FUNCTIONAL LIMIT THEOREMS FOR DEPENDENT VARIABLES

BY

RICHARD DURRETT and SIDNEY I. RESNICK

TECHNICAL REPORT NO. 95
DECEMBER 9, 1976

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MPS74-21416

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
FUNCTIONAL LIMIT THEOREMS FOR DEPENDENT VARIABLES

BY

RICHARD DURRETT and SIDNEY I. RESNICK

TECHNICAL REPORT NO. 95
DECEMBER 9, 1976

PREPARED UNDER THE AUSPICIES
OF
NATIONAL SCIENCE FOUNDATION GRANT MPS74-21416

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
FUNCTIONAL LIMIT THEOREMS FOR DEPENDENT VARIABLES

by

Richard Durrett* and Sidney I. Resnick**

University of California, Los Angeles and Stanford University

ABSTRACT

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 (Theorem 4.2) or to a process with continuous paths (Theorem 2.2). 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.

*Partially supported by NSF GRANT No. MCS 72-04591.
**Partially supported by NSF Grant No. MPS 74-21416.

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

Keywords and phrases: Weak convergence, invariance principle, Poisson process, Brownian motion; process with stationary, independent increments; extremal process, martingales.
1. Introduction

Let \( \{X_{n,i}, n \geq 1, i \geq 1\} \) be an array of random variables defined on a probability space \((\Omega, \mathcal{F}, P)\) and let \( \{\mathcal{F}_{n,i}, n \geq 1, i \geq 0\} \) be an array of sub-\(\sigma\)-fields of \( \mathcal{F} \) such that for each \( n \) and \( i \geq 1 \), \( X_{n,i} \) is \( \mathcal{F}_{n,i} \) measurable and \( \mathcal{F}_{n,1-i} \subseteq \mathcal{F}_{n,i} \). Suppose \( k_n(t) \) is a non-decreasing right continuous function with range \([0,1,2,\ldots]\). The functions \( k_n(t) \) are given time scales. Set

\[
S_{n,0} = 0, \quad S_{n,k} = \sum_{i=1}^{k} X_{n,i}, \quad k \geq 1
\]

\[
Y_{n}(t) = S_{n,k_n(t)},
\]

\[
Z_{n}(t) = S_{n,k_n(t)} - \sum_{i=1}^{k_n(t)} E(X_{n,i} \mid \{ |X_{n,i}| < \gamma \} \mid \mathcal{F}_{n,i-1})
\]

\[
M_{n,0} = -\infty, \quad M_{n,k} = \bigvee_{i=1}^{k} X_{n,i}, \quad k \geq 1
\]

\[
M_{n}(t) = M_{n,k_n(t)}
\]

where \( \gamma > 0 \). In this paper we give conditions for \( \{Y_{n}(t), 0 \leq t \leq 1\} \) and \( \{Z_{n}(t), 0 \leq t \leq 1\} \) 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], [22], [29]-[33]) have given results for a variety of time scales 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 \( \{ \mathcal{F}_{n,k'} \}_{k' \geq 1} \). 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. These results will be considered in a future publication.
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 to processes with continuous paths. Recall that in this section we assume $k_n(1)$ is a stopping time for $\{\mathcal{F}_{n,k}, k \geq 0\}$.

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

Definition. A collection of random variables $X_{n,i}$, $n \geq 1$, $i \geq 1$ and $\sigma$-fields $\mathcal{F}_{n,i}$, $n \geq 1$, $i \geq 0$ is said to be a martingale difference array if

(i) for all $n \geq 1$, $\mathcal{F}_{n,i}$, $i \geq 0$ is an increasing sequence of $\sigma$-fields
(ii) for all $n$, $i \geq 1$, $X_{n,i}$ is $\mathcal{F}_{n,i}$ measurable and
(iii) for all $n$, $i \geq 1$, $E(X_{n,i} | \mathcal{F}_{n,i-1}) = 0$. 


Theorem 2.1. (Freedman). Let \( \{X_{n,i} : \mathcal{F}_{n,i}\} \) be a martingale difference array and suppose there are numbers \( \epsilon_n \leq 0 \) so that \( |X_{n,i}| \leq \epsilon_n \) for all \( n \) and \( i \). Let \( V_n,0 = 0, V_n,j = \sum_{i=1}^{j} E(X_{n,i}^2 | \mathcal{F}_{n,i-1} \) and \( j_n(t) = \sup\{j : V_n,j \leq t\} \). If \( P(\lim_{j \to \infty} V_n,j = \infty) = 1 \) then as \( n \to \infty \), \( S_n, j_n(.) \) converges weakly 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 \( \epsilon_n \) be a sequence of positive numbers which decrease to zero and define

\[
\hat{X}_{n,i} = X_{n,i} 1\{|X_{n,i}| > \epsilon_n\}
\]

\[
\check{X}_{n,i} = X_{n,i} 1\{|X_{n,i}| \leq \epsilon_n\}
\]

\[
\bar{X}_{n,i} = \check{X}_{n,i} - E(\check{X}_{n,i} | \mathcal{F}_{n,i-1}) .
\]

Then \( |\bar{X}_{n,i}| \leq 2 \epsilon_n \) so if we let \( \bar{S}_{n,k} = \sum_{i=1}^{k} \bar{X}_{n,i} \) and

\[
j_n(t) = \sup\{j : V_n,j = \sum_{i=1}^{j} E(\bar{X}_{n,i}^2 | \mathcal{F}_{n,i-1} \leq t\},
\]

we have from Theorem 2.1 as \( n \to \infty \), \( W_n := \bar{S}_n, j_n(.) \Rightarrow W \).
Next set
\[
\hat{Y}_n(t) = \sum_{j=1}^{k_n(t)} \hat{X}_{n,j},
\]
\[
\bar{Y}_n(t) = \sum_{j=1}^{k_n(t)} \bar{X}_{n,j},
\]
\[
A_n(t) = \sum_{j=1}^{k_n(t)} E(\bar{X}_{n,j} | \mathcal{F}_{n,j-1})
\]
and let \( \varphi_n(t) \) be any strictly increasing continuous function which satisfies \( \bar{Y}_n(t) = \bar{Y}_n(\varphi_n(t)) \). From the last three definitions it is immediate that

\[
Y_n(t) = W_n(\varphi_n(t)) + \hat{Y}_n(t) + A_n(t)
\]

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

\[
Z_n(t) = W_n(\varphi_n(t)) + \hat{Y}_n(t) - B\bar{Y}(t).
\]

**Theorem 2.2.** If for some \( \epsilon_n \to 0 \)

(i) \( P( \max_{0 \leq s \leq 1} |\hat{Y}_n(s)| > 0 ) \to 0 \) and

(ii) \( (W_n, \varphi_n) \Rightarrow (W, \varphi) \) with \( P( \varphi \text{ is continuous}) = 1 \)

then \( Z_n \Rightarrow W \circ \varphi \). If in addition

(iii) \( A_n \xrightarrow{P} 0 \) or what is equivalent for some \( \lambda > 0 \)

\[
k_n(t) = \sum_{i=1}^{k_n(t)} E(X_{n,i} 1_{\{|X_{n,i}| < \lambda \}} | \mathcal{F}_{n,i-1}) \xrightarrow{P} 0
\]

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), $\hat{Y}_n \overset{P}{\to} 0$ so to prove the first result it remains to show that $B_n \overset{P}{\to} 0$.

To do this it suffices to show that

$$\sum_{j=1}^{k_n(1)} P(|X_{n,j}| > \epsilon_n | \mathcal{F}_{n,j-1}) \overset{P}{\to} 0$$

Let $\tau_n$ be the time of the first jump of size $> \epsilon_n$, that is

$$\tau_n = \sup\{t \leq 1 : \sum_{j=1}^{k_n(t)} 1\{|X_{n,j}| > \epsilon_n\} = 0\}$$

Then

$$\sum_{j=1}^{k_n(\tau_n)} P(|X_{n,j}| > \epsilon_n | \mathcal{F}_{n,j-1})$$

$$= E\left( \sum_{j=1}^{\infty} 1\{k_n(1) \geq j, |X_{n,i}| \leq \epsilon_n \text{ for } 1 \leq i < j\} P(|X_{n,j}| > \epsilon_n | \mathcal{F}_{n,j-1}) \right)$$

Since $\{k_n(1) \geq j, |X_{n,i}| \leq \epsilon_n \text{ for } 1 \leq i < j\} \in \mathcal{F}_{n,j-1}$ the last formula

$$= P\left( \max_{1 \leq j \leq k_n(1)} |X_{n,j}| > \epsilon_n \right) = P\left( \max_{0 \leq s \leq 1} |\hat{Y}_n(s)| > 0 \right) \to 0$$

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

$$\sum_{j=1}^{k_n(1)} P(|X_{n,j}| > \epsilon_n | \mathcal{F}_{n,j-1}) > \delta$$

$$\leq P(\tau_n < 1) + \delta^{-1} P\left( \max_{1 \leq j \leq k_n(1)} |X_{n,j}| > \epsilon_n \right).$$
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 \( Y_n = Z_n + B_n X_n + A_n \) 
and \( A_n = C_n - B_n \). From the first part of the proof \( B_n Y_n \overset{P}{\to} 0 \). 
Therefore the two assumptions in (iii) are equivalent and under either 
one \( Y_n - Z_n \overset{P}{\to} 0 \) 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 Lindberg-Feller theorem.

**Theorem 2.3.** Suppose \( \{X_{n,i} | F_{n,i}^k\} \) is a martingale difference array.

If

(a) for all \( t > 0 \), \( \sum_{i=1}^{k} E(X_{n,i}^2 | F_{n,i}^k) \overset{P}{\to} ct \) and

(b) for all \( \epsilon > 0 \)

\[
\sum_{i=1}^{k_n(\epsilon)} E(X_{n,i}^2 1(|X_{n,i}| > \epsilon) | F_{n,i-1}^k) \overset{P}{\to} 0
\]

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

**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 \( s_i^2 = \text{ES}_i^2 < \infty \) for all \( i \) and let \( k_n(t) = k \) when \( s_k^2 \leq t s_n^2 < s_{k+1}^2 \). Theorem 2.3 gives one of the results 
of Scott ([33], p. 120). If we let \( k_n(t) \) be the time scale \( j_n(t) \)
defined in Theorem 2.2 we get Theorem 5 of Rootzen ([30], p. 11, see 
his Remark 6). It is trivial to generalize our result to obtain 
Theorem 3.8 of McLeish ([22], p. 626).
Proof. From (b) it follows that if we let \( \epsilon_n \) decrease to zero slowly enough then

\[
(2.3) \quad \sum_{i=1}^{k_n(1)} \epsilon_n^{-2} \epsilon_n \sum_{i=1}^{k_n(1)} \mathbb{E}(X_{n,i}^2 \mathbb{1}_{\{X_{n,i} > \epsilon_n\}} | \mathcal{F}_{n,i-1}) \xrightarrow{P} 0
\]

To prove that \( S_n, k_n(\cdot) \Rightarrow W(c) \) we will show that if \( \epsilon_n \downarrow 0 \) 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 \( N \geq 1 \)

\[
P\left( \max_{1 \leq i \leq N} |X_{n,i}| > \epsilon_n \right) \leq \delta + \mathbb{P} \left( \sum_{i=1}^{N} \mathbb{P}(|X_{n,i}| > \epsilon_n | \mathcal{F}_{n,i-1}) > \delta \right)
\]

Applying this formula to the martingale difference array

\[
X'_n,i = X_{n,i} \mathbb{1}_{\{k_n(1) > 1\}}
\]

and letting \( N \to \infty \) gives

\[
(2.4) \quad \mathbb{P}\left( \max_{1 \leq i \leq k_n(1)} |X_{n,i}| > \epsilon_n \right) \leq \delta + \mathbb{P}\left( \sum_{i=1}^{k_n(1)} \mathbb{P}(|X_{n,i}| > \epsilon_n | \mathcal{F}_{n,i-1}) > \delta \right).
\]

Since

\[
\sum_{i=1}^{k_n(1)} \epsilon_n^2 \mathbb{P}(|X_{n,i}| > \epsilon_n | \mathcal{F}_{n,i-1}) \leq \sum_{i=1}^{k_n(1)} \mathbb{E}(X_{n,i}^2 \mathbb{1}_{\{X_{n,i} > \epsilon_n\}} | \mathcal{F}_{n,i-1})
\]

it follows from (2.3) that \( \lim \sup \) of the right-hand side of (2.4) is less than \( \delta \). Since \( \delta \) is an arbitrary positive number, this shows that (i) of Theorem 2.2 holds.
To check (ii) we observe that it suffices to show 
\[ \varphi_n \xrightarrow{P} \varphi \] where \( \varphi(t) = ct \) (Billingsley, 1968, Theorem 4.4).

Now \( \varphi_n \) is a monotone function and \( \varphi \) is continuous so it suffices to show that \( \varphi_n(t) \xrightarrow{P} ct \) for each \( t > 0 \) or equivalently that

\[ \sum_{j=1}^{k_n(t)} E(\bar{X}_{n,j}^2 | \mathcal{F}_{n,j-1}) \xrightarrow{P} ct. \]

The left side of the above equals

\[ \sum_{j=1}^{k_n(t)} E(\bar{X}_{n,j}^2 | \mathcal{F}_{n,j-1}) - \sum_{j=1}^{k_n(t)} (E(\bar{X}_{n,j} | \mathcal{F}_{n,j-1}))^2. \]

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

\[ \sum_{j=1}^{k_n(t)} (E(X_{n,j} | \{X_{n,j} \mid \geq \epsilon_n \} | \mathcal{F}_{n,j-1}))^2 \]

\[ \leq \sum_{j=1}^{k_n(t)} E(X_{n,j}^2 \mathbf{1}\{X_{n,j} \mid > \epsilon_n \} | \mathcal{F}_{n,j-1}) \xrightarrow{P} 0 \]

by (2.3).

To complete the proof it remains to verify that \( A_n \xrightarrow{P} 0 \).

To do this we observe that
\[ \sup_{0 \leq t \leq 1} |A_n(t)| \leq \sum_{i=1}^{k_n(1)} \left| E(X_{n,i} 1\{X_{n,i} \leq \epsilon_n\} | \mathcal{F}_{n,i-1}) \right| \]

\[ = \sum_{i=1}^{k_n(1)} \left| E(X_{n,i} 1\{X_{n,i} > \epsilon_n\} | \mathcal{F}_{n,i-1}) \right| \]

\[ \leq \sum_{i=1}^{k_n(1)} E(|X_{n,i} | 1\{X_{n,i} > \epsilon_n\} | \mathcal{F}_{n,i-1}) \]

\[ \leq \epsilon_n^{-1} \sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1\{X_{n,i} > \epsilon_n\} | \mathcal{F}_{n,i-1}) \cdot \]

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

**Corollary 2.1.** If \{X_{n,i}, \mathcal{F}_{n,i}\} is a martingale difference array and

\[ \frac{\sum_{i=1}^{k_n(1)} E(X_{n,i}^2 | \mathcal{F}_{n,i-1})}{P} \to 0 \]

then \(Y_n \xrightarrow{P} 0\).

**Example 2.1.** Chain dependent variables: Let \{J_n, n \geq 0\} be an m-state aperiodic, irreducible Markov chain with transition matrix \([p_{i,j}, 1 \leq i, j \leq m]\) and stationary distribution \(\pi = (\pi_1, \ldots, \pi_m)\).

The random variables \{X_n, n \geq 1\} are chain dependent if for \(\mathcal{J}_n = \mathcal{B}(J_0, \ldots, J_n, X_1, \ldots, X_n)\) and \(n \geq 1\) we have

\[ P(J_n = j, X_n \leq x | \mathcal{J}_{n-1}) = P(J_n = j, X_n \leq x | J_{n-1}) = p_{J_{n-1},j} H_{J_{n-1}}(x) \]
where $H_1, \ldots, H_m$ are given distribution functions. It then follows that

$$P(X_n \leq x \mid \bigcup_{i=1}^{n-1} J_i) = F(X_n \leq x \mid J_{n-1}) = H_{n-1}(x)$$

and

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

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

$$E[X_{n,i} \mid \bigcup_{i-1}^{i-1}] = \int xH_i \frac{(dx)}{i-1} = 0$$

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

$$n^{-1} \sum_{i=1}^{[nt]} E(X_i^2 \mid \bigcup_{i-1}^{i-1}) = n^{-1} \sum_{i=1}^{[nt]} \sigma_i^2 \Rightarrow \sum_{j=1}^{m} \pi_j \sigma_j^2 \text{ a.s.}$$

and condition (a) of Theorem 2.3 is satisfied.

To check condition (b) observe that

$$E(X_{n,i}^2 1\{|X_{n,i}| > \varepsilon\} \mid J_{i-1} = j) = \frac{1}{n} \int_{\varepsilon < \sqrt{n}} x^2H_j(dx)$$

11
so
\[
\sum_{i=1}^{n} E(\bar{X}_{n,i}^2 1[|X_{n,i}| > \epsilon] | \mathcal{F}_{i-1}) \leq \sum_{j=1}^{m} \int_{|x| > \epsilon \sqrt{n}} x^2 H_j(dx) \rightarrow 0.
\]

Since conditions (a) and (b) are satisfied we conclude
\[
\sum_{i=1}^{[n]} \frac{X_i}{\sqrt{n}} \Rightarrow W(\sum_{i=1}^{m} \pi_i \xi_i).
\]

It is not hard to obtain asymptotic independence of
\[
\left\{ \frac{S_{[n]}}{\sqrt{n}} = \sum_{i=1}^{[n]} \frac{X_i}{\sqrt{n}} \right\}
\]
and \{J_n\} as follows. Let \( B \) be a Borel subset of \( D[0,1] \) such that \( P[W \in \mathcal{B}] = 0 \). Then
\[
P(S_{[n]}^*/\sqrt{n} \in B, J_n = 1) = E(1_{\{J_n = 1\}} P(S_{[n]}^*/\sqrt{n} \in B | J))
\]
where \( J = \mathcal{B}(J_n, n \geq 0) \). With respect to (a regular version of) \( P(\cdot | J) \(\omega \) we have for \( a, a, \omega \) that \( \{X_n, n \geq 1\} \) may be considered independent with distributions \( H_1, \ldots, J_n \) and the asymptotic frequency with which \( H_i \) appears in this list is \( \pi_i \). We can then see, as in the proof above, that for a.a.\( \omega \), \( P(S_n^*/\sqrt{n} \in B | J) \(\omega \) \rightarrow P(W \in B) \). Since \( E 1_{\{J_n = 1\}} \rightarrow \pi_i \) we obtain the desired asymptotic independence by dominated convergence:
\[
P(S_{[n]}^*/\sqrt{n} \in B, J_n = i) \rightarrow \pi_i P(W \in B).
\]
Asymptotic normality of sums of chain dependent variables has been considered by Keilson and Wishart (1964) and O'Brien (1974).

As another example of the usefulness of these results, we can apply Corollary 2.1 to obtain an existence theorem for the Itô-integral. To describe this result we need to introduce the following definitions and notation.

Let \( \{Z(t), t \geq 0\} \) be a martingale which is adapted to the right continuous, increasing family of \( \sigma \)-fields \( \{\mathcal{F}(t), t \geq 0\} \). If \( \pi = \{0 = t_0 < t_1 < \cdots < t_k = 1\} \) is a partition of \([0,1]\) and \( f(t, \omega) \) is a function which is adapted to \( \mathcal{F}(t) \) let

\[
\text{mesh}(\pi) = \sup_{1 \leq j \leq k} (t_j - t_{j-1})
\]

\[
\sum_{f, \pi} = \sum_{j=1}^{k} f(t_{j-1}) Z(t_j) - Z(t_{j-1})
\]

Think of \( \sum_{f, \pi} \) as an approximation to the Itô-integral.

With these conventions recorded we can state our

**Proposition.** Suppose \( f \) is sample function bounded and continuous in probability, that is,

a) \( \sup_{0 \leq t \leq 1} |f(t, \omega)| = M(\omega) < \infty \) a.s.

b) for each \( t, \) if \( t_n \to t \) then \( f(t_n, \omega) \xrightarrow{P} f(t, \omega) \).
Furthermore, suppose \( f(t, \omega) \) is adapted to \( \mathcal{F}(t) \) and that there exists \( K > 0 \) such that for all \( 0 \leq s < t \leq 1 \)

c) \( \mathbb{E}((Z(t) - Z(s))^2 \mid \mathcal{F}(s)) \leq K(t-s) \).

Then if \( \pi_n \) and \( \pi'_n \) are sequences of partitions with \( \text{mesh}(\pi_n) \) and \( \text{mesh}(\pi'_n) \to 0 \) we have

\[
[\sum_{r, \pi_n} - \sum_{r, \pi'_n}] \xrightarrow{P} 0.
\]

**Proof.** Suppose \( \pi_n = \{0 = t_n^0, 0 < t_n^1 \ldots < t_n^k_n = 1\} \) and \( \pi'_n = \{0 = t'_n^0, 0 < t'_n^1 \ldots < t'_n^k'_n = 1\} \) and let \( \sigma_n = \{0 = s_n^0, 0 < s_n^1 \ldots < s_n^k_n = 1\} \) be the partition of \([0,1]\) with points \( \pi_n \cup \pi'_n \). For each \( j \) let

\( r_n, j = \sup\{t_n^i, t_n^i < s_n,j\} \) and \( r'_n, j = \sup\{t'_n^i, t'_n^i < s_n,j\} \). It is easy to see that

\[
\sum_{r, \pi_n} - \sum_{r, \pi'_n} = \sum_{j=1}^{k_n} [f(r_n, j-1) - f(r'_n, j-1)][Z(s_n,j) - Z(s_n,j-1)]
\]

Since \( f(r_n, j-1) \) and \( f(r'_n, j-1) \) are \( \mathcal{F}(s_n,j-1) \) measurable

\[
X_n, j = [f(r_n, j-1) - f(r'_n, j-1)][Z(s_n,j) - Z(s_n,j-1)]
\]

is a martingale difference array.

The remainder of the proof involves checking the condition of Corollary 2.1. Observe that
\[ \sum_{j=1}^{k_n} E(X_{n,j}^2 | \mathcal{F}(s_{n,j-l})) \]

\[ = \sum_{j=1}^{k_n} (f(r_{n,j-l}) - f(r_{n,j-l}')) E([Z(s_{n,j}) - Z(s_{n,j-l})]^2 | \mathcal{F}(s_{n,j-l})) \]

\[ \leq KM(\omega) \sum_{j=1}^{k_n} |f(r_{n,j-l}) - f(r_{n,j-l}')| |s_{n,j} - s_{n,j-l}| \text{ by (a) and (c)} \]

which \( \to 0 \) a.s. as \( n \to \infty \) since by (b) and Fubini, for a.a.\( \omega \) the discontinuities of \( f(\cdot,\omega) \) form a set of Lebesgue measure zero and hence \( f(\cdot,\omega) \) is Riemann integrable. The desired result now follows from Corollary 2.1.

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 \( (W_n, \phi_n) \) converges to \( (W, \varphi) \) where \( W \) and \( \varphi \) are independent. This is of course automatic when \( \phi_n \) converges to a constant. Our next lemma shows that \( W \) and \( \varphi \) are independent whenever \( \phi_n \) converges in probability. To prove this we have to introduce a notion of mixing.

Suppose \( \{V_n, n \geq 0\} \) are random elements of a metric space \( S \) and defined on \( (\Omega, \mathcal{F}, P) \). 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 \( \mathcal{B} \in \mathcal{F} \) with \( P(\mathcal{B}) > 0 \) we have \( (V_n | N) \Rightarrow V \).

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 $U_n \overset{P}{\rightarrow} U$ we have $(V_n, U_n) \Rightarrow (V, U)$ where $V$ and $U$ are independent.

**Theorem 2.4.** Suppose the array $(X_n, \mathcal{F}_n, i)$ has the property that for each $i \geq 0$, $\mathcal{F}_{n,i}$ increases as $n$ increases. For any $\varepsilon_n > 0$ the sequence $W_n$ is R-mixing as a sequence of random elements of $D[0, \infty)$.

**Remark.** The hypothesis is satisfied if for example $(S_i, \mathcal{F}_i)$ is a martingale $\mathcal{F}_{n,i} = \mathcal{F}_i$ 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 (1974, p. 628) under the assumption that $\mathcal{F}_{n,i}$ decreases as $n$ increases. To see that his claim is false let $W$ be a Brownian motion, let $X_n, i = W(12^{-n}) - W((i-1)2^{-n})$ and $\mathcal{F}_{n,i}$ be the $\sigma$-field generated by $(W(s), 0 \leq s \leq 12^{-n})$. The proof given below is essentially due to Rootzen (1974).

**Proof.** It suffices to check that the defining property holds for all $B \in \bigcup_{m,j} \mathcal{F}_{m,j}$ with $P(B) > 0$.

Since $W_n \Rightarrow W$ and $P(B) > 0$ it is easy to see that the conditional sequence $(W_n | B)$ is tight so to complete the proof it remains to show that the finite dimensional distributions converge.
To do this we begin by considering the convergence of 
\( (W_n(1)|B) \). From the proof of Theorem 2.1 given in Freedman (1971) there are stopping times \( \tau_1^n, \tau_2^n, \ldots \) so that (1) for each positive integer \( k \)
\[
(\tilde{S}_{n,1}, \ldots, \tilde{S}_{n,k}) \overset{d}{=} (W(\tau_1^n), \ldots, W(\tau_k^n)),
\]
and \( W(t + \tau_k^n) - W(\tau_k^n) \) is independent of \( \mathcal{F}_{n,k} \), and (2) for each \( t \geq 0 \)
\[
\tau_n(\cdot) \overset{P}{\longrightarrow} t.
\]
We have
\[
E(\exp(i\theta W_n(1))|B)
\]
\[
= E(\exp(i\theta W(\tau_n(1)))|B)
\]
\[
= E(\exp(i\theta[W(1+\tau_n^{\tau_n(1)}) - W(\tau_n^{\tau_n(1)})])|B)
\]
\[
+ E(\exp(i\theta W(\tau_n^{\tau_n(1)}))) - \exp(i\theta[W(1+\tau_n^{\tau_n(1)}) - W(\tau_n^{\tau_n(1)})])|B).
\]

Since \( B \subseteq \mathcal{F}_{n,j} \subseteq \mathcal{F}_{n,j} \) the first conditional expectation is \( \exp(-\theta^2/2) \) by property (1) above. From (2) and the fact that \( |\tilde{X}_{n,1}| \leq \epsilon_n \) we get that \( (\tau_n(\cdot)|B) \overset{P}{\longrightarrow} 1 \) and \( (\tau_j|B) \overset{P}{\longrightarrow} 0 \) so the second conditional expectation goes to 0 by bounded convergence.

To complete the proof of convergence of finite dimensional distributions it suffices to show that if \( a_1, a_2, \ldots, a_k \) are real numbers and \( 0 = s_0 < s_1, \ldots, < s_k \) then

17
converges to a normal distribution with mean zero and variance
\[ \sum_{i=1}^{k} a_i^2(s_i - s_{i-1}) \] (this is the so-called Cramer-Wold device, see Billingsley, 1968, p. 49). However since \[ \sum_{i=1}^{k} a_i(W_n(s_i) - W_n(s_{i-1})) \]
is the sum of variables \( X'_{n,i} \) in a martingale difference array \( \{X'_{n,i}, F_{n,i}\} \) which satisfies the hypotheses of the theorem, this result is a consequence of the first part of the proof.

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

**Theorem 2.5.** Suppose \( \{X_n,i, F_{n,i}\} \) is a martingale difference array and the fields \( F_{n,i} \) increase as \( n \) increases. If

(a) for all \( t > 0 \)
\[ \sum_{i=1}^{k_n(t)} E(X_{n,i}^2 | F_{n,i-1}) \overset{P}{\to} \varphi(t) \quad \text{with} \quad P(\varphi \text{ is continuous}) = 1 \]
and

(b) for all \( \epsilon > 0 \)
\[ \sum_{i=1}^{k_n(1)} E(X_{n,i}^2 1(|X_{n,i}| > \epsilon) | F_{n,i-1}) \overset{P}{\to} 0 \]

Then \( Y_n = S_n k_n(\cdot) \Rightarrow W \circ \varphi \) where \( W \) and \( \varphi \) are independent.
Proof. From the proof of Theorem 2.3 we have that $P\{ \max_{0 \leq s \leq 1} |Y_n(s)| > 0 \} \to 0$, $A_n \xrightarrow{P} 0$, and $\varphi_n \xrightarrow{P} \varphi$. 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 $\varphi$ are independent.

Recently several authors have proved martingale central limit theorems under conditions on the sum of the squares of the variables. In the next few paragraphs we will indicate how the techniques we used with assumption on the conditional variances can be extended to derive limit theorems under the new hypotheses.

The first step is easy. From formula (3.15) of McLeish (1974) it follows that if $(X_{n,i}, \mathcal{F}_{n,i})$ is a martingale difference array with $|X_{n,i}| \leq \varepsilon_n$ and $J_n(\cdot)$ is the time scale defined in Theorem 2.1 then

$$
(2.5) \quad \max_{0 \leq s \leq 1} \left| \sum_{i=1}^{J_n(s)} (X_{n,i}^2 - E(X_{n,i}^2 | \mathcal{F}_{n,i-1})) \right| \xrightarrow{P} 0
$$

so if we let $U_{n,k} = \sum_{i=1}^{k} X_{n,i}^2$ and $J_n'(t) = \sup\{k : U_{n,k} \leq t\}$ then it follows from Theorem 2.1 that $S_n, J_n'(\cdot) \Rightarrow W$.

The next step is to prove the analogue of Theorem 2.2. To do this we define $\hat{X}_{n,i}, \bar{X}_{n,i}, \tilde{X}_{n,i}$ etc. as before and set $W_n = S_n, J_n'(\cdot)$. Now if we let $\varphi_n'(t)$ be a strictly increasing continuous function so that $k_n(t) = J_n'(\varphi_n'(t))$ then using the proof of Theorem 2.2 we obtain
Theorem 2.6. If for some $\epsilon_n \downarrow 0$, (i) $P\{ \max_{0 \leq s \leq 1} |Y_n(s)| > 0 \} \rightarrow 0$ and (ii) $(W_n, \mathcal{F}_n) \Rightarrow (W, \mathcal{F})$ with $P[\mathcal{F} \text{ is continuous}] = 1$ then $Z_n \Rightarrow W \circ \varphi$. If, in addition (iii) $A_n \overset{P}{\rightarrow} 0$ then $Y_n \Rightarrow W \circ \varphi$.

We can use this result to prove an analogue of Theorem 2.3. Since the method of proof is similar to that of Theorem 2.3 and the necessary computations have been done by Rootzén (see the proof of his theorem 4 in [30]), we will omit the proof.

Theorem 2.7. If (a) for all $t > 0$, $\frac{\sum_{i=1}^{n} X_{n,i}^2}{k_n(t)} \overset{P}{\rightarrow} ct$ and (b) for some $\lambda > 0$, $\sum_{i=1}^{n} E(X_{n,i}^2 I_{|X_{n,i}| < \lambda} | \mathcal{F}_{n,i-1}) \overset{P}{\rightarrow} 0$ then $Z_n \Rightarrow W(c.)$. If, in addition, (c) for some $\lambda > 0$:

$$\sum_{i=1}^{k_n(t)} E(X_{n,i}^2 I_{|X_{n,i}| < \lambda} | \mathcal{F}_{n,i-1}) \overset{P}{\rightarrow} 0$$

then $Y_n \Rightarrow W(c.)$.

Remark. It is easy to show using the computations of Theorem 2.3 that if (b) holds for some $\lambda$ then it holds for all $\lambda > 0$. From this observation and Jensen's inequality we see that if $\{X_{n,i}, \mathcal{F}_{n,i}\}$ is a martingale difference array condition (b) is weaker than the corresponding hypothesis of Theorem 2.3.

Theorem 2.7 is a slight generalization of Theorem 4 of Rootzén (1975a) which shows that if $k_n(t) = j'_n(t)$ and (c) holds then (b) and $\max_{1 \leq j \leq k_n(1)} |X_{n,j}| \overset{P}{\rightarrow} 0$ are necessary and sufficient for $Y_n \Rightarrow W$. 

20
Theorem 2.7 contains as a special case Theorems 3.2 and 3.6 of McLeish (1974). In the first he assumes in addition to (a) that

$$\max_{1 \leq i \leq k_n(1)} |X_{n,i}| \text{ converges to 0 in } L_2.$$  

As he notes on p. 625, if we let \( X_{n,i} = \tilde{X}_{n,i} 1_{\{ |X_{n,i}| < C \}} \) then the \( L_2 \) convergence implies

$$\sum_{i=1}^{k_n(1)} \sum_{j=1}^{\infty} \mathbb{E} |\tilde{X}_{n,i}|^2 (X_{n,i} > \epsilon) \to 0 \quad \text{for all } \epsilon > 0 \quad \text{which in turn implies (b)}.$$

In Theorem 3.6 he assumes that (a) holds and there is an array of constants \( c_{n,i} \) bounded away from 0 and \( \infty \) so that

$$\sum_{i=1}^{k_n(1)} \mathbb{E} |X_{n,i}| 1_{\{ |X_{n,i}| < c_{n,i} \}} |\mathcal{F}_{n,i-1}| \to 0$$

In view of the computations in the proof of Theorem 2.2 there is no loss of generality in supposing \( c_{n,i} = \lambda \) so his condition is stronger than (b) and implies (c).

It is easy to use Theorem 2.7 to obtain a result about convergence to mixtures of Brownian motions under the assumption that \( \mathcal{F}_{n,i} \) increases as \( n \) increases. The key observation is that from Theorem 2.4 and formula (2.5), \( W_n \) is R-mixing. The theorem we obtain in this way has been given by Rootzén ([31], Theorem 1) and generalizes a theorem of Eagleson ([9], Theorem 2).
3. Weak Convergence to a Poisson Process

Suppose \( \{X_n, \mathcal{F}_n, \} \) 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 \( (R^2, \mathcal{B}(R^2)) \) which has counting function

\[
N_n((0,t] \times [a,b]) = \sum_{i=1}^{k_n(t)} 1_{\{a \leq X_n,i \leq b\}}
\]

(where as usual 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 on \( (R^2, \mathcal{B}(R^2)) \). Necessary background on the weak convergence of point processes may be found in Jagers (1974).

At the end of this section, our result is applied to obtain weak convergence of maxima of dependent variables to limit extremal processes. In the next section our Poisson convergence result is applied to derive criteria for weak convergence of sums of dependent variables to Lévy processes.

Theorem 3.1. Let \( \nu \) be a measure on \( (R, \mathcal{B}(R)) \) with the property that if \( x_0 = \inf(x | v(x, \infty) < \infty) \), then \( -\infty \leq x_0 < \infty \). Suppose for all \( t > 0 \) and for \( x > x_0 \) such that \( v([x]) = 0 \) we have as \( n \to \infty \)

\[
\sum_{j=1}^{k_n(t)} P[X_n,j > x | \mathcal{F}_n, j-1] \xrightarrow{P} t \nu(x, \infty)
\]
\[ (3.2) \quad \max_{j \leq k_n(t)} P[X_n, j > x | \mathcal{F}_{n, j-1}] \xrightarrow{P} 0. \]

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

**Remark.** We can replace (3.2) by the equivalent condition

\[ (3.2') \quad \sum_{j=1}^{k_n(t)} (P[X_n, j > x | \mathcal{F}_{n, j-1}])^2 \xrightarrow{P} 0. \]

**Proof.** Let \( \mathcal{G} \) be the class of rectangles in \( \mathbb{R}^2 \) of the form 
\((a, b] \times (c, d)\) or \((a, b] \times (c, \infty)\) where \( 0 \leq a < b, \nu([c]) = \nu([d]) = 0 \) and \( x_0 < c < d \). Suppose \( A \) is a disjoint union of rectangles in \( \mathcal{G} \) say, \( A = \sum_{i=1}^{m} R_i \) where \( R_i = (a_{i,1}, b_{i,1}] \times (c_{i,1}, d_{i,1}] \). We first show that in \( \mathbb{R}^1 \)

\[ (3.3) \quad N_n(A) \Rightarrow N(A) \]

where \( N(A) \) is Poisson distributed with mean \( \sum_{i=1}^{m} (b_{i,1} - a_{i,1}) \nu(c_{i,1}, d_{i,1}] \).

If \( \{(a_{i,1}, b_{i,1}], i \leq m\} \) are not pairwise disjoint. Then \( A \) can be rewritten \( A = \sum_{i} (\alpha_{i,1}, \beta_{i,1}) \times I_{i,1} \) where now the time intervals \( (\alpha_{i,1}, \beta_{i,1}] \) are pairwise disjoint and \( I_{i,1} \) is a finite union of disjoint intervals. Then
\begin{equation}
N_n(A) = \sum_i \sum_{j=k_n(\alpha_i)+1}^\infty 1\{X_{nj} \in I_i\}.
\end{equation}

Note that by (3.1) and (3.2)

\begin{equation}
\sum_i \sum_{j=k_n(\alpha_i)+1}^\infty \mathbb{P}(X_{nj} \in I_i | \mathcal{F}_n, j-1) \xrightarrow{P} \sum_i (\beta_i - \alpha_i) \nu(I_i),
\end{equation}

and

\begin{equation}
\sum_i \sum_{j=k_n(\alpha_i)+1}^\infty \mathbb{P}^2(X_{nj} \in I_i | \mathcal{F}_n, j-1) \xrightarrow{P} 0.
\end{equation}

The desired result now follows from Freedman (1974), Theorem 5.

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

**Corollary 3.1.** Suppose \( \nu \) satisfies the condition of (3.1) and also \( \nu(R) = +\infty \). If \( M_n(t) = \bigvee_{j \leq k_n(t)} X_{nj} \) and \( M_\infty \) 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)

\[
M_n \Rightarrow M \quad \text{in} \quad D(0, \infty)
\]

**Proof.** The result follows immediately from the continuous mapping theorem (Billingsley, 1968, p. 30) upon applying the \( D(0, \infty) \)-valued functional \( g(N_n)(t) := \bigvee_{j \leq k_n(t)} X_{n,j} \). That \( g(N) \overset{d}{\Rightarrow} Y \) is well known. (Cf. Resnick and Rubinovitch (1973), Resnick (1975), Weissman (1976).)

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

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

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

\[
\sum_{j=1}^{[nt]} \mathbb{P}(X_j > a_n x + b_n | J_{n-1}) = \sum_{j=1}^{[nt]} (1 - H_j a_n x + b_n)
\]

\[
= \sum_{i=1}^{m} \pi_i ([nt] - 1)(1 - H_i (a_n x + b_n))
\]
and since $\pi_i(n) \sim \pi_i n$ as $n \to \infty$ for $i = 1, \ldots, m$ we have that (3.1) holds iff

\[(3.4) \quad n \sum_{i=1}^{m} \frac{\pi_i}{\pi_i} (1 - H_i(a_n x + b_n)) \to \nu(x, \infty).\]

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 $dt \times dv$ where $\nu(x, \infty) = -\log F(x)$ and $F$ is one of the three classes of extreme value distributions. Also $M_n := (\sum_{j=1}^{[n]} X_j - b_n)/a_n \Rightarrow \Delta_\infty$ in $D(0, \infty)$ where $\Delta_\infty$ is an extremal-$F$ process.

Asymptotic independence of $\{J_n\}$ and $\{Y_n\}$ can be obtained as in Example 2.1.

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

Example 3.2. Random exchanges: Helland and Nilsen (1975) consider the following model of deep water exchanges in a sill fjord: Let $U = \{U_n, n \geq 1\}$, $D = \{D_n, n \geq 1\}$ be two independent sequences of iid variables with
\[ P(U_1 \leq x) = G(x), \quad P(D_1 \leq x) = H(x). \]

The density of the water in the fjord basin at time \( n \) is \( \xi_n \) where \( \xi_0 \) is independent of \( U, D \) and for \( n \geq 1 \)

\[(3.5) \quad \xi_n = (\xi_{n-1} - D_n) \lor U_n.\]

Our methods suffice to prove the following result of Helland and Nilsen: If \( 1 - G \) is regularly varying with exponent \(-\alpha\), \( 0 < \alpha < 1 \) then setting \( a_n = G^{-1}(1 - n^{-\alpha}) \) gives

\[
\lim_{n \to \infty} P[\xi_n \leq a_n x] = \begin{cases} 
0 & x < 0 \\
\exp(-x^{-\alpha}) & x \geq 0
\end{cases}
\]

provided \( E|D_1| < \infty \).

For the proof, first iterate (3.5) to obtain

\[ \xi_n = \max\{\xi_0 - \sum_{1}^{n} D_i, U_1 - \sum_{2}^{n} D_i, U_2 - \sum_{3}^{n} D_i, \ldots, U_{n-1} - D_n, U_n\}. \]

Clearly it suffices to consider the variables

\[ \xi'_n = \max\{U_1, U_2 - D_1, U_3 - \sum_{1}^{2} D_i, \ldots, U_n - \sum_{1}^{n-1} D_i\} \]

and show that \( \xi'_n / a_n \) has the indicated limit. Equation (3.1) at \( t = 1 \) becomes

27
\begin{align*}
(3.6) \quad \sum_{k=1}^{n} (U_k - \sum_{l=1}^{k-1} D_l) &> a_n x [U_1, \ldots, U_{k-1}, D_{l'}, \ldots, D_{k-1}] \\
&= \sum_{k=1}^{n} (1 - G(a_n x + \sum_{l=1}^{k-1} D_l)) \\
&\leq \sum_{k=1}^{n} \left(1 - G \left(a_n \left(x - \frac{\sum_{l=1}^{n} |D_l|}{a_n}\right)\right)\right) \\
&= n \left(1 - G \left(a_n \left(x - \frac{\sum_{l=1}^{n} |D_l|}{a_n}\right)\right)\right)
\end{align*}

Since $1 - G$ is regularly varying, $n(1 - G(a_n x)) \to x^{-\alpha}$ locally uniformly in $x$. Furthermore it is well known that $a_n$ is regularly varying with exponent $1/\alpha$ so that $n/a_n \to 0$ and therefore by the strong law of large numbers $\sum_{l=1}^{n} |D_l|/a_n \to 0$ a.s. Therefore $\limsup_{n \to \infty}$ of (3.6) is $\leq x^{-\alpha}$. A lower bound can be constructed similarly so that the limit of (3.6) is $x^{-\alpha}$. Relation (3.2) can be checked analogously and hence the desired result follows by Corollary 3.1.

**Example 3.3.** Let \{\(E_k, \kappa < k < \infty\)\} be iid exponentially distributed variables with $P[E_k > x] = e^{-x}$, $x > 0$ and set $c_k = E_k - 1$ and define $X_n = \sum_{k=0}^{\infty} \rho^k c_{n-k}$ where $0 < \rho < 1$ (so that $X_n = \rho X_{n-1} + c_n$). Set $X_n, j = X_j - \log n$, $n \geq 1$, $j \geq 1$ and $k_n(t) = [nt]$. Evaluating the left side of (3.1) we obtain
\[
\sum_{k=1}^{[nt]} e^{-x \cdot \log n} e^{\rho X_{k-1}} = e^{-x} \sum_{k=1}^{[nt]} e^{\rho X_{k-1}/n}.
\]
Since \( \{e^{-kX_t}\} \) is stationary, we obtain by the ergodic theorem that the above converges a.s. to

\[
\lim_{t \to \infty} e^{-X_t} = e^{-X_0}
\]

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

\[
E e^{X_0} = \exp\{-\rho(1-\rho)^{-1}\} \prod_{l=1}^{\infty} (1 - \rho^k).
\]

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

\[
P\left( \max_{k \leq nt} P[X_{nj} > x | \mathcal{F}_{n, j-1}] > \delta \right)
\]

\[
\leq \frac{[nt]}{\sum_{k=1}^{nt} \left( e^{-\frac{1}{n}} e^{-\rho^k} \right)} P\left( e^{\rho X_0} > \delta \right).
\]

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

\[
\frac{\xi e^{X_0}}{(nc)^{-\zeta}} \to 0, \quad n \to \infty.
\]

This verifies (3.2) and hence we conclude

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

where \( M_{\infty} \) is the extremal process generated by the distribution \( \exp\{-e^{X_0}e^{-x}\} \).

More sophisticated examples should be amenable to the analysis of example 3.3.
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 $\{Z(t), t \geq 0\}$ be a process with stationary independent increments. The characteristic function of $Z(t)$ is given by

$$E(e^{i\theta Z(t)}) = \exp\{t(ia\theta + c\theta^2/2) + \int_{|x| \geq \gamma} (e^{i\theta x} - 1) \nu(dx)$$

$$+ \int_{0 < |x| < \gamma} (e^{i\theta x} - 1 - i\theta x) \nu(dx)\}$$

where $a$ is a constant, $c$ is a nonnegative number and $\nu$ (called the Lévy measure is a $\sigma$-finite measure on $\mathbb{R}_0 = (-\infty, 0) \times (0, \infty)$ with the property that $\int (x^2 \wedge 1) \nu(dx) < \infty$ (cf. Gnedenko and Kolmogorov, 1968, p. 84).

Let $N$ be a Poisson random measure on $[0, \infty) \times \mathbb{R}_0$ with mean measure $dt \times \nu$ and points $\{(t, \xi)\}$. Let $W$ be a normalized Brown motion independent of $N$. The Itô representation of $Z$ (cf. Itô, 1969, p. 1.7.7) is
(4.2) \[ Z(t) = a + W(\eta t) + \sum_{t_k < t} \xi_k \mathbb{1}\{ |\xi_k| \geq \gamma \} \]

\[ + \lim_{t \to \infty} \sum_{t_k < t} \xi_k \mathbb{1}\{ |\xi_k| \in (\delta, \gamma) \} - t \int_0^t \mathbb{1}\{ |s| \in (\delta, \gamma) \} v(ds) \]

where for almost all \( \omega \), 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 Weiner component). The notation used in the statements below is that of section 2 except that we have added a superscript \( \delta \) to indicate the value at which the sequence is truncated (in section 2 this value was \( \epsilon_n \)). Throughout this section, we suppose \( k_n(t) \) is a stopping time for each \( t \).

**Theorem 4.1.** Let \( \nu \) 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 \( t > 0 \)

\[ k_n(t) \sum_{i=1} P[X, i > x | F_n, i-1] \xrightarrow{P} t \nu(\infty) \]

and

\[ k_n(t) \sum_{i=1} P[X, i < y | F_n, i-1] \xrightarrow{P} t \nu(-\infty, y) \]

whenever \( x > 0, y < 0 \) and \( \nu(\{x\}) = \nu(\{y\}) = 0 \),
(b) for all $\epsilon > 0$

$$\max_{1 \leq j \leq k_n(1)} P[|X_{n,i}| > \epsilon |F_{n,i-1}] \xrightarrow{P} 0$$

and

(c) for all $\epsilon > 0$

$$\lim_{n \to \infty} \limsup_{S \ni 0} P\left( \sup_{0 \leq s \leq 1} |\tilde{I}^S_n(s) - A^S_n(s)| > \epsilon \right) = 0$$

then $Z_n \Rightarrow Z$.

A sufficient condition for (c) is

(d) for all $\epsilon > 0$

$$\lim_{n \to \infty} \limsup_{S \ni 0} P\left( \sum_{i=1}^{k_n(1)} \mathbb{E}\left( (|\tilde{I}^S_{n,i}|^2 |F_{n,i-1}) \right) > \epsilon \right) = 0$$

**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

$$\sum_{i=1}^{k_n(t)} \mathbb{E}(X_{n,i} 1_{\delta \leq |X_{n,i}| \leq \gamma} |F_{n,i-1}) \xrightarrow{P} t \int_{|x| \in (\delta, \gamma)} sv(ds)$$

for all $t \geq 0$ whenever $\nu((-\delta, \delta)) = 0$. To do this recall that the
set of Borel Radon measures on \([0,\infty) \times \mathbb{R}^0\) is a complete separable metric space. From (a) and Theorem 3, p. 206 of Jagers (1974) it follows that the random measures

\[\mu_n([0,t] \times A) = \sum_{i=1}^{k_n(t)} \mathbb{P}(X_{n,i} \in A | \mathcal{F}_{n,i-1})\]

defined on \([0,\infty) \times \mathbb{R}^0\) converge weakly to the measure \(dt \times dv\).

The map

\[\mu \longrightarrow \int_{s \in (\delta, \gamma)} s_{\mu}([0,t] \times ds)\]

is continuous at each \(\mu\) with \(\mu([0,1] \times [-\delta, \delta, -\gamma, \gamma]) = 0\) so by the continuous mapping theorem (Billingsley, 1968, p.30)

\[\sum_{i=1}^{k_n(t)} E(X_{n,i} 1\{\delta \leq |X_{n,i}| < \gamma | \mathcal{F}_{n,i-1})\]

\[= \int_{s \in (\delta, \gamma)} s_{\mu_n}([0,t] \times ds) \Rightarrow t \int_{x \in (\delta, \gamma)} s_{\nu}(ds)\]

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 \([0,\infty) \times \mathbb{R}^0\) described before Theorem 4.1. If \(\nu([\delta, -\delta]) = 0\) then we can apply the continuous functional which sums ordinates of points in \([(s,x) : s \leq t, |x| > \delta]\) and conclude from the continuous mapping theorem that
\[ k_n(\cdot) \]
\[ \sum_{i=1}^{X_n,i} 1\{ |X_n,i| \geq \delta \} \Rightarrow \sum_{t_k \leq \cdot} \xi_k 1\{ |\xi_k| \geq \delta \} \]

Combining this result with (4.3) gives

\[ Z_{\delta,n} = \sum_{i=1}^{X_n,i} 1\{ |X_n,i| \geq \delta \} - \sum_{i=1}^{X_n,i} E(X_n,i) 1\{ \delta < |X_n,i| < \gamma \} |_{F_{n,i-1}} \]

\[ \Rightarrow Z_{\delta} = \sum_{t_k \leq \cdot} \xi_k 1\{ |\xi_k| \geq \delta \} - (\cdot) \int_{\delta < |\delta| < \gamma} s_{\nu}(ds). \]

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

\[ (\text{i.4}) \quad \lim_{\delta \to 0} \lim_{n \to \infty} \sup \{ \rho(Z_{\delta,n}, Z_n) > \epsilon \} = 0 \]

where \( \rho \) is the Skorohod metric for \( D[0,1] \). However \( Z_n - Z_{\delta,n} = Y_n - A_{\delta,n} \) so (c) implies (4.4) and the proof of the first statement is complete.

To prove that (d) is sufficient for (c), observe that if (d) holds and (c) does not then for some \( \epsilon > 0 \) there is a sequence \( \delta_n \downarrow 0 \) so that...
\begin{equation}
\limsup_{n \to \infty} P\left( \sup_{0 \leq s \leq 1} |\hat{X}_n^\infty(s) - A_n^\infty(s)| > \epsilon \right)
\end{equation}

\begin{equation}
= \limsup_{\delta \to 0} \limsup_{n \to \infty} P\left( \sup_{0 \leq s \leq 1} |Y_n^\delta(s) - A_n^\delta(s)| > \epsilon \right) > 0
\end{equation}

and for all \( \eta > 0 \)

\begin{equation}
k_n(1) \quad P\left( \sum_{i=1}^{k_n(1)} E\left( (\hat{X}_{n,i}^\infty)^2 \bigg| \mathcal{F}_{n,i-1} \right) > \eta \right) \to 0.
\end{equation}

The latter convergence implies by Corollary 2.1 that \( \hat{X}_n^\infty - A_n^\infty \to 0 \) which contradicts (4.5).

**Example 4.1.** Chain dependent variables. Continuing the developments of examples 2.1 and 3.1 let \( H = \sum_{m} \pi_i H_i \) and suppose

\begin{equation}
1 - H(x) + H(-x) \sim x^{-\alpha} L(x), \quad 0 < \alpha < 2
\end{equation}

\begin{align}
\frac{1 - H(x)}{1 - H(x) + H(-x)} &\to p, \\
\frac{H(-x)}{1 - H(x) + H(-x)} &\to q
\end{align}

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{align}
v(x, \infty) &= px^{-\alpha} \quad \text{for } x > 0 \\
v(-\infty, y) &= qy^{-\alpha} \quad \text{for } y > 0,
\end{align}

where

\begin{equation}
x_{n,j} = x_j/a_n \quad \text{and } a_n \text{ is chosen so that}
\end{equation}
\[ n(1 - H(a_n x) + H(-a_n x)) \to x^{-\alpha}, \quad x > 0, \]

we check condition (d) as follows:

\[
\sum_{i=1}^{n} \frac{E((x_{n,i}^5)^2 | \mathcal{F}_{n,i-1})}{\mathcal{F}_{n,i-1}} \leq \sum_{i=1}^{n} \frac{E(x_{n,i}^2 1_{\{|X_{n,i}^5| \leq \delta\}} | \mathcal{F}_{n,i-1})}{\mathcal{F}_{n,i-1}}
\]

\[
= \sum_{i=1}^{m} \int_{\{x | x \leq \delta\}} x^2 \pi_i (x) | \mathcal{F}_{n,i-1}) \sim n \int_{\{x | x \leq a_n \delta\}} x^2 d(S \pi_i H_i(x))
\]

\[
= \frac{n}{a_n^2} \int_{\{x | x \leq a_n \delta\}} x^2 dH(x)
\]

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

\[
\frac{n}{a_n^2} \left( \frac{\alpha}{2-\alpha} \right) (\delta a_n) ^2 (1 - H(a_n \delta) + H(-a_n \delta))
\]

\[
= \frac{\alpha \delta^2}{2-\alpha} n(1-H(a_n \delta) + H(-a_n \delta)) \to \frac{\alpha \delta^2}{2-\alpha} \quad \text{as } n \to \infty
\]

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

Thus from Theorem 4.1 we have

\[
\sum_{i=1}^{\lfloor n \rfloor} \frac{X_i}{a_n} - \sum_{i=1}^{\lfloor n \rfloor} \frac{E(X_i 1_{\{|X_i/a_n| < \gamma\}} | \mathcal{F}_{n,j-1})}{\mathcal{F}_{n,j-1}} \to Z_{\alpha}^\prime
\]

where \(Z_{\alpha}^\prime\) is a stable process with characteristic function given by (4.1) with \(a = 0 = c\) and \(\gamma\) specified as in the first part of the example.

When \(\alpha < 1\) we observe by formula (5.22) of p. 579 of Feller (1971) that

36
\[
\sum_{1}^{[nt]} E \left( \frac{X_j}{a_n} 1\{|X_j/a_n|<\gamma\} | F_{n,j-1} \right) \to \frac{\alpha}{1-\alpha} (p-q) \gamma^{1-\alpha} t \\
\text{a.s. and locally uniformly in } t.
\]

When \( \alpha > 1 \), we use formula (5.21) on p. 579 and (5.16) on p. 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

\[
\sum_{1}^{[nt]} E \left( \frac{X_j}{a_n} 1\{|X_j/a_n|<\gamma\} | F_{n,j-1} \right) - \sum_{1}^{m} \nu_i \int x \, dH_i
\]

\[
\to \left( \frac{\alpha}{\alpha-1} \right) (p-q) \gamma^{1-\alpha} t
\]

a.s. and locally uniformly in \( t \).

To sum up:

\[
\sum_{1}^{[n.]} \frac{X_j}{a_n} - \frac{tn}{a_n} (p-q) \gamma^{1-\alpha} \Rightarrow Z_\alpha, \quad 0 < \alpha < 1
\]

\[
\sum_{1}^{[n.]} \frac{X_j}{a_n} - nt \sum_{1}^{m} \nu_i \int x \, dH_i + \left( \frac{\alpha}{\alpha-1} \right) (p-q) \gamma^{1-\alpha} t \Rightarrow Z_\alpha, \quad 1 < \alpha < 2
\]

\[
\sum_{1}^{[n.]} \frac{X_j}{a_n} - \frac{tn}{a_n} \int_{x < \gamma a_n} x \, dH \Rightarrow Z_\alpha, \quad \alpha = 1.
\]

Wolfson (1974) has shown that partial sums of chain dependent variables can converge only to stable laws.
Theorem 4.2. Let $v$ and $Z$ be as specified in (4.1) and (4.2) with $a = 0$. Suppose that $v([-\gamma, \gamma]) = 0$. If conditions (a) and (b) of Theorem 4.1 are satisfied and (c) for all $\epsilon$ and $t > 0$

$$\lim_{n \to \infty} \limsup_{n \to \infty} P \left\{ \left| \sum_{i=1}^{\infty} E(x_{n,i})^2 \left| F_{n,i-1} \right| \right| > \epsilon \right\} = 0$$

then $Z_n \Rightarrow Z$.

If (d) for all $\epsilon > 0$

$$\lim_{n \to \infty} \limsup_{n \to \infty} P \left\{ \sum_{i=1}^{\infty} E(x_{n,i}^2 \left| F_{n,i-1} \right|^2 > \epsilon \right\} = 0$$

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

$$\lim_{n \to \infty} \sum_{i=1}^{\infty} E(x_{n,i}^2) \mathbb{1}_{\{x < |X_{n,i}| < y\}} \left| F_{n,i-1} \right| \overset{P}{\to} t[G(y) - G(x)]$$

Remark. Conditions (a)-(c) are the analogues of those given by Gnedenko and Kolmogorov (1968) in the case of independence (see p. 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

$$
\sum_{i=1}^{k_n(t)} \mathbb{E}(\hat{X}_{n,i}^2 | \mathcal{F}_{n,i-1}) \xrightarrow{P} ct
$$

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

$$
Z_n = \hat{Y}_n - B^Y_n + W_n \circ \varphi_n.
$$

By (4.6) we have $\varphi_n(t) \xrightarrow{P} ct$ so that $W_n \circ \varphi_n \Rightarrow W(c, )$. To complete the proof it suffices to show that

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

where $W$ and $\hat{Z}$ are independent. If $\Lambda \subset [0,1]$ is a $W$-continuity set and $P[W \in \Lambda] > 0$ then all we need show is $(\hat{Y}_n - B_n | W_n \in \Lambda)$ $\Rightarrow \hat{Z}$. Checking that the conditions of Theorem 4.1 hold with $P$ replaced by $P(\cdot | W_n \in \Lambda)$ we get the desired result. (For example, we can check (a) of Theorem 4.1 by noting that $\mu_n \xrightarrow{P} dt \times dv$ and therefore $(\mu_n | W_n \in \Lambda)$ converges in probability to $dt \times dv$.)
References


