The Annals of Probability
1978, Vol. 6, No. 5, 829-846

FUNCTIONAL LIMIT THEOREMS FOR DEPENDENT VARIABLES

BY RICHARD DURRETT[1] AND SIDNEY I. RESNICK[2]

University of California, Los Angeles and Stanford University

Conditions are given for a sequence of stochastic processes derived from row sums of an array of dependent random variables to converge to a process with stationary, independent increments or to a process with continuous paths. We also discuss when row maxima converge to an extremal process.

The first result is a generalization of the classical results for independent random variables. The second result gives general conditions for convergence to processes which can be obtained from Brownian motion by a random change of time. This result is used to give a unified development of most of the martingale central limit theorems in the literature. An important aspect of our methods is that after the initial result is shown, we can avoid any further consideration of tightness.

1. Introduction. Let $\{X_{n,i}, n \geqq 1, i \geqq 1\}$ be an array of random variables defined on a probability space $(\Omega, \mathscr{F}, P)$ and let $\{\mathscr{F}_{n,i}, n \geqq 1, i \geqq 0\}$ be an array of sub σ-fields of $\mathscr{F}$ such that for each n and $i \geqq 1$, $X_{n,i}$ is $\mathscr{F}_{n,i}$ measurable and $\mathscr{F}_{n,i-1} \subset \mathscr{F}_{n,i}$. Suppose $k_n(t)$ is a nondecreasing right continuous function with range $\{0, 1, 2, \cdots\}$. The functions $k_n(t)$ are given time scales. Set

$$S_{n,0} = 0\,, \qquad S_{n,k} = \sum_{i=1}^{k} X_{n,i}\,, \quad k \geqq 1$$
$$Y_n(t) = S_{n,k_n(t)}\,,$$
$$Z_n(t) = S_{n,k_n(t)} - \sum_{i=1}^{k_n(t)} E(X_{n,i} 1_{\{|X_{n,i}|<\gamma\}} \mid \mathscr{F}_{n,i-1})$$
$$M_{n,0} = -\infty\,, \qquad M_{n,k} = \bigvee_{i=1}^{k} X_{n,i}\,, \quad k \geqq 1$$
$$M_n(t) = M_{n,k_n(t)}$$

where $\gamma > 0$. In this paper we give conditions for $\{Y_n(t), 0 \leqq t \leqq 1\}$ and $\{Z_n(t), 0 \leqq t \leqq 1\}$ to converge weakly (written $\Rightarrow$) as a sequence of random elements of $D[0, 1]$ and for $\{M_n(t), t > 0\}$ to converge weakly in $D(0, \infty)$. (For weak convergence terminology and notation see Billingsley (1968). For information about $D(0, \infty)$ see, for example, Lindvall (1973)).

For sums of random variables, investigations of this type have received considerable attention since the time of Lévy. Many authors (see [1]—[3], [5]—[9], [11]—[13], [15], [25], [33]—[37]) have given results for a variety of time scales

Received November 24, 1976; revised October 5, 1977.

[1] Supported by an NSF graduate fellowship at Stanford University. A revised version of this report was prepared and distributed using funds from NSF Grant No. MCS 72-04591.

[2] Partially supported by NSF Grant No. MPS 74-21416. This author is now at Colorado State.

AMS 1970 *subject classifications.* Primary 60F05; Secondary 60G45, 60J30.

Key words and phrases. Weak convergence, invariance principle, Poisson process, Brownian motion, process with stationary, independent increments, extremal process, martingales.

and under a bewildering assortment of conditions. In this paper we develop a framework which allows us to consolidate and in many cases extend these results. Our approach will be to give general conditions for convergence on an arbitrary given time scale $k_n(t)$ and then obtain the results for specific time scales as special cases. Concerning the time scales $k_n(t)$, if these are random then in Section 2 we need only the restriction that $k_n(1)$ is a stopping time for $\{\mathscr{F}_{n,k}, k \geqq 1\}$. In Sections 3 and 4 it is convenient to suppose that $k_n(t)$ is a stopping time for each $t > 0$.

In Section 2 we treat the problem of convergence to a process with continuous paths. To do this we will start with a result of Freedman (1971, 1975) about the convergence of arrays with uniformly bounded variables and then extend this using truncation and the idea of a random change of time to obtain more general results. In the course of doing this we will obtain most of the results in the literature as special cases. We should point out that the results given here do not exhaust the possibilities. We have concentrated on convergence to Brownian motion or mixtures of Brownian motions. The methods we have used can be extended to give conditions for convergence to diffusions. To illustrate what kind of results are possible we have given two results of this type in Section 2. A systematic development of weak convergence to diffusions using the idea of a random change of time will be given in Helland [17]. For another martingale approach to these results see Stroock and Varadhan [38] (especially Theorem 10.3).

Sections 3 and 4 study convergence to processes which have jumps. The first step is taken in Section 3 where conditions are given for the convergence of a sequence of point processes associated with the array to a limit two dimensional Poisson process. Once this development is completed, it is relatively easy to "sum up the points" to show Z_n converges to a limiting Lévy process or to apply the continuous mapping theorem to get convergence of M_n to a limit extremal process.

2. Convergence to processes with continuous paths. In what follows we derive conditions for Y_n and Z_n (defined in the introduction) to converge weakly (as a sequence of random elements of $D[0, 1]$) to processes with continuous paths. In this section all the processes we define have their paths in $D[0, 1]$ and all the weak convergence results involve this space unless the contrary is indicated.

Our starting point is Theorem 2.1 which is due to Freedman [11], pages 89–93; see also [13]. To introduce this result we need the following:

DEFINITION. A collection of random variables $X_{n,i}$, $n \geqq 1$, $i \geqq 1$ and σ-fields $\mathscr{F}_{n,i}$, $n \geqq 1$, $i \geqq 0$ is said to be a martingale difference array if

(i) for all $n \geqq 1$, $\mathscr{F}_{n,i}$, $i \geqq 0$ is an increasing sequence of σ-fields;
(ii) for all $n, i \geqq 1$, $X_{n,i}$ is $\mathscr{F}_{n,i}$ measurable; and
(iii) for all $n, i \geqq 1$, $E(X_{n,i} \mid \mathscr{F}_{n,i-1}) = 0$.

THEOREM 2.1. (Freedman). *Let $\{X_{n,i}, \mathscr{F}_{n,i}\}$ be a martingale difference array and suppose there are numbers $\varepsilon_n \downarrow 0$ so that $|X_{n,i}| \leqq \varepsilon_n$ for all n and i. Let $V_{n,0} = 0$, $V_{n,j} = \sum_{i=1}^{j} E(X_{n,i}^2 \mid \mathscr{F}_{n,i-1})$ and $j_n(t) = \sup\{j \mid V_{n,j} \leqq t\}$. If $P\{\lim_{j\to\infty} V_{n,j} = \infty\} = 1$ then as $n \to \infty$, $S_{n,j_n(\cdot)}$ converges weakly as a sequence of random elements of $D[0, \infty)$ to a Brownian motion W.*

With this result in mind the next step is to write Y_n and Z_n as a sum with three terms (a) the sum of variables in an array which satisfies the hypotheses of Theorem 2.1, (b) the sum of the "large" $X_{n,i}$; and (c) a centering term. To describe these decompositions we have to introduce some notation. Let ε_n be a sequence of positive numbers which decrease to zero and define

$$\hat{X}_{n,i} = X_{n,i} 1_{\{|X_{k,i}| > \varepsilon_n\}}$$
$$\bar{X}_{n,i} = X_{n,i} 1_{\{|X_{n,i}| \leqq \varepsilon_n\}}$$
$$\bar{\bar{X}}_{n,i} = \bar{X}_{n,i} - E(\bar{X}_{n,i} \mid \mathscr{F}_{n,i-1}) .$$

Then $|\bar{\bar{X}}_{n,i}| \leqq 2\varepsilon_n$ so if we let $\bar{\bar{S}}_{n,k} = \sum_{i=1}^{k} \bar{\bar{X}}_{n,i}$ and

$$j_n(t) = \sup\{j : \bar{\bar{V}}_{n,j} = \sum_{i=1}^{j} E(\bar{\bar{X}}_{n,i}^2 \mid \mathscr{F}_{n,i-1}) \leqq t\} ,$$

we have from Theorem 2.1 as $n \to \infty$, $W_n := \bar{\bar{S}}_{n,j_n(\cdot)} \Rightarrow W$.

Next set

$$\hat{Y}_n(t) = \sum_{j=1}^{k_n(t)} \hat{X}_{n,j} ,$$
$$\bar{Y}_n(t) = \sum_{j=1}^{k_n(t)} \bar{X}_{n,j} ,$$
$$A_n(t) = \sum_{j=1}^{k_n(t)} E(\bar{X}_{n,j} \mid \mathscr{F}_{n,j-1})$$

and let $\varphi_n(t)$ be any strictly increasing continuous function which satisfies $\bar{\bar{S}}_{n,k_n(t)} = \bar{\bar{S}}_{n,j_n(\varphi_n(t))}$. From the last five definitions it is immediate that

$$Y_n(t) = W_n(\varphi_n(t)) + \hat{Y}_n(t) + A_n(t) \tag{2.1}$$

and if we let $B_n^{\gamma}(t) = \sum_{j=1}^{k_n(t)} E(X_{n,j} 1_{\{\varepsilon_n < |X_{n,j}| < \gamma\}} \mid \mathscr{F}_{n,j-1})$, then

$$Z_n(t) = W_n(\varphi_n(t)) + \hat{Y}_n(t) - B_n^{\gamma}(t) . \tag{2.2}$$

THEOREM 2.2. *If for some $\varepsilon_n \downarrow 0$*

(i) *$P\{\max_{0\leqq s\leqq 1} |\hat{Y}_n(s)| > 0\} \to 0$ and*
(ii) *$(W_n, \varphi_n) \Rightarrow (W, \varphi)$ with $P\{\varphi$ is continuous$\} = 1$*

then $Z_n \Rightarrow W \circ \varphi$. If in addition

(iii) *$A_n \to_P 0$ or what is equivalent for some $\lambda > 0$ we have*

$$C_n^{\lambda}(t) := \sum_{i=1}^{k_n(t)} E(X_{n,i} 1_{\{|X_{n,i}| < \lambda\}} \mid \mathscr{F}_{n,i-1}) \to_P 0 \qquad \textit{for all} \quad t > 0 .$$

Then $Y_n \Rightarrow W \circ \varphi$.

PROOF. From (ii) and formulas (17.7)—(17.9) in Billingsley (1968), it follows that $W_n \circ \varphi_n \Rightarrow W \circ \varphi$. From (i). $\hat{Y}_n \to_P 0$ so to prove the first result it remains to show that $B_n^{\gamma} \to_P 0$.

To do this it suffices to show that

$$\gamma \sum_{i=1}^{k_n(1)} P(|X_{n,i}| > \varepsilon_n \mid \mathscr{F}_{n,i-1}) \to_P 0 .$$

Let τ_n be the time of the first jump of size $> \varepsilon_n$, that is

$$\tau_n = \sup \{t \leqq 1 : \sum_{j=1}^{k_n(t)} 1_{\{|X_{n,j}| > \varepsilon_n\}} = 0\} .$$

Then

$$E(\sum_{j=1}^{k_n(\tau_n)} P(|X_{n,j}| > \varepsilon_n \mid \mathscr{F}_{n,j-1})) \\ = E(\sum_{j=1}^{\infty} 1_{\{k_n(1) \geqq j, |X_{n,i}| \leqq \varepsilon_n \text{ for } 1 \leqq i < j\}} P(|X_{n,j}| > \varepsilon_n \mid \mathscr{F}_{n,j-1})) .$$

Since $\{k_n(1) \geqq j, |X_{n,i}| \leqq \varepsilon_n \text{ for } 1 \leqq i < j\} \in \mathscr{F}_{n,j-1}$ the last expression equals

$$P\{\max_{1 \leqq j \leqq k_n(1)} |X_{n,j}| > \varepsilon_n\} = P\{\max_{0 \leqq s \leqq 1} |\hat{Y}_n(s)| > 0\} \to 0$$

by (i). So if $\delta > 0$

$$P\{\sum_{j=1}^{k_n(1)} P(|X_{n,j}| > \varepsilon_n \mid \mathscr{F}_{n,j-1}) > \delta\} \\ \leqq P\{\tau_n < 1\} + \delta^{-1} P\{\max_{1 \leqq j \leqq k_n(1)} |X_{n,j}| > \varepsilon_n\} .$$

Since the right-hand side of the last inequality converges to 0 as $n \to \infty$ this proves the first result.

To prove the second result observe that $Y_n = Z_n + B_n^\gamma + A_n$ and $A_n = C_n^\lambda - B_n^\lambda$. From the first part of the proof $B_n^\gamma \to_P 0$. Therefore the two assumptions in (iii) are equivalent and under either one $Y_n - Z_n \to_P 0$ so that $Y_n \Rightarrow W \circ \varphi$.

From Theorem 2.2 it is easy to obtain the following result which is the martingale analogue of the Lindeberg-Feller theorem.

THEOREM 2.3. *Suppose $\{X_{n,i}, \mathscr{F}_{n,i}\}$ is a martingale difference array. If*

(a) *for all $t \in [0, 1]$, $\sum_{i=1}^{k_n(t)} E(X_{n,i}^2 \mid \mathscr{F}_{n,i-1}) \to_P ct$ and*
(b) *for all $\varepsilon > 0$*

$$\sum_{i=1}^{k_n(t)} E(X_{n,i}^2 1_{\{|X_{n,i}| > \varepsilon\}} \mid \mathscr{F}_{n,i-1}) \to_P 0$$

then $Y_n = S_{n,k_n(\cdot)} \Rightarrow W(c\cdot)$.

REMARK. If we suppose $X_{n,i} = c_n^{-1}(S_i - S_{i-1})$ where c_n is a constant and S_i is a martingale with $s_i^2 = ES_i^2 < \infty$ for all i and let $k_n(t) = k$ when $s_k^2 \leqq ts_n^2 < s_{k+1}^2$, Theorem 2.3 gives one of the results of Scott ([37], page 120). If we let $k_n(t)$ be the time scale $j_n(t)$ defined in Theorem 2.2 we get Theorem 5 of Rootzén ([35], page 206, see his Remark 6). It is trivial to generalize our result to obtain Theorem 3.8 of McLeish ([25], page 626).

PROOF. From (b) it follows that if we let ε_n decrease to zero slowly enough then

$$\varepsilon_n^{-2} \sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1_{\{|X_{n,i}| > \varepsilon_n\}} \mid \mathscr{F}_{n,i-1}) \to_P 0 . \tag{2.3}$$

To prove that $S_{n,k_n(\cdot)} \Rightarrow W(c\cdot)$ we will show that if $\varepsilon_n \downarrow 0$ are chosen so that (2.3) holds then the hypotheses of Theorem 2.2 are satisfied.

To check that (i) holds we observe that from Lemma 3.5 of Dvoretzky (1972) if $\delta > 0$ we have for every $N \geqq 1$

$$P\{\max_{1\leqq i\leqq N} |X_{n,i}| > \varepsilon_n\} \leqq \delta + P\{\sum_{i=1}^{N} P(|X_{n,i}| > \varepsilon_n \,|\, \mathscr{F}_{n,i-1}) > \delta\} .$$

Applying this formula to the martingale difference array $X'_{n,i} = X_{n,i} 1_{\{k_n(1)\geqq i\}}$ and letting $N \to \infty$ gives

$$(2.4)\qquad P\{\max_{1\leqq i\leqq k_n(1)} |X_{n,i}| > \varepsilon_n\} \leqq \delta + P\{\sum_{i=1}^{k_n(1)} P(|X_{n,i}| > \varepsilon_n \,|\, \mathscr{F}_{n,i-1}) > \delta\} .$$

Since

$$\sum_{i=1}^{k_n(1)} \varepsilon_n{}^2 P(|X_{n,i}| > \varepsilon_n \,|\, \mathscr{F}_{n,i-1}) \leqq \sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1_{\{|X_{n,i}>\varepsilon_n\}} \,|\, \mathscr{F}_{n,i-1})$$

it follows from (2.3) that lim sup of the right-hand side of (2.4) is less than δ. Since δ is an arbitrary positive number, this shows that (i) of Theorem 2.2 holds.

To check (ii) we observe that it suffices to show $\varphi_n \to_P \varphi$ where $\varphi(t) = ct$ (Billingsley, 1968, Theorem 4.4). Now φ_n is a monotone function and φ is continuous so it suffices to show that $\varphi_n(t) \to_P ct$ for each $t > 0$. To do this we observe that $\bar{\bar{S}}_{n,k_n(t)} = \bar{\bar{S}}_{n,j_n(\varphi_n(t))}$ and from the definition of j_n

$$t - 4\varepsilon_n{}^2 \leqq \sum_{j=1}^{j_n(t)} E(\bar{\bar{X}}_{n,j}^2 \,|\, \mathscr{F}_{n,j-1}) \leqq t .$$

So to show that $\varphi_n(t) \to_P ct$ it suffices to show that

$$\sum_{j=1}^{k_n(t)} E(\bar{\bar{X}}_{n,j}^2 \,|\, \mathscr{F}_{n,j-1}) \to_P ct .$$

The left side of the above equals $\sum_{j=1}^{k_n(t)} E(\bar{X}_{n,j}^2 \,|\, \mathscr{F}_{n,j-1}) - \sum_{j=1}^{k_n(t)} (E(\bar{X}_{n,j} \,|\, \mathscr{F}_{n,j-1}))^2$. The first sum converges in probability to ct by (2.3) and (a). Because $\{X_{n,j}\}$ is a martingale difference array, the second sum equals

$$\sum_{j=1}^{k_n(t)} (E(X_{n,j} 1_{\{|X_{n,j}|>\varepsilon_n\}} \,|\, \mathscr{F}_{n,j-1}))^2 .$$

From Jensen's inequality, we have that this expression

$$\leqq \sum_{j=1}^{k_n(t)} E(X_{n,j}^2 1_{\{|X_{n,j}|>\varepsilon_n\}} \,|\, \mathscr{F}_{n,j-1}) \to_P 0$$

by (2.3).

To complete the proof it remains to verify that $A_n \to_P 0$. To do this we observe that

$$\begin{aligned}\sup_{0\leqq t\leqq 1} |A_n(t)| &\leqq \sum_{i=1}^{k_n(1)} |E(X_{n,i} 1_{\{|X_{n,i}|\leqq\varepsilon_n\}} \,|\, \mathscr{F}_{n,i-1})| \\ &= \sum_{i=1}^{k_n(1)} |E(X_{n,i} 1_{\{|X_{n,i}|>\varepsilon_n\}} \,|\, \mathscr{F}_{n,i-1})| \\ &\leqq \sum_{i=1}^{k_n(1)} |E(|X_{n,i}| 1_{\{|X_{n,i}|>\varepsilon_n\}} \,|\, \mathscr{F}_{n,i-1}) \\ &\leqq \varepsilon_n{}^{-1} \sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1_{\{|X_{n,i}|>\varepsilon_n\}} \,|\, \mathscr{F}_{n,i-1}) \to_p 0 .\end{aligned}$$

A consequence of Theorem 2.3 that we will need is the following.

COROLLARY 2.1. *If $\{X_{n,i}, \mathscr{F}_{n,i}\}$ is a martingale difference array and*

$$\sum_{i=1}^{k_n(1)} E(X_{n,i}^2 \,|\, \mathscr{F}_{n,i-1}) \to_P 0$$

then $Y_n \to_P 0$.

EXAMPLE 2.1. Chain dependent variables: let $\{J_n, n \geqq 0\}$ be an m state periodic, irreducible Markov chain with transition matrix $\{p_{i,j}, 1 \leqq i, j \leqq m\}$ and stationary distribution $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_m)$. The random variables $\{X_n, n \geqq 1\}$ are chain dependent if for $\mathscr{J}_n = \mathscr{B}(J_0, \cdots, J_n, X_1, \cdots, X_n)$ and $n \geqq 1$ we have

$$P\{J_n = j, X_n \leqq x \mid \mathscr{J}_{n-1}\} = P\{J_n = j, X_n \leqq x \mid J_{n-1}\} = p_{J_{n-1},j} H_{J_{n-1}}(x)$$

where $H_1, \cdots, H_m$ are given distribution functions. It then follows that

$$P\{X_n \leqq x \mid \mathscr{J}_{n-1}\} = P\{X_n \leqq x \mid J_{n-1}\} = H_{J_{n-1}}(x)$$

and

$$P(\bigcap_{i=1}^n [X_i \leqq x_i] \mid J_0, \cdots, J_{n-1}) = \prod_{i=1}^n H_{J_{i-1}}(x_i) .$$

Suppose that $\int xH_i(dx) = 0$, $\int x^2 H_i(dx) = \sigma_i^2 < \infty$, $i = 1, \cdots, m$ and set $X_{n,i} = X_i/n^{\frac{1}{2}}$. Then $\{X_{n,i}, i \geqq 1\}$ is a martingale difference array since

$$E\{X_{n,i} \mid \mathscr{J}_{i-1}\} = \int xH_{J_{i-1}}(dx) = 0 .$$

We will now show that the conditions of Theorem 2.3 hold if $k_n(t) = [nt]$. To do this we let $\pi_i(n) = \sum_{j=0}^{n-1} 1_{\{J_j = i\}}$ and observe that $\pi_i(n)/n$ converges to π_i almost surely so

$$n^{-1} \sum_{i=1}^{[nt]} E(X_i^2 \mid \mathscr{J}_{i-1}) = n^{-1} \sum_{j=1}^m \pi_j([nt])\sigma_j^2 \rightarrow \sum_{j=1}^m \pi_j \sigma_j^2 t \quad \text{a.s.}$$

and condition (a) of Theorem 2.3 is satisfied.

To check condition (b) observe that

$$\sum_{i=1}^n E(X_{n,i}^2 1_{\{|X_{n,i}| > \varepsilon\}} \mid \mathscr{J}_{i-1}) \leqq \sum_{j=1}^m n \int_{|x| > \varepsilon n^{\frac{1}{2}}} x^2 H_j(dx) \rightarrow 0 .$$

Since conditions (a) and (b) are satisfied we conclude

$$\sum_{i=1}^{[n\cdot]} \frac{X_i}{n^{\frac{1}{2}}} \Rightarrow W(\sum_{i=1}^m \pi_i \sigma_i^2) .$$

Asymptotic normality of sums of chain dependent variables has been considered by Keilson and Wishart (1964) and O'Brien (1974a).

The decomposition of $S_{n,k_n(\cdot)}$ given by (2.1) can also be used to compute convergence to limits other than Brownian motion. The most elementary situation occurs when (W_n, φ_n) converges to (W, φ) where W and φ are independent. This is, of course, automatic when φ_n converges to a constant. Our next lemma shows that W and φ are independent whenever φ_n converges in probability. To prove this we have to introduce a notion of mixing.

Suppose $\{V_n, n \geqq 0\}$ are random elements of a metric space S and defined on $(\Omega, \mathscr{F}, P)$. The sequence $\{V_n\}$ is mixing in the sense of Renyi (or briefly R-mixing) if there is a random element V such that for each $B \in \mathscr{F}$ with $P(B) > 0$ we have $(V_n \mid B) \Rightarrow V$.

The reason for our interest in this concept stems from the following well-known characterization (cf. Billingsley, 1968, Theorem 4.5): if $V_n \Rightarrow V$, then $\{V_n\}$ is R-mixing iff for any sequence of random elements U_n of a metric space S' such that $U_n \rightarrow_P U$ we have $(V_n, U_n) \Rightarrow (V, U)$ where V and U are independent.

A sufficient condition for W_n to be R-mixing is

THEOREM 2.4. *Suppose the array* $\{X_{n,i}, \mathscr{F}_{n,i}\}$ *has the property that for each* $i \geqq 0$, $\mathscr{F}_{n,i}$ *increases as n increases. For any* $\varepsilon_n \downarrow 0$ *the sequence* W_n *is R-mixing as a sequence of random elements of* $D[0, \infty)$.

REMARK. This result can be proved in the same way as Theorem 2.4 in Rootzén [33] and so the proof will be omitted. The reader should observe that the hypothesis is satisfied if, for example, $\{S_i, \mathscr{F}_i\}$ is a martingale $\mathscr{F}_{n,i} = \mathscr{F}_i$ and $X_{n,i} = (S_i - S_{i-1})/c_n$ where c_n is a sequence of positive constants. This result was first stated by McLeish ([25], page 628) under the assumption that $\mathscr{F}_{n,i}$ decreases as n increases. To see that his claim is false let W be a Brownian motion, let $X_{n,i} = W(i2^{-n}) - W((i-1)2^{-n})$, let $\mathscr{F}_{n,i}$ be the σ-field generated by $\{W(s), 0 \leqq s \leqq i2^{-n}\}$, and let $U_n = V_n = W(1)$.

An immediate consequence of the preceding result is the following result about convergence to mixtures of Brownian motion:

THEOREM 2.5. *Suppose* $\{X_{n,i}, \mathscr{F}_{n,i}\}$ *is a martingale difference array and the fields* $\mathscr{F}_{n,i}$ *increase as n increases. If*

(a) *for all* $t > 0$

$$\sum_{i=1}^{k_n(t)} E(X_{n,i}^2 \mid \mathscr{F}_{n,i-1}) \to_P \varphi(t) \qquad \text{with} \quad P\{\varphi \text{ is continuous}\} = 1$$

and

(b) *for all* $\varepsilon > 0$

$$\sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1_{\{|X_{n,i}| > \varepsilon\}} \mid \mathscr{F}_{n,i-1}) \to_P 0 .$$

Then $Y_n = S_{n,k_n(\cdot)} \Rightarrow W \circ \varphi$ *where* W *and* φ *are independent.*

PROOF. From the proof of Theorem 2.3 we have that $P\{\max_{0 \leqq s \leqq 1} |\hat{Y}_n(s)| > 0\} \to 0$, $A_n \to_P 0$, and $\varphi_n \to_P \varphi$. Now $W_n \Rightarrow W$ and from Theorem 2.4, W_n is R-mixing so from the characterization of R-mixing given before Theorem 2.4 it follows that $(W_n, \varphi_n) \Rightarrow (W, \varphi)$ where W and φ are independent.

Theorem 2.2 can also be used to prove results when W and φ are not independent. In this case it is usually more difficult to verify that $(W_n, \varphi_n) \Rightarrow (W, \varphi)$ but in one situation it is easy. If we can write $\varphi_n = H_n(W_n)$ and show that $H_n \to H$ locally uniformly (i.e., $H_n(f_n) \to H(f)$ whenever $f_n \to f$) then we can apply the continuous mapping theorem to conclude $(W_n, \varphi_n) \Rightarrow (W, H(W))$. The next result is an example of a theorem which can be obtained from Theorem 2.2 using this technique.

THEOREM 2.6. *Let* $\{X_{n,i}, \mathscr{F}_{n,i}\}$ *be a martingale difference array and let* $S_{n,k} = \sum_{i=1}^{k} X_{n,i}$. *Suppose that for all* $\varepsilon > 0$ *and* $n \geqq 1$ *there are functions* α_n *and* β_n^ε *so that*

$$E(X_{n,k}^2 \mid \mathscr{F}_{n,k-1}) = \alpha_n(S_{n,k-1})$$

and

$$E(X_{n,k}^2 1_{\{|X_{n,k}| > \varepsilon\}} \mid \mathscr{F}_{n,k-1}) \leqq \beta_n^\varepsilon(S_{n,k-1}) .$$

If there is a bounded continuous positive function α *so that* $n\alpha_n \to \alpha$ *and* $n\beta_n^\varepsilon \to 0$

uniformly on compact sets then

$$S_{n,[n\cdot]} \Rightarrow W \circ \varphi$$

where $\varphi^{-1}(s) = \int_0^s [\alpha(W(t))]^{-1}\,dt$.

The limit process is a diffusion with no drift and infinitesimal variance $\alpha(y)$.

PROOF. This result is a special case of Theorem 10.3 in [38] (for this version of the result see [39]) so we will only sketch the proof to show how our methods can be applied.

Let $M > 0$. Let $l_n = \min\{j \geq 0 : |S_{n,j}| \geq M\}$ and let $k_n(t) = [nt] \wedge l_n$. To prove the result it suffices to show that if $V = W \circ \varphi$ then (a) $S_{n,k_n(\cdot)} \Rightarrow V(\cdot \wedge T_M)$ where $T_M = \inf\{s \geq 0 : |V(s)| \geq M\}$ and (b) in the limit process V we have $\lim_{M\to\infty} T_M = \infty$ almost surely.

The second conclusion is a consequence of the assumption that α is bounded. To show that (a) holds we will use Theorem 2.2. To check that (i) is satisfied we observe that

$$\begin{aligned} E \sum_{i=1}^{k_n(1)} \hat{X}_{n,i}^2 &= E \sum_{i=1}^{n} E(\hat{X}_{n,i}^2 1_{\{k_n(1)\geq i\}} \mid \mathscr{F}_{n,i-1}) \\ &= E \sum_{i=1}^{n} 1_{\{k_n(1)\geq i\}} E(\hat{X}_{n,i}^2 \mid \mathscr{F}_{n,i-1}) \leq n \sup_{-M\leq x\leq M} \beta_n^{\varepsilon_n}(x) \end{aligned}$$

so

$$E \sum_{i=1}^{k_n(t)} |\hat{X}_{n,i}| \leq \varepsilon_n^{-1}(n \sup_{-M\leq x\leq M} \beta_n^{\varepsilon_n}(x)) .$$

Since we have assumed that for all $\delta > 0$ $n\beta_n^{\delta} \to 0$ uniformly on compact sets it follows from the last inequality that if we pick $\varepsilon_n \to 0$ slowly enough then the right-hand side of the last inequality approaches 0. This shows that (i) holds.

The next step is to show that condition (iii) is satisfied. To do this we let $\lambda > 0$ and observe that since $\{X_{n,i}, \mathscr{F}_{n,i}\}$ is a martingale difference array it follows from Jensen's inequality that

$$\begin{aligned} |C_n^{\lambda}(t)| &\leq \sum_{i=1}^{n} 1_{\{k_n(1)\geq i\}} E(|X_{n,i}| 1_{\{|X_{n,i}|>\lambda\}} \mid \mathscr{F}_{n,i}) \\ &\leq \lambda^{-1} \sum_{i=1}^{n} E(X_{n,i}^2 1_{\{|X_{n,i}|>\lambda\}} \mid \mathscr{F}_{n,i}) \leq \lambda^{-1} n \sup_{-M\leq x\leq M} \beta_n^{\lambda}(x) \to 0 . \end{aligned}$$

Finally we need to show that (ii) holds. Our first step is to observe that from (i), (iii), and the proof of Theorem 2.2 we have that

$$\sup_{0\leq t\leq 1} |S_{n,j_n(t)} - W_n(t)| \to 0.$$

From this it follows that if we let

$$\sigma_n(s) = \int_0^s (n\alpha_n(S_{n,j_n(t)}))^{-1}\,dt$$

and

$$\tau_n(s) = \int_0^s (n\alpha_n(W_n(t)))^{-1}\,dt$$

then $\sup_{0\leq s\leq 1} |\sigma_n(s) - \tau_n(s)| \to 0$. Since α is continuous and $n\alpha_n \to \alpha$ uniformly it follows from the definition of τ_n that

$$(W_n, \tau_n) \Rightarrow (W, \tau) \qquad \text{where} \quad \tau(s) = \int_0^s \alpha(W(t))^{-1}\,dt$$

and hence that $(W_n, \sigma_n) \Rightarrow (W, \tau)$.

To complete the proof we have to show that $(W_n, \varphi_n) \Rightarrow (W, \varphi)$. To do this we start by observing that $[n\sigma_n(s)] \wedge l_n = j_n(s) \wedge l_n$ (to check this note that if $t = \alpha_n(S_{n,0})$ then $\sigma_n(t) = 1/n$ and then use induction). Now τ is strictly increasing so it follows from the last observation that if the φ_n are defined so that $[ns] = j_n(\varphi_n(s))$ then

$$(W_n, \varphi_n) \Rightarrow (W, \varphi) \qquad \text{where} \quad \varphi^{-1}(s) = \textstyle\int_0^s \alpha(W(t))^{-1}\,dt$$

and this completes the proof.

The hypotheses of Theorem 2.6 may be weakened to allow the function α to be unbounded or equal to 0. This is accomplished by considering a collection of processes $\{S_{n,k_n(t\wedge T_{m,n})}, n \geqq 1, m \geqq 1\}$ with $T_{m,n} = \inf\{s \geqq 0 : S_{n,k_n(s)} \notin [a_m, b_m]\}$ and letting $n \to \infty$ then $m \to \infty$. One of the theorems we can obtain in this way is the following result about critical branching processes which is due to Lamperti [21] and Lindvaal [23]. Since the proof we obtain in this way is harder than the original one, it is omitted.

THEOREM 2.7. *For each $n \geqq 1$ let $\{Z_{n,k}; k \geqq 0\}$ be a Galton-Watson process with offspring distribution $\{p_m, m \geqq 0\}$ and suppose that $\sum_{m=0}^{\infty} mp_m = 1$ and $\sum_{m=0}^{\infty} (m-1)^2 p_m = \sigma^2 \in (0, \infty)$. If $Z_{n,0}/n \to c > 0$ then $Z_{n,[n\cdot]}/n$ converges weakly to a diffusion on $[0, \infty)$ which has no drift and has infinitesimal variance $\alpha(y) = \sigma^2 y$.*

3. Weak convergence to a Poisson process. Suppose $\{X_{n,i}, \mathscr{F}_{n,i}\}$ is an array of variables satisfying the conditions given in the introduction and let the given time scales be $k_n(t)$ as before where we now suppose for each $t > 0$ that $k_n(t)$ is a stopping time. For each n, form the point process on $(R^2, \mathscr{B}(R^2))$ which has counting function

$$N_n((0, t] \times [a, b]) = \textstyle\sum_{i=1}^{k_n(t)} 1_{\{X_{n,i} \in [a,b] \cap (x_0, \infty)\}}$$

where x_0 is specified below and the right side is zero if $k_n(t) = 0$. In this section, we give conditions for N_n to converge weakly to a limit two dimensional Poisson process. At the end of this section, our result is applied to obtain weak convergence of maxima of dependent variables to limit extremal processes. In Section 4 the Poisson convergence result is applied to derive criteria for weak convergence of sums of dependent variables to Lévy processes. Necessary background on the weak convergence of point processes may be found in Jagers (1974).

THEOREM 3.1. *Let ν be a σ-finite measure on $(R, \mathscr{B}(R))$ with the property that if $x_0 = \inf\{x \mid \nu(x, \infty) < \infty\}$ then $-\infty \leqq x_0 < \infty$. Suppose for all $t > 0$ and for $x > x_0$ such that $\nu(\{x\}) = 0$ we have as $n \to \infty$*

$$\textstyle\sum_{j=1}^{k_n(t)} P\{X_{n,j} > x \mid \mathscr{F}_{n,j-1}\} \to_P t\nu(x, \infty) \tag{3.1}$$

and

$$\max_{j \leqq k_n(t)} P\{X_{n,j} > x \mid \mathscr{F}_{n,j-1}\} \to_P 0\,. \tag{3.2}$$

Then $N_n \Rightarrow N$ *where* N *is a Poisson process on* $[0, \infty) \times (x_0, \infty)$ *with mean measure* $dt \times d\nu$.

REMARK. We can replace (3.2) by the equivalent condition

$$\sum_{j=1}^{k_n(t)} (P[X_{n,j} > x \mid \mathscr{F}_{n,j-1}])^2 \to_P 0 . \tag{3.2'}$$

PROOF. Let $\mathscr{E}$ be the class of rectangles in R^2 of the form $(a, b] \times (c, d]$ or $(a, b] \times (c, \infty)$ where $0 \leqq a < b$, $\nu(\{c\}) = \nu(\{d\}) = 0$ and $x_0 < c < d$. Suppose A is a disjoint union of rectangles in $\mathscr{E}$ say, $A = \sum_1^m R_i$ where $R_i = (a_i, b_i] \times (c_i, d_i]$. We first show that

$$N_n(A) \Rightarrow N(A) \tag{3.3}$$

where $N(A)$ is Poisson distributed with mean $\sum_1^m (b_i - a_i)\nu(c_i, d_i]$.

If $\{(a_i, b_i], i \leqq m\}$ are not pairwise disjoint then A can be rewritten $A = \sum_i (\alpha_i, \beta_i] \times I_i$ where now the time intervals $(\alpha_i, \beta_i]$ are pairwise disjoint and I_i is a finite union of disjoint intervals. Then

$$N_n(A) = \sum_i \sum_{j=k_n(\alpha_i)+1}^{k_n(\beta_i)} 1_{\{X_{n,j} \in I_i\}} .$$

Note that by (3.1) and (3.2)

$$\sum_i \sum_{j=k_n(\alpha_i)+1}^{k_n(\beta_i)} P\{X_{n,j} \in I_i \mid \mathscr{F}_{n,j-1}\} \to_P \sum_i (\beta_i - \alpha_i)\nu(I_i) ,$$

and

$$\sum_i \sum_{j=k_n(\alpha_i)+1}^{k_n(\beta_i)} P^2\{X_{n,j} \in I_i \mid \mathscr{F}_{n,j-1}\} \to_P 0 ,$$

so (3.3) now follows from Freedman (1974), Theorem 5.

Based on (3.3), the random measures $\{N_n\}$ are tight (cf. Jagers (1974), page 209). Let $N_{n'}$ be a weakly convergent subsequence and suppose for some random measure $\tilde{N}$ that $N_{n'} \Rightarrow \tilde{N}$. From (3.3) we must have $\tilde{N}(A) =_d N(A)$ for any A which is a finite disjoint union of rectangles in $\mathscr{E}$. By a result of Renyi (1967) (see Jagers, 1974, Proposition 4.2), it follows that $\tilde{N} =_d N$. This shows that every convergent subsequence has the same limit. Since N_n is tight we have $N_n \Rightarrow N$ as desired.

COROLLARY 3.1. *Suppose* ν *satisfies the condition of* (3.1) *and also* $\nu(x_0, \infty) = +\infty$. *If* $M_n(t) = \bigvee_{j \leqq k_n(t)} X_{n,j}$ *and* M_∞ *is an extremal process generated by the distribution* $F(x) = e^{-\nu(x,\infty)}$ (cf. *Resnick and Rubinovitch*, 1973), *then under* (3.1) *and* (3.2)

$$M_n \Rightarrow M_\infty \qquad in \quad D(0, \infty) .$$

PROOF. The result follows from the continuous mapping theorem (Billingsley, 1968, page 30) upon applying the $D(0, \infty)$ valued functional $g(N_n)(t) := \sup\{X_{n,j} \mid X_{n,j} > x_0, j \leqq k_n(t)\}$. That $g(N) =_d M_\infty$ is well known (cf. Resnick and Rubinovitch (1973), Resnick (1975), Weissman (1976)) and it remains to show $M_n - g(N_n) \to_P 0$. However

$$\begin{aligned} &P[\sup_{0 \leqq t \leqq 1} |M_n(t) - g(N_n)(t)| > 0] \\ &\qquad \leqq P[N_n((0, t] \times (x_0, \infty)) = 0] \to e^{-\nu(x_0,\infty)} = 0 . \end{aligned}$$

Of course, applying the appropriate functional, one would get joint convergence of the maxima, second maxima, etc.

EXAMPLE 3.1. Chain dependent variables (continued). Recall the setup of Example 2.1 but now we make no assumptions about moments of the H's. Define

$$X_{n,j} = (X_j - b_n)/a_n\,, \qquad k_n(t) = [nt]$$

where $a_n > 0$, b_n are normalizing constants. We seek conditions for $N_n \Rightarrow N$. Since $\mathscr{F}_{n,j} = \mathscr{J}_j$ the left side of (3.1) becomes

$$\begin{aligned}\textstyle\sum_{j=1}^{[nt]} P\{X_j > a_n x + b_n \mid J_{j-1}\} &= \textstyle\sum_{j=1}^{[nt]} (1 - H_{J_{j-1}}(a_n x + b_n)) \\ &= \textstyle\sum_{1=i}^{m} \pi_i([nt] - 1)(1 - H_i(a_n x + b_n))\end{aligned}$$

and since $\pi_i(n) \sim \pi_i n$ as $n \to \infty$ for $i = 1, \cdots, m$ we have that (3.1) holds iff

$$(3.4) \qquad n \textstyle\sum_{i=1}^{m} \pi_i(1 - H_i(a_n x + b_n)) \to \nu(x, \infty)\,.$$

It is well known that (3.4) requires $\exp\{-\nu(x, \infty)\}$ to be an extreme value distribution and the distribution $\sum_{i=1}^{m} \pi_i H_i(x)$ to be in the domain of attraction of $\exp\{-\nu(x, \infty)\}$. From (3.4) it is easy to check (3.2).

To summarize: if (3.4) holds then $N_n \Rightarrow N$ where N is Poisson with mean measure $dt \times d\nu$ where $\nu(x, \infty) = -\log F(x)$ and F is one of the three classes of extreme value distributions. Also $M_n := (\bigvee_{j=1}^{[n\cdot]} X_j - b_n)/a_n \Rightarrow M_\infty$ in $D(0, \infty)$ where M_∞ is an extremal F process.

One dimensional convergence of maxima of chain dependent variables has been considered by Resnick and Neuts (1970) and O'Brien (1974b), and Denzel and O'Brien (1975).

EXAMPLE 3.2. Let $\{E_k, -\infty < k < \infty\}$ be i.i.d. exponentially distributed variables with $P[E_k > x] = e^{-x}$, $x > 0$ and set $\varepsilon_k = E_k - 1$ and define $X_n = \sum_{k=0}^{\infty} \rho^k \varepsilon_{n-k}$ where $0 < \rho < 1$ (so that $X_n = \rho X_{n-1} + \varepsilon_n$). Set $X_{n,j} = X_j - \log n$, $n \geqq 1$, $j \geqq 1$ and $k_n(t) = [nt]$. Evaluating the left side of (3.1) we obtain

$$\textstyle\sum_{k=1}^{[nt]} e^{-(x+1)} e^{-\log n} e^{\rho X_{k-1}} = e^{-(x+1)} \sum_{k=1}^{[nt]} e^{\rho X_{k-1}}/n\,.$$

Since $\{e^{\rho X_k}\}$ is stationary, we obtain by the ergodic theorem that the above converges a.s. to

$$te^{-(x+1)} E e^{\rho X_0}$$

provided $Ee^{\rho X_0} < \infty$. The finiteness is not hard to check and in fact

$$Ee^{\rho X_0} = \exp\{-\rho(1 - \rho)^{-1}\} \textstyle\prod_1^\infty (1 - \rho^k)^{-1}\,.$$

To check (3.2) we compute for any $\delta > 0$

$$\begin{aligned}&P\{\max_{k \leqq nt} P\{X_{n,j} > x \mid \mathscr{F}_{n,j-1}\} > \delta\} \\ &\quad = P\{\textstyle\bigcup_{k=1}^{[nt]} [e^{-1} e^{-x} e^{\rho X_{k-1}}/n > \delta]\} \leqq [nt] P[e^{\rho X_0} > nc]\,, \qquad \text{where } c > 0\,.\end{aligned}$$

Now pick $\rho < \zeta < 1$ and the above has the Chebychev bound

$$\frac{[nt]Ee^{\zeta X_0}}{(nc)^{\zeta/\rho}} \to 0\,, \quad n \to \infty\,.$$

This verifies (3.2) and hence we conclude

$$\bigvee_{j=1}^{[nt]} X_j - \log n \Rightarrow M_\infty$$

where M_∞ is the extremal process generated by the distribution $\exp\{-(Ee^{\rho X_0})e^{-(x+1)}\}$.

4. Convergence to processes with stationary independent increments. In this section we will give conditions for Z_n to converge to a process with stationary independent increments. Our method of proof will be to obtain the convergence by applying the continuous mapping theorem to the convergence of the point processes obtained in Section 3. To do this we need the following facts about processes with stationary independent increments and their relationships to an associated Poisson random measure.

Let $\{Z(t), t \geqq 0\}$ be a process with stationary independent increments. The characteristic function of $Z(t)$ is given by

$$E(e^{i\theta Z(t)}) = \exp\left\{t\left[ia\theta + \frac{c\theta^2}{2} + \int_{|x|\geqq\gamma}(e^{i\theta x} - 1)\nu(dx) + \int_{0<|x|<\gamma}(e^{i\theta x} - 1 - i\theta x)\nu(dx)\right]\right\} \tag{4.1}$$

where a is a contsant, c is a nonnegative number and ν (called the Lévy measure) is a σ-finite measure on $R^0 = (-\infty, 0) \cup (0, \infty)$ with the property that $\int (x^2 \wedge 1)\nu(dx) < \infty$ (cf. Gnedenko and Kolmogorov, 1968, page 84).

Let N be a Poisson random measure on $[0, \infty) \times R^0$ with mean measure $dt \times d\nu$ and points $\{(t_k, \xi_k)\}$. Let W be a normalized Brownian motion independent of N. The Itô representation of Z (cf. Itô, 1969, page 1.7.7) is

$$Z(t) = at + W(ct) + \sum_{t_k\leqq t} \xi_k 1_{\{|\xi_k|\geqq\gamma\}} + \lim_{\delta\downarrow 0}\left[\sum_{t_k\leqq t} \xi_k 1_{\{|\xi_k|\in(\delta,\gamma)\}} - t\int_{|s|\in(\delta,\gamma)} s\nu(ds)\right] \tag{4.2}$$

where, for almost all ω, the convergence is uniform on compact t sets.

Having introduced the necessary preliminaries on processes with stationary independent increments, we are ready to state and prove theorems for convergence to these processes. In both of the results given the limits will have $a = 0$. In the first we will also suppose $c = 0$ (i.e., there is no Wiener component). The notation used in the statements below is that of Section 2 except that we have added a superscript δ to indicate the value at which the sequence is truncated (in Section 2 this value was ε_n). Throughout this section, we suppose $k_n(t)$ is a stopping time for each t.

THEOREM 4.1. *Let ν and Z be as specified in* (4.1) *and* (4.2) *with $a = c = 0$. Suppose that $\nu(\{-\gamma, \gamma\}) = 0$. If*

(a) *for all* $t > 0$

$$\sum_{i=1}^{k_n(t)} P[X_{n,i} > x \mid \mathscr{F}_{n,i-1}] \to_P t\nu(x, \infty)$$

and

$$\sum_{i=1}^{k_n(t)} P[X_{n,i} < y \mid \mathscr{F}_{n,i-1}] \to_P t\nu(-\infty, y)$$

whenever $x > 0$, $y < 0$ *and* $\nu(\{x, y\}) = 0$,

(b) *for all* $\varepsilon > 0$

$$\max_{1 \leq i \leq k_n(1)} P\{|X_{n,i}| > \varepsilon \mid \mathscr{F}_{n,i-1}] \to_P 0$$

and

(c) *for all* $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} P\{\sup_{0\leq s\leq 1} |\bar{Y}_n^\delta(s) - A_n^\delta(s)| > \varepsilon\} = 0$$

then $Z_n \Rightarrow Z$ *in* $D[0, 1)$.

A sufficient condition for (c) *is*

(d) *for all* $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} P\{\sum_{i=1}^{k_n(1)} E((\bar{\bar{X}}_{n,i}^\delta)^2 \mid \mathscr{F}_{n,i-1}) > \varepsilon\} = 0 .$$

REMARK. Other sufficient conditions for (c) can be obtained by using Doob's maximal inequality for martingales.

PROOF. We begin by disposing of a technical matter. We show that (a) implies

$$\sum_{i=1}^{k_n(t)} E(X_{n,i} 1_{\{\delta < |X_{n,i}| < \gamma\}} \mid \mathscr{F}_{n,i-1}) \to_P t \int_{|s| \in (\delta,\gamma)} s\nu(ds) \tag{4.3}$$

for all $t \geqq 0$ whenever $\nu(\{-\delta, \delta\}) = 0$. One can proceed as in Brown and Eagleson (1971) by approximating above and below by step functions but we will use the following approach which will be useful later as well: observe that the set of Borel-Radon measures on $[0, \infty) \times R^0$ is a complete separable metric space. From (a) and Theorem 3, page 206 of Jagers (1974) it follows that the random measures

$$\mu_n([0, t] \times A) := \sum_{i=1}^{k_n(t)} P\{X_{n,i} \in A \mid \mathscr{F}_{n,i-1}\}$$

defined on $[0, \infty) \times R^0$ converge weakly to the measure $dt \times d\nu$. The map

$$\mu \to \int_{|s| \in (\delta,\gamma)} s\mu([0, t] \times ds)$$

is continuous at each μ with $\mu([0, 1] \times \{-\delta, \delta, -\gamma, \gamma\}) = 0$ so by the continuous mapping theorem (Billingsley, 1968, page 30)

$$\begin{aligned}\sum_{i=1}^{k_n(t)} E(X_{n,i} 1_{\{\delta < |X_{n,i}| < \gamma\}} \mid \mathscr{F}_{n,i-1}) &= \int_{|s| \in (\delta,\gamma)} s\mu_n([0, t] \times ds) \\ &\Rightarrow \int_{|s| \in (\delta,\gamma)} s\nu(ds)\end{aligned}$$

and convergence in probability is automatic since the limit is constant.

From (a), (b) and Theorem 3.1 we get convergence of the point processes $N_n \Rightarrow N$ where N is the Poisson process on $[0, \infty) \times R^0$ described before Theorem 4.1 and $N_n([0, t] \times [a, b]) = \sum_{i=1}^{k_n(t)} 1_{\{X_{n,i} \in [a,b]\}}$. If $\nu(\{-\delta, \delta\}) = 0$ then we can apply the continuous functional which sums ordinates of points in

$\{(s, x): s \leqq t, |x| > \delta\}$ and conclude from the continuous mapping theorem that

$$\sum_{i=1}^{k_n(\cdot)} X_{n,i} 1_{\{|X_{n,i}|>\delta\}} \Rightarrow \sum_{t_k \leqq \cdot} \xi_k 1_{\{|\xi_k|>\delta\}} .$$

Combining this result with (4.3) gives

$$\begin{aligned} Z_{\delta,n} &= \sum_{i=1}^{k_n(\cdot)} X_{n,i} 1_{\{|X_{n,i}|>\delta\}} - \sum_{i=1}^{k_n(\cdot)} E(X_{n,i} 1_{\{\delta<|X_{n,i}|<\gamma\}} \mid \mathscr{F}_{n,i-1}) \\ &\Rightarrow Z_\delta = \sum_{t_k \leqq \cdot} \xi_k 1_{\{|\xi_k|>\delta\}} - (\cdot) \int_{\delta<|s|<\gamma} s\nu(ds) . \end{aligned}$$

Since $Z_\delta \to Z$ almost surely and uniformly on compact t sets we have $Z_\delta \Rightarrow Z$ and applying Theorem 4.2 in Billingsley, 1968, we will have $Z_n \Rightarrow Z$ provided

$$(4.4) \qquad \lim_{\delta\to 0} \limsup_{n\to\infty} P\{\rho(Z_{\delta,n}, Z_n) > \varepsilon\} = 0$$

where ρ is the Skorohod metric for $D[0, 1]$. However $Z_n - Z_{\delta,n} = \sum_{j=1}^{k_n(\cdot)} (\bar{X}_{n,j}^\delta - E\bar{X}_{n,j}^\delta) = \bar{Y}_n^\delta - A_n^\delta$ so (c) implies (4.4) and the proof of the first statement is complete.

To prove that (d) is sufficient for (c) observe that by the Doob decomposition of square integrable submartingales and a formula on page 150 of Neveu (1975) the probability in (c) is dominated by

$$\begin{aligned} P\{\textstyle\sum_{i=1}^{k_n(1)} E((\bar{\bar{X}}_{n,i}^\delta)^2 \mid Z_{n,i-1}) > \varepsilon^2\} + E\{1 \wedge \varepsilon^{-2} \sum_1^{k_n(1)} E((\bar{\bar{X}}_{n,i}^\delta) \mid Z_{n,i-1})\} \\ = (1) + (2) . \end{aligned}$$

However for any $1 > \xi > 0$, (2) is

$$\begin{aligned} &\leqq E\{1 \wedge \varepsilon^{-2} \textstyle\sum_1^{k_n(1)} E((\bar{\bar{X}}_{n,i}^\delta)^2 \mid Z_{n,i-1}); \varepsilon^{-2} \sum_1^{k_n(1)} < \xi\} + P\{\varepsilon^{-2} \sum_1^{k_n(1)} > \xi\} \\ &\leqq \xi + P[\varepsilon^{-2} \textstyle\sum_1^{k_n(1)} > \xi] . \end{aligned}$$

So because of (d)

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} P\{\sup_{0\leqq s\leqq 1} |Y_n^\delta(s) - A_n^\delta(s)| > \varepsilon\} < \xi$$

and since ξ is arbitrarily small, (c) is verified.

EXAMPLE 4.1. Chain dependent variables. Continuing the developments of Examples 2.1 and 3.1 let $H = \sum_1^m \pi_i H_i$ and suppose

$$1 - H(x) + H(-x) \sim x^{-\alpha} L(x) , \qquad 0 < \alpha < 2$$

$$\frac{1 - H(x)}{1 - H(x) + H(-x)} \to p , \qquad \frac{H(-x)}{1 - H(x) + H(-x)} \to q$$

as $x \to \infty$ where L is slowly varying and $p + q = 1$. As in Example 3.1 it is readily verified that (a) and (b) of Theorem 4.1 hold with

$$\begin{aligned} \nu(x, \infty) &= px^{-\alpha} \quad \text{for } x > 0 \\ \nu(-\infty, -y) &= qy^{-\alpha} \quad \text{for } y > 0 , \end{aligned}$$

where

$$X_{n,j} = X_j/a_n \quad \text{and} \quad a_n \quad \text{is chosen so that}$$
$$n(1 - H(a_n x) + H(-a_n x)) \to x^{-\alpha} , \quad x > 0 .$$

We check condition (d) as follows:

$$\begin{aligned}\textstyle\sum_1^n E((\bar{\bar{X}}_{n,i}^{\delta})^2 \mid \mathscr{F}_{n,i-1})&\\ &\leqq \textstyle\sum_1^n E\{X_{n,i}^2 1_{\{|X_{n,i}|\leqq\delta\}} \mid \mathscr{F}_{n,i-1}\}\\ &= \textstyle\sum_1^m \int_{|x|\leqq\delta} x^2\pi_i(n)\, dH_i(a_n x) \sim n \int_{|x|\leqq a_n\delta} \frac{x^2}{a_n^2}\, d(\sum_1^m \pi_i H_i(x))\\ &= \frac{n}{a_n^2} \int_{|x|\leqq a_n\delta} x^2\, dH(x)\end{aligned}$$

which by the lemma on page 578 of Feller (1971) is asymptotic to

$$\begin{aligned}\frac{n}{a_n^2}\left(\frac{\alpha}{2-\alpha}\right)(\delta a_n)^2(1 - H(a_n\delta) + H(-a_n\delta))&\\ &= \frac{\alpha\delta^2}{2-\alpha}\, n(1 - H(a_n\delta) + H(-a_n\delta)) \to \frac{\alpha\delta^{2-\alpha}}{2-\alpha} \qquad \text{as} \quad n \to \infty\end{aligned}$$

and since $2 - \alpha > 0$ we have as $\delta \downarrow 0$ that the above $\to 0$.

Thus from Theorem 4.1 we have

$$\textstyle\sum_1^{[n\cdot]} \frac{X_j}{a_n} - \sum_1^{[n\cdot]} E\left(\frac{X_j}{a_n} 1_{\{|X_j/a_n|<\gamma\}} \middle| \mathscr{F}_{n,j-1}\right) \Rightarrow Z_\alpha$$

where Z_α is a stable process with characteristic function given by (4.1) with $a = 0 = c$ and ν specified as in the first part of the example.

When $\alpha < 1$ we observe by formula (5.22) on page 579 of Feller (1971) that

$$\textstyle\sum_1^{[nt]} E\left(\frac{X_j}{a_n} 1_{\{|X_j/a_n|<\gamma\}} \middle| \mathscr{F}_{n,j-1}\right) \to t\, \frac{\alpha}{1-\alpha}(p - q)\gamma^{1-\alpha}$$

a.s. and locally uniformly in t.

When $\alpha > 1$, we use formula (5.21) on page 579 and (5.16) on page 577 of Feller to obtain:

$$\frac{n}{a_n} \int_{|x|>\gamma a_n} x\, dH \to \left(\frac{\alpha}{\alpha-1}\right)(p - q)\gamma^{1-\alpha}$$

so that

$$\textstyle\sum_1^{[nt]} E\left(\frac{X_j}{a_n} 1_{\{|X_j/a_n|<\gamma\}} \middle| \mathscr{F}_{n,j-1}\right) - nt \sum_1^m \pi_i \int x\, dH_i \to \left(\frac{\alpha}{\alpha-1}\right)(p - q)\gamma^{1-\alpha} t$$

a.s. and locally uniformly in t.

To sum up:

$$\textstyle\sum_1^{[n\cdot]} X_j/a_n - \frac{t\alpha}{1-\alpha}(p - q)\gamma^{1-\alpha} \Rightarrow Z_\alpha\,, \quad 0 < \alpha < 1$$

$$\textstyle\sum_1^{[n\cdot]} X_j/a_n - nt \sum_1^m \pi_i \int x\, dH_i + \left(\frac{\alpha}{\alpha-1}\right)(p - q)\gamma^{1-\alpha} t \Rightarrow Z_\alpha\,, \quad 1 < \alpha < 2$$

$$\textstyle\sum_1^{[n\cdot]} X_j/a_n - \frac{tn}{a_n} \int_{|x|<\gamma a_n} x\, dH \Rightarrow Z_\alpha\,, \quad \alpha = 1\,.$$

Wolfson (1974) has shown that partial sums of chain dependent variables can converge only to stable laws.

THEOREM 4.2. *Let* ν *and* Z *be as specified in* (4.1) *and* (4.2) *with* $a = 0$. *Suppose that* $\nu(\{-\gamma, \gamma\}) = 0$. *If conditions* (a) *and* (b) *of Theorem* 4.1 *are satisfied and* (c) *for all* ε *and* $t > 0$

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} P\{|\sum_{i=1}^{k_n(t)} E((\bar{\bar{X}}_{n,i}^\delta)^2 \mid \mathscr{F}_{n,i-1}) - ct| > \varepsilon\} = 0$$

then $Z_n \Rightarrow Z$.

If (d) *for all* $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} P\{\sum_{i=1}^{k_n(1)} E(\bar{X}_{n,i}^\delta \mid \mathscr{F}_{n,i-1})^2 > \varepsilon\} = 0$$

then (a) *and* (c) *are equivalent to*

(e) *there is a nondecreasing function* G *such that if* x *and* y *are continuity points of* G *then*

$$\sum_{i=1}^{k_n(t)} E(X_{n,i}^2 1_{\{x<|X_{n,i}|<y\}} \mid \mathscr{F}_{n,i-1}) \to_P t[G(y) - G(x)]\,.$$

REMARK. Conditions (a)—(c) are the analogues of those given by Gnedenko and Kolmogorov (1968) in the case of independence (see page 124). Condition (e) is from Brown and Eagleson (1971) who studied the case in which the limit law has finite variance. They assumed some other conditions which include (b) and imply (d).

PROOF. From assumption (c) if we let $\varepsilon_n \downarrow 0$ slowly enough then

$$\sum_{i=1}^{k_n(t)} E(\bar{\bar{X}}_{n,i}^2 \mid \mathscr{F}_{n,i-1}) \to_P ct \tag{4.5}$$

for all $t > 0$. (Here we have returned to the practice of deleting the superscript ε_n.) Checking Theorem 4.1 we find that $\hat{Y}_n - B_n^\gamma$ converges weakly to $\hat{Z} := Z - W(c\cdot)$ (replace $X_{n,i}$ in Theorem 4.1 by $X_{n,i} 1_{\{\varepsilon_n < |X_{n,i}|\}}$). From formula (2.2)

$$Z_n = \hat{Y}_n - B_n^\gamma + W_n \circ \varphi_n\,.$$

By (4.5) we have $\varphi_n(t) \to_P ct$ so that $W_n \circ \varphi_n \Rightarrow W(c\cdot)$.

To complete the proof we have to show that

$$(W_n, \hat{Y}_n - B_n^\gamma) \Rightarrow (W, \hat{Z})$$

where W and $\hat{Z}$ are independent. To do this it suffices to show that if ε_n is the sequence we have used above then for all $M < \infty$ which are continuity points of ν we have that the finite dimensional distributions of

$$\sum_{i=1}^{k_n(t)} X_{n,i} 1_{\{X_{n,i} \in [-M,M]\}}$$

converge to those of a process which is the sum of a Brownian motion with variance c and an independent pure jump process with Lévy measure $\nu(\cdot \cap [-M, M])$. The desired result is then a consequence of Theorem 1 of Brown and Eagleson (1971) so the proof is complete.

REFERENCES

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

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

[3] BROWN, B. M. and EAGLESON, G. K. (1971). Martingale convergence to infinitely divisible laws with finite variances. *Trans. Amer. Math. Soc.* **162** 449–453.

[4] DENZEL, G. E. and O'BRIEN, G. L. (1975). Limit theorems for extreme values of chain dependent process. *Ann. Probability* **3** 773–780.

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

[6] DVORETZKY, A. (1969). Central limit theorems for dependent random variables and some applications. Abstract: 1969 Rietz Lecture, Institute of Mathematical Statistics.

[7] DVORETZKY, A. (1972). Asymptotic normality for sums of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probability* **2** 513–535. Univ. of California Press.

[8] EAGLESON, G. K. (1975). Martingale convergence to mixtures of infinitely divisible laws. *Ann. Probability* **3** 557–562.

[9] EAGLESON, G. K. (1976). The martingale central limit theorem and random norming. Preprint: Statistical Laboratory, Cambridge University.

[10] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications Vol. II, 2nd ed.* Wiley, New York.

[11] FREEDMAN, D. (1971). *Brownian Motion and Diffusion.* Holden-Day, San Francisco.

[12] FREEDMAN, D. (1974). The Poisson approximation for dependent events. *Ann. Probability* **2** 256–269.

[13] FREEDMAN, D. (1975). On tail probabilities for martingales. *Ann. Probability* **3** 100–118.

[14] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). *Limit Distributions for Sums of Independent Random Variables.* Translated by K. L. Chung, Addison-Wesley, Menlo Park, California.

[15] HALL, P. (1975). Martingale invariance principles. Preprint: Mathematics Institute, Oxford University.

[16] HELLAND, I. S. (1976). Continuity of a class of random time transformations. Preprint: University of Bergen.

[17] HELLAND, I. S. (1977). Weak convergence to diffusions. (In preparation).

[18] ITÔ, K. (1969). *Stochastic Processes.* Lecture Notes Series No. 16, Matematisk Institut, Aarhus Universitet.

[19] JAGERS, P. (1974). Aspects of random measures and point processes. *Advances in Probability* **3** (eds., Ney and Port). Marcel Dekker, New York.

[20] KEILSON, J. and WISHART, D. M. G. (1964). A central limit theorem for processes defined on a Markov chain. *Proc. Cambridge Phil. Soc.* **60** 547–567.

[21] LAMPERTI, J. (1967). Limiting distributions for branching processes. *Proc. Fifth Berkeley Symp. Math. Statist. Probability* Vol. II, Part II, 225–242.

[22] LÉVY, P. (1937). *Theorie de l'Addition de Variables Aleatories.* Gauthier-Villars, Paris.

[23] LINDVAAL, T. (1972). Convergence of critical Galton-Watson processes. *J. Appl. Probability* **9** 445–450.

[24] LINDVAAL, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0, \infty)$. *J. Appl. Probability* **10** 109–121.

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

[26] NEVEU (1975). *Discrete parameter martingales.* North Holland.

[27] O'BRIEN, G. (1974a). Limit theorems for sums of chain dependent processes. *J. Appl. Probability* **11** 582–587.

[28] O'BRIEN, G. (1974b). The limiting distribution of maxima of random variables defined on a denumerable Markov chain. *Ann. Probability* **2** 103–111.

[29] RENYI, A. (1967). Remarks on the Poisson process. *Studia Sci. Math. Hungar.* **2** 119–123.

[30] RESNICK, S. I. (1975). Weak convergence to extremal processes. *Ann. Probability* **3** 951–960.

[31] RESNICK, S. I. and NEUTS, M. F. (1970). Limit laws for maxima of a sequence of random variables defined on a Markov chain. *Advances in Appl. Probability* **2** 323–343.

[32] RESNICK, S. I. and RUBINOVITCH, M. (1973). The structure of extremal processes. *Advances in Appl. Probability* **5** 287–307.

[33] ROOTZÉN, H. (1974). Renyi-mixing and functional limit theorems for martingales. Preprint 1974: 8, Depart. Math. Statistics, Univ. Lund and Lund Institute of Technology.

[34] ROOTZÉN, H. (1976). Fluctuations of sequences which converge in distribution. *Ann. Probability* **4** 456–463.

[35] ROOTZÉN, H. (1977a). On the functional central limit theorem for martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** 199–210.

[36] ROOTZÉN, H. (1977b). A note on convergence to mixtures of normal distributions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** 211–216.

[37] SCOTT, D. J. (1973). Central limit theorems for martingales and for processes with stationary increments using a Skorohod representation approach. *Advances in Appl. Probability* **5** 119–137.

[38] STROOCK, D. W. and VARADHAN, S. R. S. (1969). A martingale approach to defining diffusions, II. *Comm. Pure Appl. Math.* **22** 479–530.

[39] STROOCK, D. W. and VARADHAN, S. R. S. (1977). *Diffusion Processes.* Springer-Verlag. To appear.

[40] WEISSMAN, I. (1976). On weak convergence of extremal processes. *Ann. Probability* **4** 470–473.

[41] WOLFSON, D. (1974). Limit theorems for sums of a sequence of random variables defined on a Markov chain. Ph. D. Thesis, Purdue University.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305