right-continuous real-valued functions on $[0, \infty)$

Received January 24, 1977; revised April 27, 1977.
AMS 1970 *sbject classifications.* Primary 60F05; Secondary 60G45.
Key words and phrases. Invariance principle, Donsker's theorem, central limit theorem, martingales.

endowed with Stone's (1963) separable metric topology. Let X_{ni}, $i = 1, 2, 3, \cdots, n = 1, 2, 3, \cdots$ be an infinite array of random variables on the probability triple $(\Omega, \mathscr{F}, P)$ such that $\sum_i X_{n,i}^2 = \infty$ a.s. and let $\mathscr{F}_{n,i}$, $i = 1, 2, 3, \cdots, n = 1, 2, 3, \cdots$ be an infinite array of sigma-algebras contained in $\mathscr{F}$ such that for each n, i, $\mathscr{F}_{n,i} \subset \mathscr{F}_{n,i+1}$ and $X_{n,i}$ is measurable with respect to $\mathscr{F}_{n,i}$. We will denote the conditional expectation with respect to $\mathscr{F}_{n,i}$ by $E_i(\cdot)$ (the index n to be understood). Let $k_n(t)$, $n = 1, 2, 3, \cdots$ be random elements of $D[0, \infty)$ which are almost surely nondecreasing and integer-valued with $k_n(0) = 0$. Define $W_n(t) = \sum_{i=1}^{k_n(t)} X_{n,i}$ for each n and for each $t \geqq 0$ and observe that $W_n(t)$ is a random element of $D[0, \infty)$. Let $W(t)$ denote a standard Brownian motion process on $D[0, \infty)$. Finally assume U_n, $n = 1, 2, 3, \cdots, U$ are random elements of an arbitrary separable metric space M and $U_n \to_{\mathscr{D}} U$, where for convenience we choose U independent of all $\mathscr{F}_{n,i}$ and W. Suppose $\mathscr{A}$ is an algebra or semi-algebra of U-continuity subsets of M which generate the Borel sets in the support of U. Then the main result of this note is the following theorem.

THEOREM 1. *Assume for each* $A \in \mathscr{A}$ *and for each* $t > 0$, *there exists a sequence* A_n *of elements of* $\mathscr{F}$ *satisfying* $P([U_n \in A] \triangle A_n) \to 0$ *as* $n \to \infty$ (*where* $\triangle$ *denotes symmetric difference*), *and a finite constant* d *which may depend on any of* A, $\{A_n\}$, *or* t, *such that the following two conditions hold*;

$$\sum_{i=1}^{k_n(t)} |E_{i-1} X_{ni} I(A_n \cap [|X_{ni}| \leqq \eta_n])| \to_p 0 \tag{1.1}$$

where $\eta_n \leqq \infty$ *is some positive sequence of constants bounded away from* 0 *with*

$$\eta_n = O[P(\sup_{i \leqq k_n(t)} |X_{ni}| > d)]^{-\frac{1}{2}} .$$

(For the purpose of this definition and the proof of the theorem $\infty \cdot 0 = 0$ and so η_n may be $= \infty$ if $P\{\sup_i |X_{ni}| > d\} = 0$.)

There exists a sequence $t_n \to t$ *for which*

$$\sum_{i=1}^{k_n(t_n)} X_{ni}^2 \to_p t . \tag{1.2}$$

Then (W_n, U_n) *converges to* (W, U) *as random elements of the product metric space* $D[0, \infty) \times M$.

REMARK. It should be noted that $k_n(t)$ need not be a stopping time as was assumed in [3], that $\sup_{i \leqq k_n(t)} |X_{ni}| \to_p 0$ as a consequence of (1.2) and so η_n is permitted to diverge to ∞ at some maximal rate, and that in the presence of (1.2), (1.1) follows from the condition

$$\sum_{i=1}^{k_n(t)} |E_{i-1} X_{ni} I(A_n)| \to_p 0$$

together with either

$$\text{the sequence} \quad \sum_{i=1}^{k_n(t)} E_{i-1} X_{ni}^2 \quad \text{is tight,} \tag{1.1'}$$

or

$$E(\sup_{i \leqq k_n(t)} X_{ni}^2) \quad \text{is a bounded sequence,} \tag{1.1''}$$

These implications, with the fact that (1.1″) implies (1.1′) are contained in the proof of the theorem.

If we put all $U_n = U = 1$, an invariance principle results which generalizes Theorem (3.2) of [3].

Contained in the proof of Theorem 1 is the verification of the following central limit theorem:

THEOREM 2. *Assume k_n is a sequence of positive integer-valued random variables such that there exists a finite positive d satisfying*

(a) $\sum_{i=1}^{k_n} |E_{i-1}X_{ni}I(|X_{ni}| \leqq \eta_n)| \to_p 0$ *where η_n is some positive sequence of constants bounded away from* 0 *with*

$$\eta_n = O(P[\sup_{i \leqq k_n} |X_{ni}| > d]^{-\frac{1}{2}}) .$$

(b) $\sup_{i \leqq k_n} |X_{ni}| \to_p 0$.

(c) $\sum_{i=1}^{k_n} X_{ni}^2 \to_p \sigma^2 \geqq 0$.

Then $\sum_{i=1}^{k_n} X_{ni}$ converges in distribution to a normal $(0, \sigma^2)$ *variable.*

In the following corollary, the symbol $\mathscr{F}_{n,k_n(t)}$ will have two different interpretations depending on whether or not $k_n(t)$ is a stopping time for each t. If it is, $\mathscr{F}_{n,k_n(s)}$ is the sigma-algebra of all sets A such that $A \cap [k_n(s) = j] \in \mathscr{F}_{nj}$ for each j. If $k_n(s)$ is not a stopping time, $[k_n(s) = j]$ is replaced by the set $[\sum_{i=1}^{j} X_{ni}^2 > s]$ in the above definition.

COROLLARY 1. *Assume for each $t > 0$, there exists a sequence of random variables $\hat{U}_n(t)$ measurable with respect to $\mathscr{F}_{n,k_n(t)}$ such that $\hat{U}_n(t) - U_n \to_p 0$. Assume also that* 1.2 *holds and*

$$\sum_{i=1}^{k_n(t)} |E_{i-1}X_{ni}I(|X_{ni}| \leqq \eta_n)| \to_p 0 \tag{1.3}$$

where $\eta_n = O(P(\sup_{i \leqq k_n(t)} |X_{ni}| > d)^{-\frac{1}{2}})$.

Then (W_n, U_n) converge weakly to (W, U) as random elements of the product metric space.

2. Proofs. We begin with three elementary lemmas: Lemma 3 is a trivial extension of Theorem 3.2 of [3] along the lines mentioned in that paper. As observed, we may assume $k_n(t)$ is a stopping time and replace P by P_n with no major changes in the proof.

LEMMA 1. *Let $F_n(t)$ $n = 1, 2, \cdots$ and $F(t)$ be random elements of $D[0, \infty)$ such that $F_n(t)$ is almost surely nondecreasing in t, $F(t)$ is* a.s. *continuous and for each $t \geqq 0$, there exists $t_n \to t$ such that*

$$F_n(t_n) \to_p F(t) .$$

Then $\sup_t |F_n(t) - F(t)| \to_p 0$ *where the supremum is taken over any compact subset of* $[0, \infty)$.

PROOF. For simplicity we assume the compact subset is [0, 1]. For arbitrary

$\varepsilon > 0$, choose $\{t_{ni}; i = 0, 1, 2, \cdots, k\}$ where k is an integer strictly greater than $1/\varepsilon$ such that $t_{ni} \to i\varepsilon$ for each $i \leqq k$, and $F_n(t_{ni}) \to_p F(i\varepsilon)$ for $i \leqq k$ as $n \to \infty$. Then for n sufficiently large,

$$\sup_t |F_n(t) - F(t)| \leqq \sup_i |F_n(t_{n,i+1}) - F_n(t_{n,i})| + \sup_i |F_n(t_{n,i}) - F(t_{n,i})| + \sup_i |F(t_{n,i+1}) - F(t_{n,i})| .$$

The right-hand side converges in probability to $2 \sup_i |F(t_{n,i+1}) - F(t_{n,i})|$ which can be made arbitrarily small in probability by choosing a sufficiently small ε due to the uniform continuity of $F(t)$.

LEMMA 2. *If $E(\sup_i X_{ni}^2) \leqq K < \infty$ and $\sum_i X_{ni}^2$ is tight in the real line, then $\sum_i E(X_{ni}^2 \mid \mathscr{F}_{n,i-1})$ is also tight in the line.*

PROOF. For arbitrary $\varepsilon > 0$, choose c such that

$$P[\textstyle\sum_i X_{ni}^2 > c] \leqq \varepsilon .$$

Then define

$$Y_{ni} = X_{ni}^2 I(\textstyle\sum_{j=1}^{i-1} X_{nj}^2 \leqq c) \qquad \text{for each} \quad n, i .$$

Then

$$\begin{aligned} P\left[\textstyle\sum_i E_{i-1} X_{ni}^2 > \frac{c + K}{\varepsilon}\right] &\leqq P\left[\textstyle\sum_i E_{i-1} Y_{ni} > \frac{c+K}{\varepsilon}\right] + P[\textstyle\sum E_{i-1} Y_{ni} \neq \sum E_{i-1} X_{ni}^2] \\ &\leqq \frac{\varepsilon E(\sum_i Y_{ni})}{c + K} + \varepsilon \\ &\leqq 2\varepsilon . \end{aligned}$$

LEMMA 3. *Consider X_{ni}, $i = 1, 2, \cdots, k_n(t)$ defined on probability space $(\Omega_n, \mathscr{F}_n, P_n)$ and adapted to the increasing (in i) sequence of sigma-algebras $\mathscr{F}_{ni} \subset \mathscr{F}_n$ for each n, i. Suppose $E(X_{n,i} \mid \mathscr{F}_{n,i-1}) = 0$ a.s. and assume $k_n(t)$ is a stopping time for each t (i.e., satisfies 2.2). Moreover, assume for each $t > 0$, $\varepsilon > 0$, the following two conditions hold:*

(a) $\limsup_{n\to\infty} \int \sup_i X_{ni}^2 I(i \leqq k_n(t))\, dP_n = 0$.

(b) $P_n\{|\sum_{i=1}^{k_n(t)} X_{ni}^2 - t| > \varepsilon\} \to 0$.

Then $W_n(t) \to_{\mathscr{D}} W(t)$ on $D[0, \infty)$.

PROOF OF THEOREM 1. It follows from Stone (1963) that we need only show weak convergence for the functions restricted to some arbitrary compact interval contained in the nonnegative reals. We may assume the pertinent subset to be $[0, 1]$, since a simple change of scale will prove the more general result. Moreover, if we permit the probability measure to change with n (denoted $P_n : P_n(\cdot) = P(\cdot \mid A_n)$) we may assume without loss of generality that each $A_n = \Omega$ in condition (1.1). Indeed, as long as $P(A_n)$ is bounded away from 0, (1.1) and (1.2) continue to hold with these substitutions.

Observe first by Lemma 1 that

(2.1) $$\sup_{1\leq k_n(1)} X_{n,i}^2 \leqq \sup_{t\leq 1} |\textstyle\sum_{i=1}^{k_n(t)} X_{n,i}^2 - t| \to_p 0 .$$

We assume initially that $k_n(t)$ is a stopping time; viz.,

(2.2) $$[k_n(t) = j] \in \mathscr{F}_{nj} \qquad \text{for each} \quad t, n, \quad \text{and} \quad j .$$

Now choose a positive sequence $\bar{\eta}_n \to \infty$ so that $\bar{\eta}_n = o(P[\sup_{i\leq k_n(1)} |X_{ni}| > d])^{-\frac{1}{2}}$ and observe that

$$E \sup_i X_{ni}^2 I(|X_{ni}| \leqq \eta_n) \leqq d^2 + \eta_n{}^2 P_n[\sup_{i\leq k_n(1)} |X_{ni}| > d] ,$$

which sequence is bounded in n. It now follows from the conditional form of Jensen's inequality, Lemma 2, and (1.2) that

$$\textstyle\sum_{i=1}^{k_n(1)} |E_{i-1} X_{ni} I(\bar{\eta}_n < |X_{ni}| \leqq \eta_n)| \leqq \bar{\eta}_n{}^{-1} \sum_{i=1}^{k_n(1)} |E_{i-1} X_{ni}^2 I(|X_{ni}| \leqq \eta_n)| \to_p 0 .$$

Therefore, (1.1) continues to hold when η_n is replaced by $\bar{\eta}_n$. Therefore, if we define $\bar{X}_{ni} = X_{ni} I(|X_{ni}| \leqq \bar{\eta}_n)$, $\bar{X}_{ni}$ is an equivalent array satisfying

(2.3) $$\textstyle\sum_{i=1}^{k_n(1)} |E_{i-1}\bar{X}_{ni}| \to_p 0$$

(2.4) $$P(\bar{X}_{ni} = X_{ni},\ i = 1, 2, \cdots, k_n(1)) \to 1$$

(2.5) $$\textstyle\sum_{i=1}^{k_n(t)} \bar{X}_{ni}^2 \to_p t , \qquad \text{and}$$

$\sup_{i\leq k_n(1)} \bar{X}_{ni}^2 \leqq d^2 + \bar{\eta}_n{}^2 I(\sup_i |X_{ni}| > d)$, where the second term in the majorant converges in expectation to 0, implying that $\sup_i \bar{X}_{ni}^2$ is a uniformly integrable sequence and $E \sup_{i\leq k_n(1)} \bar{X}_{ni}^2 \to 0$. The weak convergence of $\sum_{i=1}^{k_n(t)} (\bar{X}_{ni} - E_{i-1}\bar{X}_{ni})$ to $W(t)$ (under the measures P_n) now follows by Lemma 3. The convergence of $W_n(t)$ follows from (2.3) and (2.4).

We now remove the condition (2.2) that $k_n(t)$ be a stopping time. Define $\bar{k}_n(t) = \inf\{j : \sum_{i=1}^{j} X_{ni}^2 > t\}$ where the infimum is finite and $\bar{k}_n(t)$ well defined with probability 1. Now for any $t, \delta > 0$,

$$P_n(\textstyle\sum_{i=1}^{\bar{k}_n(t)} X_{ni}^2 > \sum_{i=1}^{k_n(t+\delta)} X_{ni}^2) \leqq P_n(\sum_{i=1}^{k_n(t+\delta)} X_{ni}^2 < t) \to 0 .$$

It follows from this and

$$\textstyle\sum_{i=1}^{\bar{k}_n(t)} X_{ni}^2 > t , \qquad \text{that} \quad \sum_{i=1}^{\bar{k}_n(t)} X_{ni}^2 \to_p t ,$$

and the convergence is uniform in $t \in [0, 1]$ by Lemma 1.

Now by the argument above and the uniform convergence, we have, for each positive δ, $P_n[\bar{k}_n(t-\delta) \leqq k_n(t) \leqq \bar{k}_n(t+\delta)$ for all $t \leqq 1] \to 1$. It therefore follows that (1.1) also holds with $k_n(t)$ replaced by $\bar{k}_n(t)$ and by the first part of the proof, that

$$\bar{W}_n(t) = \textstyle\sum_{i=1}^{\bar{k}_n(t)} X_{ni} \to_{\mathscr{D}} W(t) \qquad \text{on the space } D[0, 1] .$$

But on the set $[\bar{k}_n(t-\delta) \leqq k_n(t) \leqq \bar{k}_n(t+\delta)$ for all $t \leqq 1]$, we have

$$\sup_{t\leq 1} |W_n(t) - \bar{W}_n(t)| \leqq \sup_{0\leq s<t\leq s+\delta\leq 1} |\bar{W}_n(t) - \bar{W}_n(s)|$$

and the tightness of $\bar{W}_n$ implies that this can be made arbitrarily small in probability by choosing δ sufficiently small. It follows that $W_n(t) \to_{\mathscr{D}} W(t)$ on $D[0, 1]$, again with respect to the measures P_n.

The proof of the theorem is now completed if we observe that for each U-continuity set A with $P(U \in A) > 0$, and for any W-continuity set B, we have shown $P(W_n \in B \mid A_n) \to P(W \in B)$. This and the fact that $P[(U_n \in A) \triangle A_n] \to 0$ imply $P(W_n \in B \text{ and } U_n \in A) \to P(W \in B)P(U \in A)$, which is the desired conclusion.

PROOF OF COROLLARY. We begin by assuming $k_n(t)$ is a stopping time. It is easy to see by a simple change of time scale (and taking limits as $\sigma^2 \to 0$) that if (1.1) holds with A_n replaced by Ω, and

$$(1.2') \qquad \sum_{i=1}^{k_n(t_n)} X_{ni}^2 \to_p \sigma^2 t \qquad \text{for} \quad \sigma^2 \geqq 0 \,,$$

W_n converges weakly to $W(\sigma^2 t)$.

Let $\delta(\cdot, \cdot)$ be a metric defining convergence in probability, e.g., $\delta(X, Y) = E(|X - Y|/1 + |X - Y|)$. Then by assumption, for each $t > 0$, there exists $\mathscr{F}_{n,k_n(t)}$ measurable random variables $\hat{U}_n(t)$ with $f(t) = \delta(\hat{U}_n(t), U_n) \to 0$.

Therefore by Lemma 1, page 188 of Chung (1968), there exists a sequence $s_n \downarrow 0$ with

$$f(s_n) = \delta(\hat{U}_n(s_n), U_n) \to 0 \,.$$

Now by the above extension of Theorem 1 applied to the random variables $X_{ni} I(i \leqq k_n(s_n))$ with $\sigma^2 = 0$, we conclude that

$$\sup_{t \leq 1} \left|\sum_{i=1}^{k_n(t)} X_{ni} I(i \leqq k_n(s_n))\right| \to_p 0 \,.$$

Consider the random variables $X_{ni} I(k_n(s_n) < i)$. Since $\hat{U}_n(s_n)$ is measurable with respect to $\mathscr{F}_{n,k_n(s_n)}$ we have, with $A_n = [\hat{U}_n(s_n) \in A]$; that $A_n \cap [k_n(s_n) = j] \in \mathscr{F}_{nj}$ and $A_n \cap [k_n(s_n) < i] \in \mathscr{F}_{n,i-1}$. Therefore, by assumption,

$$\begin{aligned} \sum_{i=1}^{k_n(t)} |E_{i-1} X_{ni} I(A_n \text{ and } k_n(s_n) < i \text{ and } |X_{ni}| < \eta_n)| \\ = I(A_n) \sum_{i=k_n(s_n)+1}^{k_n(t)} |E_{i-1} X_{ni} I(|X_{ni}| < \eta_n)| \to_p 0 \,. \end{aligned}$$

Moreover (1.2) continues to hold with X_{ni} replaced by $X_{ni} I(k_n(s_n) < i)$, and so putting

$$\bar{W}_n(t) = \sum_{k_n(s_n)+1}^{k_n(t)} X_{ni} \,,$$

we conclude that $(\bar{W}_n, U_n)$ converges jointly to (W, U). This and the fact that $\sup_t |\bar{W}_n(t) - W_n(t)| \to_p 0$ shows (W_n, U_n) converges weakly to the same limit.

The extension of this corollary to include the case when $k_n(t)$ is not a stopping time follows the same line as the last part of the proof of Theorem 1.

REFERENCES

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

[2] CHUNG, K. L. (1968). *A Course in Probability Theory.* Harcourt, Brace and World, New York.

[3] MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probability* **2** 620–628.

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, CANADA T6G 2G1