Z. Wahrscheinlichkeitstheorie verw. Gebiete
40, 59–82 (1977)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete


# Limit Theorems for Subsequences of Arbitrarily-Dependent Sequences of Random Variables

D.J. Aldous*

University of Cambridge, Statistical Laboratory, 16 Mill Lane, Cambridge, England

## 1. Introduction

Komlós [11] proved the following. Suppose $\{X_i\}$ is a sequence of r.v.'s such that $\sup E|X_i| < \infty$: then there exists a r.v. $\alpha$ and an increasing sequence of integers $\{n_i\}$ such that

$$N^{-1}\sum_{1}^{N} X_{n_i} \to \alpha \text{ a.s.}$$

Chatterji [3] has formulated the following heuristic principle. Given a limit theorem for i.i.d.r.v.'s under certain moment conditions, there exists an analogous theorem such that an arbitrarily-dependent sequence (under the same moment conditions) always contains a subsequence satisfying this analogous theorem.

The result of Komlós may be regarded as the precise form of Chatterji's principle for the strong law of large numbers. The principle has been verified for several other classical theorems ([1, 3, 4, 8]) by purely *ad hoc* arguments. Theorem 3 of this paper implies the truth of the principle for *any* a.s. limit theorem, and Theorem 6 for *any* weak limit (convergence in distribution) theorem; subject only to mild technical conditions on the nature of the theorem.

To describe the technique used, define an exchangeable sequence of r.v.'s to be a mixture of i.i.d. sequences. Obviously exchangeable sequences satisfy theorems analogous to those for i.i.d. sequences. Given an arbitrarily-dependent sequence $\{X_i\}$, assuming only that the distributions are tight, we will (Proposition 11) extract a certain subsequence $\{Y_i\} = \{X_{n_i}\}$, and associate with it an exchangeable sequence $\{Z_i\}$. We then show (Proposition 13) that certain types of property of $\{Z_i\}$ are shared by suitable subsequences of $\{Y_i\}$.

This technique also proves (Theorem 9) an assertion of Révész concerning unconditionally a.s. convergent subsequences. These results will be stated properly in Section 3.

* This research was supported by a S.R.C. Studentship

## 2. Notation and Definitions

Let $S$ be a separable metrisable space. Topological spaces will always be equipped with the $\sigma$-field generated by the open sets, and product spaces will be given product topology and product $\sigma$-field.

Let $C(S)$ (resp. $L^\infty(S)$) be the set of continuous (resp. measurable) bounded real-valued functions on $S$. Let $\mathscr{P}(S)$ be the space of probability measures on $S$, equipped with the weak topology

$$\lambda_n \Rightarrow \lambda \quad \text{if and only if} \quad \int f d\lambda_n \to \int f d\lambda \qquad \text{for each } f \in C(S).$$

Recall ([13], Theorem 6.2) that $\mathscr{P}(S)$ is itself a separable metrisable space. For measurable $A \subset \mathbb{R}$, the map $\lambda \to \lambda(A)$ from $\mathscr{P}(\mathbb{R})$ to $\mathbb{R}$ is measurable. Conversely [9], if $\mathbb{R}_0$ is a dense subset of $\mathbb{R}$ then the collection of maps

$$\lambda \to \lambda((-\infty, x]); \qquad x \in \mathbb{R}_0 \tag{2.1}$$

generate the $\sigma$-field on $\mathscr{P}(\mathbb{R})$.

Let $(\Omega, \mathscr{B}, P)$ be a probability space. A measurable function $T: \Omega \to S$ will be called a *random measure* when $S = \mathscr{P}(\mathbb{R})$; a *random vector* when $S = \mathbb{R}^\infty$; a *random variable* (r.v) when $S = \mathbb{R}$; and in general a *random map*. Let $\mathscr{L}(T)$ denote the distribution of $T$. Write $T_n \to_{\mathscr{D}} T$ for $\mathscr{L}(T_n) \Rightarrow \mathscr{L}(T)$. Write $\mathscr{F}(T)$ for the $\sigma$-field generated by $T$. Write $I(A)$ for the indicator function of $A$. For $s \in S$, write $\delta_s$ for the measure $\delta_s(A) = I\ (s \in A)$.

Suppose $\mathscr{B}_0$ is a sub-$\sigma$-field of $\mathscr{B}$. A random map $\pi: \Omega \to \mathscr{P}(S)$ is called a regular conditional distribution (r.c.d.) for $T$ given $\mathscr{B}_0$ if, for each measurable $A \subset S$,

$$\pi(\omega, A) = E(I(T \in A) | \mathscr{B}_0)(\omega) \text{ a.s.} \tag{2.2}$$

Here we have written $\pi(\omega, A)$ for $(\pi(\omega))(A)$, though the author prefers to regard $\pi$ as random map rather than as a kernel function. An earlier draft of this paper was full of r.c.d.'s, but these have been largely eliminated for the reasons of [7, page viii]. However, frequent use is made of the following Lemma, which enables versions of conditional expectations to be computed.

**Lemma 1.** *Let $X, Y$ be random maps into spaces $S_1, S_2$. Let $h: S_1 \times S_2 \to \mathbb{R}$ be a measurable function such that $E|h(X, Y)| < \infty$. Let $X$ be $\mathscr{B}_0$-measurable. Suppose $\pi$ is a r.c.d. for $Y$ given $\mathscr{B}_0$. Then*

$$\int_{S_2} h(X(\omega), y)\, \pi(\omega, dy) = E(h(X, Y) | \mathscr{B}_0)(\omega) \text{ a.s.} \tag{2.3}$$

For $\lambda \in \mathscr{P}(\mathbb{R})$, let $\lambda^* \in \mathscr{P}(\mathbb{R}^\infty)$ be the product measure $\lambda \times \lambda \times \cdots$. For a random measure $\mu$, let $\mu^*$ be the random map into $\mathscr{P}(\mathbb{R}^\infty)$ such that $\mu^*(\omega) = \mu(\omega) \times \mu(\omega) \times \cdots$.

It is often convenient to regard a sequence $\{X_i\}$ of random variables as a random vector $\mathbf{X}$. For example, the assertion $\mathscr{L}(\mathbf{X}) = \lambda^*$ is a concise way of saying that $\{X_i\}$ is an independent identically distributed (i.i.d.) sequence of r.v.'s. Because this paper concerns subsequences, it is worth introducing special notation for these. Write $\mathbf{n} = \{n_i\}$ and $\mathbf{m} = \{m_i\}$ to denote strictly increasing sequences of positive integers (constants, not random variables). Write $\mathbf{m} \subset \mathbf{n}$ to mean that $\{m_i\}$ is a subsequence of $\{n_i\}$. Write $\mathbf{X}_n$ for the random vector $\{X_{n_1}, X_{n_2}, \ldots\}$. For typographical convenience, we often write $X_{n,i}$ for $X_{n_i}$.

**Definition.** A random vector $\mathbf{Z} = \{Z_i\}$ is *exchangeable* if there exists a random measure $\mu$ such that

$$\mu^* \text{ is a r.c.d. for } \mathbf{Z} \text{ given } \mathscr{F}(\mu). \tag{2.4}$$

Following Kallenberg [10], call $\mu$ the *canonical* random measure for $\mathbf{Z}$. For the purposes of this paper, "exchangeable" is merely a name for sequences which are mixtures of i.i.d. sequences in the sense of (2.4). However, the reader may recall the usual definition:

$$\mathscr{L}(Z_1, \ldots, Z_j) = \mathscr{L}(Z_{i_1}, \ldots, Z_{i_j}) \tag{2.5}$$

for each $j \in \mathbb{Z}^+$ and each permutation $(i_1, \ldots, i_j)$ of $(1, \ldots, j)$. That (2.4) implies (2.5) is straightforward. The converse, loosely known as de Finetti's theorem, is non-trivial (see [10]) but will not be needed, though Proposition 11 uses an extension of some of the ideas in its proof.

**Definition.** $X_i \to X\, \sigma(L^1, L^\infty)$ means

$$E(X_i \mid B) \to E(X \mid B) \qquad \text{for each } B \in \mathscr{B},\ P(B) > 0. \tag{2.6}$$

It is well known ([12], T 23) that for $\{X_i\}$ to be $\sigma(L^1, L^\infty)$ relatively sequentially compact it is necessary and sufficient that $\{X_i\}$ be uniformly integrable, and in particular it is sufficient that $\{X_i\}$ be $L^p$-bounded for some $p > 1$.

**Definition.** For $\lambda \in \mathscr{P}(\mathbb{R})$, write

$$|\lambda|^p = \int |x|^p \lambda(dx), \qquad 0 < p < \infty. \tag{2.7}$$

Also, write $|\lambda|_1$ for the mean and $|\lambda|_2$ for the variance of $\lambda$. For definiteness, write $|\lambda|_1 = \infty$ if $|\lambda|^1 = \infty$, and write $|\lambda|_2 = \infty$ if $|\lambda|^2 = \infty$.

## 3. Statement and Discussion of Results

Let $\mathbf{X} = \{X_i\}$ be a random vector such that

$$\{\mathscr{L}(X_i)\} \qquad \text{is tight.} \tag{3.1}$$

In Section 4 we shall construct a random measure $\mu$ associated with $\mathbf{X}$. The details of the construction need not concern us now, except for the following technical result to be proved in Section 5.

**Lemma 2.** $E|\mu(\omega)|^p \leq \limsup E|X_i|^p,\ 0 < p < \infty.$

Define $\alpha(\omega)=|\mu(\omega)|_1$, and $\beta(\omega)=|\mu(\omega)|_2$. Then

$$\sup E|X_i|<\infty \quad \text{implies } \alpha(\omega)<\infty \text{ a.s.,} \tag{3.2}$$

$$\sup E|X_i|^2<\infty \quad \text{implies } \beta(\omega)<\infty \text{ a.s.} \tag{3.3}$$

Our results are stated in this Section in terms of $(\mathbf{X},\mu)$. When these results are applied to specific theorems, $\alpha$ and $\beta$ play the role played by the mean and variance in the i.i.d. case.

In our main results the only assumption on $\mathbf{X}$ is (3.1), which will not be mentioned again. In order to motivate the form of these results, however, consider the following special property $\mathfrak{P}$ which $\mathbf{X}$ might possess.

There exists an exchangeable random vector $\mathbf{Z}$ such that

(i) $\mu$ is the canonical random measure for $\mathbf{Z}$;

(ii) $\sum_1^\infty |X_{m,i}-Z_i|<\infty$ a.s., for some $\mathbf{m}$.

Our results will be formulated so as to be almost obvious if $\mathbf{X}$ satisfies $\mathfrak{P}$. An example in Section 7 will show that in general $\mathfrak{P}$ is not satisfied.

It is now necessary to say precisely what we mean by an a.s. limit theorem for i.i.d. random variables. Define a *statute* $A$ to be a measurable subset of $\mathscr{P}(\mathbb{R})\times\mathbb{R}^\infty$ such that, for each $\lambda\in\mathscr{P}(\mathbb{R})$,

$$(\lambda,\mathbf{V}(\omega))\in A \text{ a.s.} \quad \text{when } \mathscr{L}(\mathbf{V})=\lambda^*. \tag{3.4}$$

Equivalently, (3.4) can be written as

$$\lambda^*\{\mathbf{x}\in\mathbb{R}^\infty\colon (\lambda,\mathbf{x})\in A\}=1, \tag{3.5}$$

or, writing (3.4) out fully, as

$$P((\lambda, V_1(\omega), V_2(\omega), \ldots)\in A)=1 \tag{3.6}$$

when $\{V_i\}$ is an i.i.d. sequence with distribution $\lambda$.

Let us give examples of the statutes representing some well-known theorems.

$$A_1=\left\{(\lambda,\mathbf{x})\colon \lim_{N\to\infty} N^{-1}\sum_1^N x_i=|\lambda|_1\right\}\cup\{(\lambda,\mathbf{x})\colon |\lambda|_1=\infty\}.$$

$$A_2=\left\{(\lambda,\mathbf{x})\colon \limsup_{N\to\infty} \left(\sum_1^N x_i-N|\lambda|_1\right)/(2N\log\log N)^{1/2}=|\lambda|_2\right\}$$
$$\cup\{(\lambda,\mathbf{x})\colon |\lambda|_2=\infty\}.$$

$$A_3=\left\{(\lambda,\mathbf{x})\colon N^{-1}\sum_1^N \delta_{x_i}\Rightarrow\lambda \text{ as } N\to\infty\right\}.$$

Let us also give two examples of statutes representing trivial theorems.

$$A_4=\bigcup_k \{(\lambda,\mathbf{x}): x_i\in\text{support}(\lambda) \text{ for all } i>k\}.$$

$$A_5=\left\{(\lambda,\mathbf{x}):\ \lambda\in\Lambda_a,\ \lim_{N\to\infty}\sum_1^N a_i x_i \text{ exists}\right\}$$
$$\cup\left\{(\lambda,\mathbf{x}):\ \lambda\notin\Lambda_a,\ \lim_{N\to\infty}\sum_1^N a_i x_i \text{ does not exist}\right\},$$

where $a_i\to 0$ is a fixed sequence of reals and

$$\Lambda_a=\left\{\lambda\in\mathscr{P}(\mathbb{R}):\ \sum_1^\infty a_i V_i \text{ converges a.s. for } \mathscr{L}(\mathbf{V})=\lambda^*\right\}.$$

Call $A$ a *limit* statute if the following additional condition is satisfied.

$$\text{If}\quad (\lambda,\mathbf{x})\in A \quad\text{and}\quad \sum|x_i'-x_i|<\infty \quad\text{then } (\lambda,\mathbf{x}')\in A. \tag{3.7}$$

It is easily seen that this condition is satisfied by the above examples, except for $A_4$.

Now suppose $\mu$ happens to be the canonical random measure for an exchangeable vector $\mathbf{Z}$. Then for any statute $A$,

$$(\mu(\omega),\mathbf{Z}(\omega))\in A \text{ a.s.} \tag{3.8}$$

Because by Lemma 1 and (2.4)

$$E(I((\mu,\mathbf{Z})\in A)|\mu)(\omega)=\mu^*(\omega)(\{\mathbf{x}:(\mu(\omega),\mathbf{x})\in A\}) \text{ a.s.}$$
$$=1 \text{ a.s.}\quad \text{by (3.5).}$$

So (3.8) shows that classical a.s. theorems for i.i.d.r.v.'s extend immediately to exchangeable r.v.'s: of course this idea is well known. However, (3.8) provides the motivation for our first result, which is immediate if $\mathbf{X}$ has property $\mathfrak{P}$.

**Theorem 3.** *Let $A$ be any limit statute. Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{n}\subset\mathbf{m}$,

$$(\mu(\omega),\mathbf{X}_n(\omega))\in A \text{ a.s.} \tag{3.9}$$

Let us write out in detail one special case. Consider the statute $A_1$ defined above, representing the strong law of large numbers. Suppose $\{X_i\}$ is such that $\sup E|X_i|<\infty$: then certainly (3.1) holds. Then (3.9) asserts that, on a set of probability 1, either $|\mu(\omega)|_1=\infty$ or

$$N^{-1}\sum_1^N X_{n,i}(\omega)\to|\mu(\omega)|_1.$$

But from (3.2), $\alpha(\omega)=|\mu(\omega)|_1<\infty$ a.s., and so we have obtained the result of Komlós [11] mentioned earlier, in the following form.

**Corollary 4.** *Suppose* $\sup E|X_i| < \infty$. *Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{n} \subset \mathbf{m}$,

$$N^{-1} \sum_{1}^{N} X_{n,i} \to \alpha \text{ a.s.} \tag{3.10}$$

Similarly, applying Theorem 3 to statute $A_2$ gives the following result of Chatterji [3] and Gaposhkin [8].

**Corollary 5.** *Suppose* $\sup E X_i^2 < \infty$. *Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{n} \subset \mathbf{m}$,

$$\limsup_{N\to\infty} \left(\sum_{1}^{N} X_{n,i} - N\alpha\right) \Big/ (2N \log\log N)^{1/2} = \beta \text{ a.s.}$$

Here $\beta$ is as in (3.3). The analogous result for Strassen's functional law of the iterated logarithm (Berkes [1]) follows equally easily from Theorem 3. It should be clear how to apply Theorem 3 to any a.s. limit theorem, subject only to the technical condition (3.7), concerning which we make the following remark. It might seem more natural to express the idea of a *limit* theorem by the weaker condition: the set $\{\mathbf{x}: (\lambda, \mathbf{x}) \in A\}$ is in the tail $\sigma$-field of $\mathbb{R}^\infty$ for each $\lambda \in \mathscr{P}(\mathbb{R})$. However, this latter condition is satisfied by statute $A_4$ above, but Theorem 3 fails for $A_4$. For suppose $X_i = 2^{-i}$ a.s.: then by Lemma 1, $\mu(\omega) = \delta_0$ a.s. and assertion (3.9) fails. Fortunately, condition (3.7) seems to be satisfied by the statutes representing most non-trivial theorems.

The paper of Berkes [1] exemplifies the technique previously used to prove the special cases mentioned. Briefly, the idea was to reduce from general $\mathbf{X}$ to the case where $\mathbf{X}$ is a martingale difference sequence (m.d.s.); then to show that the particular theorem under consideration has a version for m.d.s.'s; and finally to show that an arbitrary m.d.s. has a subsequence satisfying these conditions. This is a non-trivial task, and clearly requires more work than the original proof of the theorem for i.i.d.r.v.'s. On the other hand, to apply Theorem 3 we need know only the statement of the result for i.i.d.r.v.'s. Clearly Theorem 3 cannot be proved by the above technique.

The assertions of (3.8) and Theorem 3 may be informally stated as follows. *Given any a.s. limit theorem for i.i.d.r.v.'s, there is an analogous theorem satisfied by exchangeable sequences and by all sub-subsequences of some subsequence of an arbitrarily-dependent sequence of r.v.'s.* The same result for weak limit theorems is given in (3.17) and Theorem 6, but in a somewhat less concise manner.

Let $\mathscr{P}_0(\mathbb{R})$, $S$, $\{h_k\}$, $\Phi$ be such that

$$\mathscr{P}_0(\mathbb{R}) \text{ is a measurable subset of } \mathscr{P}(\mathbb{R}); \tag{3.11}$$

$$S \text{ is a separable metrisable space;} \tag{3.12}$$

$$h_k: \mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^\infty \to S \text{ is continuous, for each } k; \tag{3.13}$$

$$\Phi: \mathscr{P}_0(\mathbb{R}) \to \mathscr{P}(S) \text{ is measurable;} \tag{3.14}$$

$$\text{if } \lambda \in \mathscr{P}_0(\mathbb{R}) \text{ and } \mathscr{L}(\mathbf{V}) = \lambda^* \text{ then } \mathscr{L}(h_k(\lambda, \mathbf{V})) \Rightarrow \Phi(\lambda). \tag{3.15a}$$

The last condition conveys the essential idea of a weak limit theorem: examples will be given after the statement of Theorem 6. For $f \in L^\infty(S)$ and $\phi \in \mathscr{P}(S)$, write

$$\langle f, \phi \rangle = \int f \, d\phi.$$

Then (3.15a) is equivalent to:

$$\int f(h_k(\lambda, \mathbf{x}))\, \lambda^*(d\mathbf{x}) \to \langle f, \Phi(\lambda) \rangle \tag{3.15b}$$

for each $\lambda \in \mathscr{P}_0(\mathbb{R})$ and each $f \in C(S)$. Suppose, as in the discussion before Theorem 3, that $\mu$ happens to be the canonical random measure for an exchangeable random vector $\mathbf{Z}$. Standard arguments show the existence of an element $E\Phi(\mu)$ of $\mathscr{P}(S)$ such that

$$\langle f, E\Phi(\mu) \rangle = E \langle f, \Phi(\mu) \rangle \quad \text{for each } f \in L^\infty(S). \tag{3.16}$$

We assert that, provided $\mu(\omega) \in \mathscr{P}_0(\mathbb{R})$ a.s.,

$$\mathscr{L}(h_k(\mu, \mathbf{Z})) \Rightarrow E\Phi(\mu). \tag{3.17}$$

To prove this, consider $f \in C(S)$. By (2.4) and Lemma 1,

$$\begin{aligned} E(f(h_k(\mu, \mathbf{Z})) \mid \mu)(\omega) &= \int f(h_k(\mu(\omega), \mathbf{x}))\, \mu^*(\omega, d\mathbf{x}) \text{ a.s.} \\ &\to \langle f, \Phi(\mu(\omega)) \rangle \text{ a.s.} \quad \text{by (3.15b)} \end{aligned}$$

and so, using (3.16),

$$E f(h_k(\mu, \mathbf{Z})) \to \langle f, E\Phi(\mu) \rangle.$$

This establishes (3.17). In order to extend this to subsequences we will make use of the following, somewhat arbitrary, technical condition.

There exist constants $\{c_{k,i}\}$ and a metric $d$ generating the topology on $S$ such that

$$d(h_k(\mathbf{x}), h_k(\mathbf{x}')) \leqq \sum_1^\infty c_{k,i} |x_i - x_i'|; \tag{3.18}$$

$$0 \leqq c_{k,i} \leqq 1 \tag{3.19}$$

$$\lim_{k \to \infty} c_{k,i} = 0 \quad \text{for each } i. \tag{3.20}$$

In fact these conditions can be weakened, their purpose only being to ensure that, for large $k$, $h_k(\mathbf{x})$ does not depend much on the first few coordinates of $\mathbf{x}$. The following result is easy if $\mathbf{X}$ happens to satisfy $\mathfrak{P}$.

**Theorem 6.** *Suppose that* (3.11)–(3.15) *and* (3.18)–(3.20) *hold. Suppose also that* $\mu(\omega) \in \mathscr{P}_0(\mathbb{R})$ *a.s. Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{n} \subset \mathbf{m}$,

$$\mathscr{L}(h_k(\mu, \mathbf{X}_n)) \Rightarrow E\Phi(\mu). \tag{3.21}$$

Theorem 6 has been formulated so as to be applicable to Donsker's theorem (see [2] for background). Let

$\mathscr{P}_0(\mathbb{R})=\{\lambda: |\lambda|^2<\infty \text{ and } |\lambda|_1=0\}$;

$S=D[0,1]$ with the Skorokhod $J_1$ topology;

$h_k(\lambda,\mathbf{x})$ be the function taking the value $k^{-1/2}\sum_1^{[kt]} x_i$ at $t$;

$\Phi(\lambda)=\mathscr{L}(|\lambda|_2\cdot W)$, where $W$ is a random map into $D[0,1]$ distributed as Wiener measure.

Applying Theorem 6 gives the following result (recall that $\beta(\omega)=|\mu(\omega)|_2$).

**Corollary 7.** *Suppose* $\sup EX_i^2<\infty$, *and* $X_i\to 0\,\sigma(L^1,L^\infty)$. *Then there exists an* $\mathbf{m}$ *such that, for each* $\mathbf{n}\subset\mathbf{m}$,

$$k^{-1/2}\sum_1^{[kt]} X_{n,i}\xrightarrow[\mathscr{D}]{}\beta\cdot W$$

*where* $W$ *is taken independent of* $\beta$.

This perhaps requires a few words of proof. Conditions (3.11)–(3.14) are straightforward, and (3.15) is Donsker's theorem. To verify (3.18)–(3.20), note that the $J_1$ metric may be taken smaller then the uniform metric $d_0$ which satisfies

$$d_0(h_k(\lambda,\mathbf{x}),h_k(\lambda,\mathbf{x}'))\leqq k^{-1/2}\sum_1^k |x_i-x_i'|.$$

To verify that $\mu(\omega)\in\mathscr{P}_0(\mathbb{R})$ a.s., it is necessary to show that $\alpha(\omega)=0$ and $\beta(\omega)<\infty$ a.s. The latter follows from (3.3): for the former, observe that $X_i\to X_\infty\,\sigma(L^1,L^\infty)$ implies $X_\infty=\alpha$ a.s., by (3.11). Finally, it follows from (3.16) that $E\Phi(\mu)$ is indeed the distribution of $\beta\cdot W$.

Let us remark that the hypothesis $X_i\to 0\,\sigma(L^1,L^\infty)$ is no real restriction. The $L^2$-boundedness implies that $X_i'=X_{q,i}-X_\infty\to 0\,\sigma(L^1,L^\infty)$ for some $\mathbf{q}, X_\infty$ and so the Corollary may be applied to this sequence. The same remark holds for Theorem 9.

From Corollary 7 we may deduce in the usual manner analogues for subsequences of the classical convergence in distribution theorems, in particular the central limit theorem due to Chatterji [4], Gaposhkin [8].

Theorem 6 has also a somewhat different application.

**Corollary 8.** *Suppose* $\mu$ *is the canonical random measure for an exchangeable vector* $\mathbf{Z}$. *Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{n}\subset\mathbf{m}$,

$$\mathscr{L}(\mu, X_{n,i}, X_{n,i+1},\ldots)\Rightarrow\mathscr{L}(\mu,\mathbf{Z})\quad \text{as } i\to\infty. \tag{3.22}$$

This follows by considering

$\mathscr{P}_0(\mathbb{R})=\mathscr{P}(\mathbb{R})$;

$S=\mathbb{R}^\infty$;

$h_k(\lambda,x_1,x_2,\ldots)=(\lambda,x_{k+1},x_{k+2},\ldots)$;

$\Phi(\lambda)=\lambda^*$;

$d(\mathbf{x},\mathbf{x}')=\sum 2^{-i}\min(1,|x_i-x_i'|)$.

The conditions are easily verified. It is possible to prove Corollary 8 directly. Indeed, the slightly weaker result in which $\mu$ is omitted in (3.22) is known, being implicit in Dacunha-Castelle [6], and Kingman has an unpublished proof. Now (3.22) expresses a weak sense in which $\mathbf{X}_n$ has similar properties to $\mathbf{Z}$, but this does not seem strong enough to be used to prove the more general results.

Theorems 3 and 6 show that most "deep" results for i.i.d.r.v.'s have analogues for subsequences. Paradoxically, analogues are harder to find for certain rather superficial results. Let us give one positive result and mention an open problem. Let

$$\ell_2 = \{\mathbf{a} \in \mathbb{R}^\infty : \sum a_i^2 < \infty\}.$$

If $|\lambda|_1 = 0$, $|\lambda|_2 < \infty$ and $\mathscr{L}(\mathbf{V}) = \lambda^*$, it is well known that

$$\sum a_i V_i \quad \text{converges a.s., for each } \mathbf{a} \in \ell_2. \tag{3.23}$$

Now suppose $\mu$ is the canonical random measure for the exchangeable vector $\mathbf{Z}$, and that $\alpha(\omega) = 0$ and $\beta(\omega) < \infty$ a.s. From (2.4) and Lemma 1,

$$\sum a_i Z_i \quad \text{converges a.s., for each } \mathbf{a} \in \ell_2. \tag{3.24}$$

Let $\mathscr{S}$ denote the set of permutations $\sigma = \{\sigma(i)\}$ of $\mathbb{Z}^+$. Then (2.5) extends (3.24) to

$$\sum a_{\sigma(i)} Z_{\sigma(i)} \quad \text{converges a.s., for each } \mathbf{a} \in \ell_2,\ \sigma \in \mathscr{S}. \tag{3.25}$$

Thus the following result is trivial if $\mathbf{X}$ happens to satisfy $\mathfrak{P}$.

**Theorem 9.** *Suppose* $\sup E X_i^2 < \infty$ *and* $X_i \to 0\, \sigma(L^1, L^\infty)$. *Then there exists* $\mathbf{m}$ *such that, for each* $\mathbf{a} \in \ell_2$ *and each* $\sigma \in \mathscr{S}$,

$$\sum a_{\sigma(i)} X_{m,\sigma(i)} \quad \textit{converges a.s.} \tag{3.26}$$

This is the assertion of Révész [14] Theorem 5.1.1, but the proof given there is fallacious, as observed in [8].

To state the open problem, write

$$\mathscr{A} = \{\mathbf{a} \in \mathbb{R}^\infty : a_i \to 0\},$$

$$\mathscr{C}(\mathbf{V}) = \{\mathbf{a} \in \mathscr{A} : \sum a_i V_i \text{ converges a.s.}\}.$$

For exchangeable $\mathbf{Z}$ it is clear that $\mathscr{C}(\mathbf{Z}_n) = \mathscr{C}(\mathbf{Z})$ for all $\mathbf{n}$. So if $\mathbf{X}$ happens to satisfy $\mathfrak{P}$, then there exists $\mathbf{m}$ such that

$$\mathscr{C}(\mathbf{X}_n) = \mathscr{C}(\mathbf{X}_m) \quad \text{for each } \mathbf{n} \subset \mathbf{m}. \tag{3.27}$$

But it is not known whether this holds in general. A partial result can be obtained from Theorem 3. Fix $\mathbf{a} \in \mathscr{A}$, and consider the statute $A_5$ defined earlier. Theorem 3 asserts: there exists $\mathbf{m}$ such that, for each $\mathbf{n} \subset \mathbf{m}$,

$$P(\mu(\omega) \in \Lambda_a, \sum a_i X_{n,i} \text{ converges}) + P(\mu(\omega) \notin \Lambda_a, \sum a_i X_{n,i} \text{ does not converge}) = 1.$$

Hence $\sum a_i X_{n,i}$ converges a.s. if and only if $\mu(\omega) \in \Lambda_a$ a.s. In other words,

*for a prescribed* $\mathbf{a} \in \mathscr{A}$, *there exists* $\mathbf{m}$ *such that*

$$\mathscr{C}(\mathbf{X}_n) \cap \{\mathbf{a}\} = \mathscr{C}(\mathbf{X}_m) \cap \{\mathbf{a}\} \quad \textit{for each } \mathbf{n} \subset \mathbf{m}.$$

Because a countable intersection of statutes is a statute, this extends to a prescribed countable subset $\mathscr{A}_0$ of $\mathscr{A}$. However, for the general case it is necessary to consider an uncountable number of conditions, and this would require the type of uniformity arguments which are to be used in the proof of Theorem 9.

To end this Section, let us discuss property $\mathfrak{P}$. Since the probabilistic properties of a random sequence $\mathbf{V}=\{V_i\}$ depend only on $\mathscr{L}(\mathbf{V})$, it is natural to extend $\mathfrak{P}$ as follows.

**Definition.** Say random vectors $\mathbf{V}$ and $\mathbf{Y}$ are *asymptotically equivalent* if there exist, on some probability space, random vectors $\mathbf{V}', \mathbf{Y}'$ such that

$$\mathscr{L}(\mathbf{V}')=\mathscr{L}(\mathbf{V}), \quad \mathscr{L}(\mathbf{Y}')=\mathscr{L}(\mathbf{Y}), \quad \sum |V_i'-Y_i'|<\infty \text{ a.s.} \tag{3.28}$$

The reader may easily verify:

$$\text{asymptotic equivalence is an equivalence relation.} \tag{3.29}$$

**Definition.** Say $\mathbf{X}$ has property $\mathfrak{P}'$ if there exist $\mathbf{m}$ and an exchangeable $\mathbf{Z}$ such that $\mathbf{X}_m$ is asymptotically equivalent to $\mathbf{Z}$.

We have remarked that all our results, as well as (3.27), are easy for $\mathbf{X}$ satisfying $\mathfrak{P}$, and that in Section 7 an example of $\mathbf{X}$ not satisfying $\mathfrak{P}$ will be given: replacing $\mathfrak{P}$ by $\mathfrak{P}'$ involves no essential change in the arguments. This suggests two questions. Is there a simple property which holds for all $\mathbf{X}$ and is strong enough to imply the results of this paper? Such a property would have to be intermediate between $\mathfrak{P}'$ and (3.22). And can one give sufficient conditions on $\mathbf{X}$ for $\mathfrak{P}'$ to hold? A partial answer to the latter is given below.

**Definition.** $X_i \underset{\mathscr{D}}{\longrightarrow} v$ *(mixing)* if $P(X_i \leqq x \mid B) \to v((-\infty, x])$ for each continuity point $x$ of $v$ and each measurable $B$, $P(B)>0$.

**Theorem 10.** *Suppose $X_i \underset{\mathscr{D}}{\longrightarrow} v$ (mixing). Then there exist $\mathbf{m}$ and $\mathbf{V}$ such that $\mathbf{X}_m$ is asymptotically equivalent to $\mathbf{V}$, and $\mathscr{L}(\mathbf{V})=v^*$.*

It is well known ([5], page 354) that if $\mathscr{L}(X_i) \Rightarrow v$ and $\{X_i\}$ has trivial tail $\sigma$-field, then $X_i \underset{\mathscr{D}}{\longrightarrow} v$ (mixing). It easily follows from Theorem 10 that if $\{X_i\}$ has purely atomic tail $\sigma$-field then $\{X_i\}$ satisfies $\mathfrak{P}'$. The author knows no more general sufficient condition.

Finally, let us say that although the results have been stated and will be proved for real-valued r.v.'s, Theorems 3 and 6 hold *mutatis mutandis* for sequences of separable metric space valued maps. But in general there are no interesting results to which they may be applied.

## 4. Construction of $\mu$

The first step in the proof of the results is the construction of the random measure $\mu$. This is somewhat similar to a construction in Révész [14] Theorem 6.1.1, but Proposition 11 below carries some additional information.

We are given a sequence $\mathbf{X}=\{X_i\}$ whose distributions are tight. So no generality is lost in Proposition 11 by assuming

$$\mathscr{L}(X_i) \Rightarrow v. \tag{4.1}$$

Let $D_\infty$ be a countable dense set of continuity points of $v$. For each $k$, choose a finite set $D_k=\{x_{k,i}\}_{i=1}^{q_k+1}$ such that

$$x_{k,i}<x_{k,i+1}<x_{k,i}+2^{-k}; \quad i\leqq q_k,\ k\in\mathbb{Z}^+. \tag{4.2}$$

$$P(x_{k,1}<X_n<x_{k,q_k})\geqq 1-2^{-k}; \quad k,n\in\mathbb{Z}^+. \tag{4.3}$$

$$D_k\subset D_{k+1} \quad \text{for each } k. \tag{4.4}$$

$$D_\infty=\bigcup_k D_k. \tag{4.5}$$

Let $\mathscr{J}_k$ be the set of intervals $(-\infty, x_{k,1}]$, $(x_{k,1}, x_{k,2}]$, ... $(x_{k,q_k}, \infty)$, and let $\mathscr{J}=\bigcup_k \mathscr{J}_k$. Define $\rho_k:\mathbb{R}\to D_k$ by

$$\begin{aligned}\rho_k(x)&=\inf\{x_{k,i}\in D_k: x_{k,i}\geqq x\} \quad \text{for } x\leqq x_{k,q_k}\\ &=x_{k,q_k+1} \quad \text{otherwise.}\end{aligned} \tag{4.6}$$

So $\rho_k$ is constant on each $J\in\mathscr{J}_k$. For future reference, note that (4.2) and (4.3) imply

$$P(|\rho_k(X_n)-X_n|\geqq 2^{-k})\leqq 2^{-k}; \quad k,n\in\mathbb{Z}^+ \tag{4.7}$$

and moreover, for any $\mathbf{V}$ such that $\mathscr{L}(\mathbf{V})=v$,

$$P(|\rho_k(V)-V|\geqq 2^{-k})\leqq 2^{-k}; \quad k\in\mathbb{Z}^+. \tag{4.8}$$

**Proposition 11.** *There exists a subsequence $\mathbf{Y}=\mathbf{X}_n$ and a random measure $\mu$ such that, for each $J\in\mathscr{J}$,*

$$E(I(Y_k\in J)|\mathscr{F}_{k-1})(\omega)\to\mu(\omega, J) \text{ a.s.}$$

*where*

$$\mathscr{F}_{k-1}=\mathscr{F}(\rho_1(Y_1), \dots, \rho_{k-1}(Y_{k-1})).$$

*Proof.* For each $x\in D_\infty$, the random variables $\{I(X_i\leqq x)\}$ are uniformly bounded and hence $\sigma(L^1, L^\infty)$ sequentially compact. Using the diagonal argument, we may assume without loss of generality that for each $x\in D_\infty$,

$$I(X_i\leqq x)\to G_x\sigma(L^1, L^\infty) \tag{4.9}$$

for some random variables $G_x$. Suppose, inductively, that $1=n_1<\dots<n_{k-1}$ have been defined. It follows from (4.9) that

$$E(I(X_i\leqq x)|A)\to E(G_x|A)$$

for each $x \in D_k$ and each $A \in \mathscr{F}_{k-1}$. But $D_k$ and $\mathscr{F}_{k-1}$ are finite, and so $n_k$ may be chosen sufficiently large that, putting $Y_k = X_{n_k}$,

$$|E(I(Y_k \leqq x)|A) - E(G_x|A)| \leqq 2^{-k} \tag{4.10}$$

for each $x \in D_k$ and each $A \in \mathscr{F}_{k-1}$ with $P(A) > 0$.

Having thus constructed $\mathbf{n}$ and $\mathbf{Y}$, let $\mathscr{F} = \bigvee_k \mathscr{F}_k$. Consider some fixed $x \in D_\infty$: then $x \in D_k$ for large $k$. Then for such $k$, (4.10) shows that

$$|E(I(Y_k \leqq x)|\mathscr{F}_{k-1})(\omega) - E(G_x|\mathscr{F}_{k-1})(\omega)| \leqq 2^{-k} \text{ a.s.}$$

But

$$E(G_x|\mathscr{F}_{k-1})(\omega) \to E(G_x|\mathscr{F})(\omega) \text{ a.s.}$$

So to finish the proof of the Proposition it is necessary only to construct $\mu$ such that, for each $x \in D_\infty$,

$$\mu(\omega, (-\infty, x]) = E(G_x|\mathscr{F})(\omega) \text{ a.s.} \tag{4.11}$$

This is quite straightforward. From (4.9) we see that $x_1, x_2 \in D_\infty$ and $x_1 < x_2$ imply that $G_{x_1} \leqq G_{x_2}$ a.s. and $E(G_{x_2} - G_{x_1}) = v((x_1, x_2])$. So, as in Révész [14] Lemma 6.1.4, we can construct versions $\lambda_x$ of $E(G_x|\mathscr{F})$ such that, for each fixed $\omega$, the function $\lambda_x(\omega)$ extends to a distribution function. Let $\mu(\omega)$ be the corresponding measure. Then (4.11) is satisfied, and the measurability of $\mu$ (considered as a random map into $\mathscr{P}(\mathbb{R})$) follows from (2.1).

*Remarks.* Let $\theta_k$ be a r.c.d. for $Y_k$ given $\mathscr{F}_{k-1}$. It follows easily from Proposition 11 that $\theta_k(\omega) \Rightarrow \mu(\omega)$ a.s. This suggests, informally, that a sample path $\{Y_k(\omega)\}$ should look asymptotically like a typical sample path from an i.i.d. sequence with distribution $\mu(\omega)$. The main line of argument continues in the next Section by constructing an exchangeable sequence $\mathbf{Z}$ whose canonical random measure is $\mu$, and comparing properties of $\mathbf{Y}$ and $\mathbf{Z}$. But we now digress to prove Theorem 10.

*Proof of Theorem 10.* Let $\{T_i\}$ be an independent sequence of r.v.'s distributed uniformly on $[0,1]$ and defined on some space $(\Omega', \mathscr{B}', P')$. For $\lambda \in \mathscr{P}(\mathbb{R})$, let $F_\lambda^{-1}$ denote the inverse distribution function

$$\begin{aligned} F_\lambda^{-1}(t) &= \inf\{x : \lambda(-\infty, x] \geqq t\} \quad \text{for } 0 < t < 1 \\ &= 0, \quad \text{say, otherwise.} \end{aligned} \tag{4.12}$$

It is well known that $\mathscr{L}(F_\lambda^{-1}(T_1)) = \lambda$ and

$$\lambda_j \Rightarrow \lambda \quad \text{if and only if} \quad F_{\lambda_j}^{-1}(T_1) \to F_\lambda^{-1}(T_1) \text{ a.s.} \tag{4.13}$$

The proof of Theorem 10 depends on the following construction, which does not seem helpful in the general case. We define inductively random measures $\theta_i$ on $\Omega$, and random measures $\phi_i$ and r.v.'s $W_i$ on $\Omega'$. Let $\theta_1$ and $\phi_1$ be identically $\mathscr{L}(Y_1)$. Let $W_1 = F_{\phi_1}^{-1}(T_1)$. Now suppose $i > 1$. For each atom $G = \{\rho_1(Y_1) = x_1, \ldots, \rho_{i-1}(Y_{i-1}) = x_{i-1}\}$ of $\mathscr{F}_{i-1}$, let $G' = \{\rho_1(W_1) = x_1, \ldots, \rho_{i-1}(W_{i-1}) = x_{i-1}\}$ be the corresponding atom of $\mathscr{F}'_{i-1}=\mathscr{F}(\rho_1(W_1),\ldots,\rho_{i-1}(W_{i-1}))$; for $\omega\in G$ and $\omega'\in G'$ let $\theta_i(\omega)$ and $\phi_i(\omega')$ be the conditional distribution of $Y_i$ given $G$ (or $\delta_0$, say, if $P(G)=0$). Let $W_i=F_{\phi_i}^{-1}(T_i)$.

It should be clear from the construction that

$$\mathscr{L}(\{\rho_i(W_i)\}_1^\infty)=\mathscr{L}(\{\rho_i(Y_i)\}_1^\infty); \tag{4.14}$$

$$\mathscr{L}(W_i)=\mathscr{L}(Y_i) \quad \text{for each } i; \tag{4.15}$$

$$\mathscr{L}(\theta_i)=\mathscr{L}(\phi_i) \quad \text{for each } i. \tag{4.16}$$

Now by (4.7) and the Borel-Cantelli Lemma, $\mathbf{Y}$ is asymptotically equivalent to $\{\rho_i(Y_i)\}$ in the sense of (3.28); and similarly for $\mathbf{W}$, by (4.15). So it follows from (4.14) and (3.29) that

$$\mathbf{W} \text{ is asymptotically equivalent to } \mathbf{Y}. \tag{4.17}$$

Recall that the hypothesis of Theorem 10 is that $\mathscr{L}(X_i) \Rightarrow \nu$ (mixing). So in (4.9) we may take $G_x$ to be identically $\nu((-\infty,x])$. And in (4.11) we may take $\mu(\omega)$ to be identically $\nu$. Proposition 11 shows that for each $J\in\mathscr{J}$,

$$\theta_i(\omega)(J)=E(I(Y_i\in J)|\mathscr{F}_{i-1})(\omega)\to\nu(J) \text{ a.s.}$$

and so from [2] Theorem 2.2

$$\theta_i(\omega) \Rightarrow \nu \text{ a.s.} \tag{4.18}$$

Consider the metric on $\mathscr{P}(\mathbb{R})$ given by

$$d(\lambda,\lambda')=E\min(1,|F_\lambda^{-1}(T_1)-F_{\lambda'}^{-1}(T_1)|).$$

Using (4.13), we see that $d$ generates the weak topology. Now

$$E(\min(1,|F_{\phi_i}^{-1}(T_i)-F_\nu^{-1}(T_i)|)|\mathscr{F}'_{i-1})(\omega')=d(\phi_i(\omega'),\nu) \text{ a.s.}$$

because $\phi_i$ is $\mathscr{F}'_{i-1}$-measurable and $T_i$ is independent of $\mathscr{F}'_{i-1}$. So from the definition of $W_i$,

$$\begin{aligned} E\min(1,|W_i-F_\nu^{-1}(T_i)|)&=Ed(\phi_i,\nu)\\ &=Ed(\theta_i,\nu) \quad \text{by (4.16)}\\ &\to 0 \quad \text{by (4.18)}. \end{aligned}$$

So $W_i-F_\nu^{-1}(T_i)\to 0$ in probability, and so for some $\mathbf{n}$,

$$\sum|W_{n,i}-F_\nu^{-1}(T_{n,i})|<\infty.$$

Hence $\mathbf{W}_n$ is asymptotically equivalent to the random vector $\{F_\nu^{-1}(T_{n,i})\}$ whose distribution is $\nu^*$. Theorem 10 now follows from (4.17) and (3.29).

*Remark.* The hypothesis of mixing is used only to show that $\mu$ is essentially constant. A modification of the argument shows that if $\mu$ is essentially countably-valued, then $\mathbf{X}$ satisfies $\mathfrak{P}'$.

## 5. The Fundamental Result

We may now restrict attention to the sequence $\mathbf{Y}$ produced in Proposition 11. Recall that $\mu$ is $\mathscr{F}$-measurable, where

$$\mathscr{F}=\bigvee_k \mathscr{F}_k \subset \mathscr{F}(\mathbf{Y}). \tag{5.1}$$

Let $\mathbf{Z}$ be a random vector such that

$$\mu^* \text{ is a r.c.d. for } \mathbf{Z} \text{ given } \mathscr{F}(\mathbf{Y},\mu). \tag{5.2}$$

To be more precise, let $\mathbf{Y}', \mu', \mathbf{Z}'$ be the natural maps on the canonical space $\Omega'=\mathbb{R}^\infty \times \mathscr{P}(\mathbb{R}) \times \mathbb{R}^\infty$; by a standard construction (Meyer [12] T.14) there exists a probability measure $P'$ on $\Omega'$ under which $\mathscr{L}(\mu',\mathbf{Y}')=\mathscr{L}(\mu,\mathbf{Y})$ and $\mu'^*$ is a r.c.d. for $\mathbf{Z}'$ given $F(\mathbf{Y}',\mu')$. But the results to be proved depend only on the distribution $\mathscr{L}(\mu, \mathbf{Y})$, and so we may suppose $\Omega'$ to be the original space and drop the prime.

Note that (5.2) is equivalent, because of (5.1), to

$$\mu^* \text{ is a r.c.d. for } \mathbf{Z} \text{ given } \mathscr{F}(\mathbf{Y}) \tag{5.3}$$

and implies that $\mu^*$ is a r.c.d. for $\mathbf{Z}$ given $\mathscr{F}(\mu)$, so that $\mu$ is the canonical random measure for the exchangeable vector $\mathbf{Z}$.

**Lemma 12.** *Let $V$ be an $\mathscr{F}(\mathbf{Y})$-measurable random map into some separable metrisable space $S$. Let $j\in\mathbb{Z}^+$. Then*

(i) $(V, Y_m)\xrightarrow[\mathscr{D}]{}(V,Z_j)$ *as* $m\to\infty$.

*Moreover, suppose $f\in L^\infty(S\times\mathbb{R})$ is such that, for each $s$,*

(ii) $f(s,\cdot)$ *is constant on each* $J\in\mathscr{J}_j$.

*Then* $Ef(V,Y_m)\to Ef(V,Z_j)$ *as* $m\to\infty$.

*Proof.* Let $J\in\mathscr{J}$ and let $A$ be a measurable subset of $S$. We shall prove

$$P(V\in A, Y_m\in J)\to P(V\in A, Z_j\in J). \tag{5.4}$$

Then case (i) will follow from [2] Theorem 2.2, and case (ii) will follow by approximating $f$ by simple functions.

To prove (5.4), define random variables

$$K_m=E(I(Y_m\in J)\cdot E(I(V\in A)|\mathscr{F}_{m-1})|\mathscr{F}_{m-1})$$
$$K'_m=E(I(Y_m\in J)\cdot E(I(V\in A)|\mathscr{F})|\mathscr{F}_{m-1})$$
$$K(\omega)=\mu(\omega,J)\cdot E(I(V\in A)|\mathscr{F})(\omega).$$

Because $Y_m$ is $\mathscr{F}$-measurable,

$$P(V\in A, Y_m\in J)=EK'_m. \tag{5.5}$$

And

$$K_m=E(I(Y_m\in J)|\mathscr{F}_{m-1})\cdot E(I(V\in A)|\mathscr{F}_{m-1})$$
$$\to K \text{ a.s.}$$

by Proposition 11 and the martingale limit theorem. As everything is bounded,

$$E K_m \to E K. \tag{5.6}$$

Using the martingale limit theorem again,

$$E|K_m - K'_m| \leqq E|E(I(V\in A)|\mathscr{F}_{m-1}) - E(I(V\in A)|\mathscr{F})| \to 0. \tag{5.7}$$

Finally, because $V$ is $\mathscr{F}(\mathbf{Y})$-measurable,

$$\begin{aligned} E(I(Z_j\in J, V\in A)|\mathbf{Y}) &= E(I(Z_j\in J)|\mathbf{Y})\cdot I(V\in A) \\ &= \mu(\omega, J)\cdot I(V\in A) \qquad \text{by (5.3)} \end{aligned}$$

where we omit the a.s. in equalities between conditional expectations. Now because $\mu$ is $\mathscr{F}$-measurable,

$$E(I(Z_j\in J, V\in A)|\mathscr{F}) = K$$

and so

$$P(Z_j\in J, V\in A) = EK. \tag{5.8}$$

The relations (5.5)–(5.8) imply (5.4) and thence the Lemma.

*Remark.* The reader familiar with conditional independence will see that $V$ and $\mathbf{Z}$ are conditionally independent given $\mathscr{F}$.

In particular, Lemma 12 implies that $Y_m \underset{\mathscr{D}}{\longrightarrow} Z_j$, and so from (4.1)

$$\mathscr{L}(Z_j) = \nu \qquad \text{for each } j. \tag{5.9}$$

Hence for $0<p<\infty$

$$E|Z_1|^p \leqq \liminf E|Y_m|^p \leqq \limsup E|X_i|^p.$$

But from (5.3) and Lemma 1,

$$E(|Z_1|^p|\mathbf{Y})(\omega) = |\mu(\omega)|^p \text{ a.s.}$$

in the notation of (2.7): this proves Lemma 2.

We are almost ready for the fundamental result of the paper, from which the results stated in Section 3 will be deduced in the next Section. Consider a function $g: \mathscr{P}(\mathbb{R})\times\mathbb{R}^\infty\to\mathbb{R}$. We would like to use the "asymptotic conditional independence" property given by Lemma 12 to show that $E\,g(\mu, \mathbf{Y}_n)$ is close to $E\,g(\mu, \mathbf{Z})$ whenever $\mathbf{n}=\{n_i\}$ increases rapidly. The situation is simplest when g is continuous, but the alternative conditions below will sometimes be needed.

For each $\mathbf{x}\in\mathbb{R}^\infty$, $\mathbf{y}\in\mathbb{R}^\infty$ and $\lambda\in\mathscr{P}(\mathbb{R})$:

$$g(\lambda, \mathbf{x}) \leqq \liminf g(\lambda, x_1, \ldots, x_i, y_{i+1}, y_{i+2}, \ldots); \tag{5.10}$$

$$\text{if } \rho_i(x_i) = \rho_i(y_i) \text{ for each } i \text{ then } g(\lambda, \mathbf{x}) = g(\lambda, \mathbf{y}), \tag{5.11}$$

where $\rho_i$ is as defined in (4.6).

One final definition is needed. Let $\mathfrak{Q}$ denote an assertion applicable to increasing sequences **n** of positive integers, and let $\mathfrak{Q}'$ denote the set of **n** for which $\mathfrak{Q}$ is true. If there exists a function $L$: {finite sequences in $\mathbb{Z}^+$} $\to \mathbb{Z}^+$ such that

$$\text{if } n_i \geqq L(n_1, \ldots, n_{i-1}) \text{ for each } i \text{ then } \mathbf{n} = \{n_i\} \in \mathfrak{Q}' \tag{5.12}$$

then we will say that $\mathfrak{Q}$ holds *for all* **n** *increasing sufficiently rapidly*. For example, suppose $V_i \to 0$ in probability. Then we can say that $\sum V_{n,i}$ converges a.s. for all **n** increasing sufficiently rapidly, because (5.12) is satisfied by the function

$$L(n_1, \ldots, n_{i-1}) = \min\{m > n_{i-1} : P(|V_q| > 2^{-i}) < 2^{-i} \text{ for all } q \geqq m\}.$$

**Proposition 13.** *Let $\mathscr{P}_0(\mathbb{R})$ be a measurable subset of $\mathscr{P}(\mathbb{R})$ such that*

$$\mu(\omega) \in \mathscr{P}_0(\mathbb{R}) \text{ a.s.} \tag{5.13}$$

*Let $\mathscr{P}_0(\mathbb{R})$ be equipped with a separable metrisable topology $\tau$ such that*

$$\lambda_i \xrightarrow[\tau]{} \lambda \text{ implies } \quad \lambda_i \Rightarrow \lambda; \tag{5.14}$$

$$\tau \text{ and the weak topology generate the same } \sigma\text{-field.} \tag{5.15}$$

*Let $\varepsilon > 0$ be given. Let $g: \mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^\infty \to \mathbb{R}$ satisfy either* (i) *$g$ is bounded and continuous, or* (ii) *$g$ is bounded, measurable and satisfies* (5.10) *and* (5.11). *Then*

$$E\,g(\mu, \mathbf{Y}_n) \leqq E\,g(\mu, \mathbf{Z}) + \varepsilon$$

*for all* **n** *increasing sufficiently rapidly.*

*Moreover, suppose $\{g^\theta\}$ is a uniformly bounded family of measurable functions $\mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^\infty \to \mathbb{R}$, each satisfying* (5.10), *and such that condition* (5.25) *below is satisfied. Then*

$$\sup_\theta (Eg^\theta(\mu, \mathbf{Y}_n) - Eg^\theta(\mu, \mathbf{Z})) \leqq \varepsilon$$

*for all* **n** *increasing sufficiently rapidly.*

*Proof.* Suppose g satisfies (i) or (ii). Suppose, inductively, that $n_1 < \cdots < n_k$ have been specified (the first step of the induction is similar to the general step). Define random variables

$$\begin{aligned} G_0 &= g(\mu, \mathbf{Z}) \\ G_i &= g(\mu, Y_{n,1}, \ldots, Y_{n,i}, Z_{i+1}, Z_{i+2}, \ldots), \qquad 1 \leqq i \leqq k. \end{aligned} \tag{5.16}$$

For $m > n_k$ define

$$\beta_m = g(\mu, Y_{n,1}, \ldots, Y_{n,k}, Y_m, Z_{k+2}, Z_{k+3}, \ldots).$$

We shall prove

$$E\beta_m \to EG_k \qquad \text{as } m \to \infty. \tag{5.17}$$

Let us accept (5.17) for the moment. Then there exists a number $L(n_1, \ldots, n_k)$ such that, by choosing $n_{k+1}$ to be any integer larger than $L(n_1, \ldots, n_k)$ and defining $G_{k+1}$ as in (5.16), we obtain

$$EG_{k+1} \leqq EG_k + \varepsilon 2^{-k-1}.$$

This shows that

$$\sup EG_k \leqq EG_0 + \varepsilon = Eg(\mu, \mathbf{Z}) + \varepsilon$$

for all $\mathbf{n}$ increasing sufficiently rapidly. But from (5.10), which follows from continuity in case (i),

$$g(\mu, \mathbf{Y}_n) \leqq \liminf G_k.$$

The Proposition follows from Fatou's Lemma.

It remains to prove (5.17). Define $g_i: \mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^i \to \mathbb{R}$ by

$$g_i(\lambda, y_1, \ldots, y_i) = \int g(\lambda, y_1, \ldots, y_i, x_{i+1}, x_{i+2}, \ldots) \lambda^*(d\mathbf{x}). \tag{5.18}$$

Certainly, $g_i$ is bounded and measurable. Define

$$V = (\mu, Y_{n,1}, \ldots, Y_{n,k}).$$

By (5.1), $V$ is an $\mathscr{F}(\mathbf{Y})$-measurable random map into $\mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^k$. For a random variable $W$, we may regard $(V, W)$ as a random map into $\mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^{k+1}$. Using Lemma 1, the definitions of the quantities involved, and (5.3),

$$E(G_k | \mathbf{Y}) = g_k(V) \quad \text{a.s.,} \tag{5.19}$$

$$E(\beta_m | Y) = g_{k+1}(V, Y_m) \quad \text{a.s.,} \tag{5.20}$$

$$\begin{aligned} E(g_{k+1}(V, Z_{k+1}) | \mathbf{Y}) &= \int g_{k+1}(V, t)\, \mu(\omega, dt) \quad \text{a.s.} \\ &= \iint g(\mu, Y_{n,1}, \ldots, Y_{n,k}, t, x_{k+2}, x_{k+3}, \ldots)\ \mu(\omega, dt) \mu^*(\omega, d\mathbf{x}) \quad \text{a.s.} \\ &= g_k(V) \quad \text{a.s.} \end{aligned} \tag{5.21}$$

Identities (5.19)–(5.21) show that assertion (5.17) is equivalent to the assertion:

$$Eg_{k+1}(V, Y_m) \to Eg_{k+1}(V, Z_{k+1}) \qquad \text{as } m \to \infty. \tag{5.22}$$

We can prove (5.22) from Lemma 12, considering the cases separately. In case (ii) it follows from (5.11) that $g_{k+1}(\lambda, y_1, \ldots, y_{k+1})$, considered as a function of $y_{k+1}$ for fixed $(\lambda, y_1, \ldots, y_k)$, is constant on each $J \in \mathscr{J}_{k+1}$, and so (5.22) follows from part (ii) of Lemma 12.

In case (i) it is sufficient to show that $g_{k+1}$ is continuous, for then (5.22) follows from part (i) of Lemma 12. Now the definition (5.18) of $g_{k+1}$ may be rewritten, in the notation of (4.13), as

$$g_{k+1}(\lambda, y_1, \ldots, y_{k+1}) = Eg(\lambda, y_1, \ldots, y_{k+1}, F_\lambda^{-1}(T_{k+2}), F_\lambda^{-1}(T_{k+3}), \ldots). \tag{5.23}$$

So if $\lambda_q \xrightarrow[\tau]{} \lambda$ then $\lambda_q \Rightarrow \lambda$ and so

$$F_{\lambda_q}^{-1}(T_i) \to F_\lambda^{-1}(T_i) \text{ a.s.}$$

Hence the continuity of $g_{k+1}$ follows from that of $g$.

We have now established the Proposition for $g$ satisfying (i) or (ii). Consider now a uniformly bounded family $\{g^\theta\}$ of functions, each satisfying (5.10). Define $g_i^\theta$ as in (5.18). If we can establish, in place of (5.22),

$$E g_{k+1}^\theta(V, Y_m) \to E g_{k+1}^\theta(V, Z_{k+1}) \quad \text{uniformly in } \theta, \tag{5.24}$$

then the desired result follows with only minor changes in the above argument. But from part (ii) of Lemma 12,

$$(V, Y_m) \xrightarrow[\mathscr{D}]{} (V, Z_{k+1}).$$

Hence ([2] page 17) a sufficient condition for (5.24) is that, for each $j$, the family $\{g_j^\theta\}$ be uniformly equicontinuous. To be more definite, we take the following condition to be a hypothesis of the Proposition.

*There exists a metric $d_\tau$ generating the topology $\tau$ on $\mathscr{P}_0(\mathbb{R})$ such that*

$$|g_j^\theta(\lambda, y_1, \ldots, y_j) - g_j^\theta(\lambda', y_1', \ldots, y_j')| \leqq d_\tau(\lambda, \lambda') + \sum_1^j |y_i - y_i'| \tag{5.25}$$

*for each $\theta$, each $j \in \mathbb{Z}^+$, each $\lambda, \lambda' \in \mathscr{P}_0(\mathbb{R})$ and each $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^\infty$.*

*Remarks.* The separability of $\tau$ is essential. The Proposition is false for $d_\tau(\lambda, \lambda') = \sup |\lambda(A) - \lambda'(A)|$, because Lemma 12 fails for non-separable-valued $V$.

Condition (5.25) can certainly be weakened, but the form given will be sufficient for proving Theorem 9. However, it always seems necessary to impose some form of equicontinuity on $\{g_j^\theta\}$ as a function of $\lambda$, and this is the difficulty in trying to prove (3.27).

## 6. Proofs of the Main Results

In this Section Theorems 3, 6 and 9 will be deduced from Proposition 13 and the following form of the diagonal argument.

**Lemma 14.** *Let $\{\mathfrak{Q}_{j,k}\}$, $1 \leqq k \leqq q_j$, $j \in \mathbb{Z}^+$ be a collection of properties, each of which holds for all $\mathbf{n}$ increasing sufficiently rapidly. Then there exists $\mathbf{m}$ satisfying the following assertion.*

*For each $\mathbf{n} \subseteq \mathbf{m}$ and each $j \in \mathbb{Z}^+$, there exists $\mathbf{n}^j$ such that:*

$$\mathbf{n}^j \text{ satisfies properties } \mathfrak{Q}_{j,k},\ 1 \leqq k \leqq q_j; \tag{6.1a}$$

$$n_i = n_i^j, \quad \text{for all } i > j. \tag{6.1b}$$

*Proof.* Consider first a single such property $\mathfrak{Q}$ and the associated function $L$ of (5.12). It is easy to construct inductively a function $L'$ such that, for all finite sequences $I_1 \subset I_2$ in $\mathbb{Z}^+$, $L'(I_2) \geqq L'(I_1) \geqq L(I_1)$. Now if $\mathbf{m}$ is such that $m_i \geqq L'(m_1, \ldots, m_{i-1})$ for all $i$, then each $\mathbf{n} \subset \mathbf{m}$ has property $\mathfrak{Q}$.

To prove the Lemma, we lose no generality in assuming $q_j = 1$ for all $j$, because $\mathfrak{Q}_{k,1}$ *and* $\mathfrak{Q}_{k,2} \ldots$ *and* $\mathfrak{Q}_{k,q_k}$ denotes a property holding for all $\mathbf{n}$ increasing sufficiently rapidly. By the remark above, we can choose inductively $\mathbf{m}^j \subset \mathbf{m}^{j-1}$ such that each $\mathbf{n} \subset \mathbf{m}^j$ has property $\mathfrak{Q}_{j,1}$. Then the diagonal sequence $m_i = m_i^i$ satisfies the assertion of the Lemma.

*Proof of Theorem 3.* Let $A$ be any limit statute, and let $A'$ be the complement of $A$ in $\mathscr{P}(\mathbb{R}) \times \mathbb{R}^\infty$. Clearly $A'$ satisfies (3.7). Accept for the moment the following Lemma.

**Lemma 15.** *Let* $j \in \mathbb{Z}^+$ *be given. Then* $P((\mu, \mathbf{Y}_n) \in A') \leqq 2^{-j}$ *for all* $\mathbf{n}$ *increasing sufficiently rapidly.*

Applying Lemma 14, there exists $\mathbf{m}$ satisfying the following assertion.
For each $\mathbf{n} \subset \mathbf{m}$ and each $j \in \mathbb{Z}^+$, there exists $\mathbf{n}^j$ such that:

$$P((\mu, \mathbf{Y}_{n^j}) \in A') \leqq 2^{-j}; \tag{6.2a}$$

$$n_i = n_i^j, \quad \text{for all } i > j. \tag{6.2b}$$

Fix $\mathbf{n} \subset \mathbf{m}$. By (3.7) and (6.2b),

$$P((\mu, \mathbf{Y}_n) \in A') = P((\mu, \mathbf{Y}_{n^j}) \in A') \quad \text{for each } j.$$

So by (6.2a),

$$P((\mu, \mathbf{Y}_n) \in A') = 0.$$

This establishes Theorem 3.

*Proof of Lemma 15.* We cannot apply Proposition 13 directly to the indicator function $I(A')$ because (5.11) is not satisfied, so a more devious approach is required.

Consider the functions $\rho_i$ defined in (4.6). It follows from (4.7), (4.8) and (5.9) that

$$\sum |Z_i - \rho_i(Z_i)| < \infty \text{ a.s.,}$$
$$\sum |Y_{n,i} - \rho_i(Y_{n,i})| < \infty \text{ a.s.} \quad \text{for each } \mathbf{n}.$$

Define $\rho: \mathbb{R}^\infty \to \mathbb{R}^\infty$ by

$$\rho(\mathbf{x}) = (\rho_1(x_1), \rho_2(x_2), \ldots).$$

Then from (3.7),

$$P((\mu, \rho(\mathbf{Y}_n)) \in A') = P((\mu, \mathbf{Y}_n) \in A') \quad \text{for each } \mathbf{n}; \tag{6.3}$$
$$P((\mu, \rho(\mathbf{Z})) \in A') = P((\mu, \mathbf{Z}) \in A')$$
$$= 0 \quad \text{by (3.8).}$$

Now a probability measure on a metric space is regular, so there exists an open set $G \supset A'$ such that

$$P((\mu, \rho(\mathbf{Z}) \in G) \leqq 2^{-j-1}. \tag{6.4}$$

Let $H=\{(\lambda, \mathbf{x}):(\lambda, \rho(\mathbf{x}))\in G\}$. Then $I(H)$ satisfies the hypotheses of part (ii) of Proposition 13, where we take $\mathscr{P}_0(\mathbb{R})=\mathscr{P}(\mathbb{R})$ and $\tau$ the weak topology: condition (5.10) follows from the openness of $G$. So Proposition 13 and (6.4) show that

$$P((\mu, \rho(\mathbf{Y}_n))\in G)\leqq 2^{-j}$$

for all $\mathbf{n}$ increasing sufficiently rapidly. The Lemma now follows from (6.3), since $A'\subset G$.

*Proof of Theorem 6.* Let $\mathscr{P}_0(\mathbb{R})$, $S$, $\{h_k\}$, $\Phi$, $\{c_{k,i}\}$ be as in the hypotheses (3.11)–(3.15) and (3.18)–(3.20) of the Theorem. The latter technical conditions are needed only to establish (6.5) below. For $\{x_{k,i}\}$ defined in (4.3), let

$$t_j=x_{j,q_j}-x_{j,1}.$$

Then

$$P(|Y_{i_1}-Y_{i_2}|>t_j)\leqq 2^{-j+1}; \qquad i_1, i_2, j\in\mathbb{Z}^+.$$

Using (3.20), choose $r_j$ such that

$$c_{k,i}\leqq (j^2 t_j)^{-1}; \qquad 1\leqq i\leqq j,\ k\geqq r_j.$$

Suppose $m_i=n_i$ for all $i>j$. Then from (3.18) and (3.19),

$$P(d(h_k(\mu, \mathbf{Y}_m), h_k(\mu, \mathbf{Y}_n))>j^{-1})\leqq j2^{-j+1}; \qquad k\geqq r_j. \tag{6.5}$$

Clearly we may assume that $\{r_j\}$ is strictly increasing.

Now $S$ is separable, and so by [13] Theorem 6.6 there exists a countable subset $\{f_q\}$ of $C(S)$ such that

$$\phi_k \Rightarrow \phi \quad \text{in } \mathscr{P}(S) \text{ if } \int f_q\, d\phi_k\to\int f_q\, d\phi \qquad \text{for each } q. \tag{6.6}$$

We may assume

$$f_{2q}=\underline{f}_{2q-1}, \qquad q\geqq 1. \tag{6.7}$$

Consider the function $f_q\, h_k$: $\mathscr{P}_0(\mathbb{R})\times\mathbb{R}^\infty\to\mathbb{R}$. This is bounded and continuous, because $h_k$ is by hypothesis (3.13) continuous with respect to the weak topology (in fact any topology satisfying (5.14) and (5.15) will suffice). Apply Proposition 13. We deduce that the property $\mathfrak{Q}_{j,k,q}$ defined by

$$Ef_q h_k(\mu, \mathbf{Y}_n)-Ef_q h_k(\mu, \mathbf{Z})\leqq j^{-1}$$

holds for all $\mathbf{n}$ increasing sufficiently rapidly. Now apply Lemma 14, and deduce that there exists $\mathbf{m}$ satisfying the following assertion.

For each $\mathbf{n}\subset\mathbf{m}$ and each $j\in\mathbb{Z}^+$ there exists $\mathbf{n}^j$ such that:

$$\mathbf{n}^j \text{ satisfies } \mathfrak{Q}_{j,k,q},\ 1\leqq q\leqq 2j,\ r_j\leqq k<r_{j+1}; \tag{6.8a}$$

$$n_i^j=n_i, \qquad i>j. \tag{6.8b}$$

Fix $\mathbf{n} \subset \mathbf{m}$. We must prove (3.21), that is

$$\mathscr{L}(h_k(\mu, \mathbf{Y}_n)) \Rightarrow E\Phi(\mu). \tag{6.9}$$

Let $j(k)$ denote the integer $j$ such that $r_j \leqq k < r_{j+1}$. Then $j(k) \to \infty$ as $k \to \infty$. And (6.7) and (6.8) imply, writing $\mathbf{n}^k$ for $\mathbf{n}^{j(k)}$,

$$|E f_q h_k(\mu, \mathbf{Y}_{n^k}) - E f_q h_k(\mu, \mathbf{Z})| \leqq 1/j(k); \qquad 1 \leqq q \leqq 2j(k). \tag{6.10a}$$

$$n_i^k = n_i, \qquad i > j(k). \tag{6.10b}$$

Recall that (3.17) asserted

$$\mathscr{L}(h_k(\mu, \mathbf{Z})) \Rightarrow E\Phi(\mu).$$

So by (6.6) and (6.10a),

$$\mathscr{L}(h_k(\mu, \mathbf{Y}_{n^k})) \Rightarrow E\Phi(\mu).$$

But by (6.5) and (6.10b),

$$d(h_k(\mu, \mathbf{Y}_n), h_k(\mu, \mathbf{Y}_{n^k})) \to 0 \qquad \text{in probability.}$$

Now (6.9) follows, using [2] Theorem 4.1.

*Proof of Theorem 9.* Recall that $\mathscr{S}$ denotes the set of permutations $\sigma$ of $\mathbb{Z}^+$. Let $\ell_2' = \{\mathbf{a} \in \mathbb{R}^\infty : \sum a_i^2 < 1\}$. It is clearly sufficient to prove convergence in (3.26) for $\mathbf{a} \in \ell_2'$ only. For $\mathbf{x} \in \mathbb{R}^\infty$, let $\sigma(\mathbf{x})$ denote the element $(x_{\sigma(i)})$ of $\mathbb{R}^\infty$ and let $\mathbf{a} \cdot \mathbf{x}$ denote $(a_i x_i)$. Define $H_k$: $\mathbb{R}^\infty \to \mathbb{R}$ by

$$H_k(\mathbf{x}) = \min\left(1, \sup_{q > k} \left|\sum_k^q x_i\right|\right).$$

Then for any random vector $\mathbf{V}$,

$$\sum V_i \text{ converges a.s. if and only if } EH_k(\mathbf{V}) \to 0. \tag{6.11}$$

Write $\theta$ for a generic element $(k, \mathbf{a}, \sigma)$ of $\mathbb{Z}^+ \times \ell_2' \times \mathscr{S}$. Define $g^\theta$: $\mathbb{R}^\infty \to \mathbb{R}$ by

$$g^\theta(\mathbf{x}) = H_k(\sigma(\mathbf{a} \cdot \mathbf{x})),$$

and observe that

$$|g^\theta(\mathbf{x}) - g^\theta(\mathbf{x}')| \leqq g^\theta(\mathbf{x} - \mathbf{x}'). \tag{6.12}$$

Using (6.11), the assertion of the Theorem becomes: there exists $\mathbf{m}$ such that

$$\lim_{k \to \infty} E g^{(k, \mathbf{a}, \sigma)}(\mathbf{Y}_m) = 0 \qquad \text{for each } (\mathbf{a}, \sigma) \in \ell_2' \times \mathscr{S}. \tag{6.13}$$

We shall deduce (6.13) from Proposition 13. Let

$$\mathscr{P}_0(\mathbb{R}) = \{\lambda : |\lambda|^2 < \infty, |\lambda|_1 = 0\}.$$

As in Corollary 7, the hypotheses ensure that $\mu \in \mathscr{P}_0(\mathbb{R})$ a.s. Define a metric on $\mathscr{P}_0(\mathbb{R})$, using the notation of (4.12), by

$$d(\lambda, \phi) = (E\,|F_\lambda^{-1}(T_1) - F_\phi^{-1}(T_1)|^2)^{1/2}. \tag{6.14}$$

It is easy to see that

$$d(\lambda_i, \lambda) \to 0 \text{ if and only if } \quad |\lambda_i|^2 \to |\lambda|^2 \quad \text{and} \quad \lambda_i \Rightarrow \lambda,$$

and thence that (5.14) and (5.15) are satisfied. Also, considering $\{g^\theta\}$ as functions on $\mathscr{P}_0(\mathbb{R}) \times \mathbb{R}^\infty$ in the obvious manner, it is easy to check (5.10), but (5.25) requires some estimation.

Fix $\phi \in \mathscr{P}_0(\mathbb{R})$, and let $\mathscr{L}(\mathbf{V}) = \phi^*$. Consider the quantity

$$h = E g^\theta(y_1, \ldots, y_j, V_{j+1}, V_{j+2}, \ldots). \tag{6.15}$$

From the definitions,

$$\begin{aligned} h &\leqq E \sup_{q>k} \left| \sum_{i=k}^{q} a_{\sigma(i)}\, y_{\sigma(i)}\, I(\sigma(i) \leqq j) + a_{\sigma(i)}\, V_{\sigma(i)}\, I(\sigma(i) > j) \right| \\ &\leqq \sum_1^j |a_i\, y_i| + \left( E \sup_{q>k} \left| \sum_{i=k}^{q} a_{\sigma(i)}\, V_{\sigma(i)}\, I(\sigma(i)>j) \right|^2 \right)^{1/2} \\ &\leqq \sum_1^j |y_i| + 2 \left( E \left( \sum_{i=k}^{\infty} a_{\sigma(i)}\, V_{\sigma(i)}\, I(\sigma(i)>j) \right)^2 \right)^{1/2} \end{aligned}$$

using a well-known inequality ([5] Theorem 9.5.4); hence

$$h \leqq \sum_1^j |y_i| + 2\,|\phi|_2 \left( \sum_k^\infty a_{\sigma(i)}^2 \right)^{1/2}. \tag{6.16}$$

In particular we deduce from (5.3) and Lemma 1 that

$$E g^\theta(\mathbf{Z}) \leqq 2K \left( \sum_k^\infty a_{\sigma(i)}^2 \right)^{1/2}$$

where $K = E\,|\mu(\omega)|_2 \leqq (E\,|\mu(\omega)|^2)^{1/2} < \infty$ by Lemma 2; so

$$\lim_{k \to \infty} E g^{(k,\mathbf{a},\sigma)}(\mathbf{Z}) = 0 \qquad \text{for each } (\mathbf{a}, \sigma) \in \ell_2' \times S. \tag{6.17}$$

In fact, (6.11) shows that (6.17) is equivalent to (3.25) and so follows quickly from (3.23). However, we do need (6.16) in order to verify condition (5.25) of Proposition 13. Let $\beta$ denote the quantity to be estimated in (5.25). Then by (5.23),

$$\begin{aligned} \beta &= |E g^\theta(y_1, \ldots, y_j, F_\lambda^{-1}(T_{j+1}), \ldots) - E g^\theta(y_1, \ldots, y_j, F_{\lambda'}^{-1}(T_{j+1}), \ldots)| \\ &\leqq E g^\theta(y_1 - y_1', \ldots, y_j - y_j', F_\lambda^{-1}(T_{j+1}) - F_{\lambda'}^{-1}(T_{j+1}), \ldots) \end{aligned} \tag{6.18}$$

by (6.12). Put

$$\phi = \mathscr{L}(F_\lambda^{-1}(T_1) - F_{\lambda'}^{-1}(T_1)).$$

Then (6.18) is of the form (6.15), and so by (6.16)

$$\beta \leqq \sum_1^j |y_i - y_i'| + 2|\phi|_2$$
$$\leqq \sum_1^j |y_i - y_i'| + 2d(\lambda, \lambda')$$

by (6.14). Hence (5.25) holds, and Proposition 13 may be applied to $\{g^\theta\}$. We deduce that the property $\mathfrak{Q}_j$ defined by

$$E g^\theta(\mathbf{Y}_n) \leqq E g^\theta(\mathbf{Z}) + 1/j \qquad \text{for each } \theta = (k, \mathbf{a}, \sigma)$$

holds for all $\mathbf{n}$ increasing sufficiently rapidly. So from (6.17), the property $\mathfrak{R}_j$ defined by

$$\limsup_{k \to \infty} E g^{(k, \mathbf{a}, \sigma)}(\mathbf{Y}_n) \leqq 1/j \qquad \text{for each } (\mathbf{a}, \sigma) \in \ell_2' \times \mathscr{S}$$

holds for all $\mathbf{n}$ increasing sufficiently rapidly. Applying Lemma 14, there exists $\mathbf{m}$ satisfying the following assertion.

For each $j$, there exists $\mathbf{m}^j$ such that:

$$\mathbf{m}^j \text{ has property } \mathfrak{R}_j; \tag{6.19a}$$

$$m_i^j = m_i \qquad \text{for all } i > j. \tag{6.19b}$$

But now $\mathbf{m}$ satisfies (6.13), because for each $(\mathbf{a}, \sigma)$ and each $j$,

$$\limsup_{k \to \infty} E g^{(k, \mathbf{a}, \sigma)}(\mathbf{Y}_m) = \limsup_{k \to \infty} E g^{(k, \mathbf{a}, \sigma)}(\mathbf{Y}_{m^j}) \qquad \text{by (6.19b)}$$
$$\leqq 1/j \qquad \text{by (6.19a)}.$$

## 7. A Counter-Example

Let $\beta_i$: $(0, 1) \to \{0, 1\}$ denote the $i$'th term of the nonterminating binary expansion

$$t = \sum 2^{-i} \beta_i(t).$$

Let $T$ be distributed uniformly on (0, 1). Define

$$X_i = (2\beta_i(T) - 1)\, T. \tag{7.1}$$

Suppose there exist $\mathbf{n}$ and an exchangeable $\mathbf{Z}$ such that

$$\sum |X_{n,i} - Z_i| < \infty \text{ a.s.} \tag{7.2}$$

Then

$$X_{n,i} - Z_i \to 0 \text{ a.s.} \quad \text{and in } L^1, \tag{7.3}$$

the latter because $|X_i| \leqq 1$ and $\{Z_i\}$ is identically distributed. Let $\mu$ be the canonical random measure for $\mathbf{Z}$. Applying statute $A_3$ of Section 3 to (3.8),

$$N^{-1} \sum_1^N \delta_{Z_i} \Rightarrow \mu(\omega) \text{ a.s.}$$

and then it follows easily from (7.3) that

$$N^{-1}\sum_{1}^{N}\delta_{X_{n,i}}\Rightarrow\mu(\omega)\ \text{a.s.}$$

From (7.1) and the strong law of large numbers for the i.i.d. sequence $\{\beta_i(T)\}$ we deduce

$$\mu(\omega)=1/2\,\delta_{T(\omega)}+1/2\,\delta_{-T(\omega)}\ \text{a.s.} \tag{7.4}$$

Now from (7.1) and (7.4) we see that $X_i$ is a.s. equal to some measurable function of $\mu$. And $\mu$ is a r.c.d. for $Z_i$ given $\mathscr{F}(\mu)$. So by Lemma 1,

$$\begin{aligned}E(|X_{n,i}-Z_i|\,|\,\mu)(\omega)&=1/2\,|X_{n,i}(\omega)-T(\omega)|+1/2\,|X_{n,i}(\omega)+T(\omega)|\ \text{a.s.}\\&\geqq T(\omega)\ \text{a.s.,}\end{aligned}$$

and so $E|X_{n,i}-Z_i|\geqq ET=1/2$, contradicting (7.3). This shows that $\{X_i\}$ does not satisfy $\mathfrak{P}$.

*Acknowledgements.* The author would like to thank Professor J.F.C. Kingman for raising the question of whether $\mathfrak{P}'$ was satisfied in general, and his research supervisors D.J.H. Garling and G.K. Eagleson for their encouragement.

## References

1. Berkes, I.: The law of the iterated logarithm for subsequences of random variables. Z. Wahrscheinlichkeitstheorie verw. Gebiete **30**, 209–215 (1974)
2. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
3. Chatterji, S.D.: A subsequence principle in probability theory. Bull. Amer. Math. Soc. **80**, 495–497 (1974)
4. Chatterji, S.D.: A principle of subsequences in probability theory. Advances in Math. **13**, 31–54 (1974)
5. Chung, K.L.: A course in probability theory. Second edition. New York-London: Academic Press 1974
6. Dacunha-Castelle, D.: Indiscernability and exchangeability in $L^p$-spaces. In: Proceedings of the seminar on random series, convex sets and geometry of Banach spaces, pp. 50–56. Aarhus Universitet Matematisk Institut 1975
7. Freedman, D.: Markov chains. San Francisco: Holden Day 1971
8. Gaposhkin, V.F.: Convergence and limit theorems for sequences of random variables. Theor. Probability Appl. **17**, 379–399 (1972)
9. Jagers, P.: Aspects of random measures and point processes. In: Advances in Probability 3. New York: Marcel Dekker 1974.
10. Kallenberg, O.: Canonical representations and convergence criteria for processes with interchangeable increments. Z. Wahrscheinlichkeitstheorie verw. Gebiete **27**, 23–36 (1973)
11. Komlós, J.: A generalisation of a problem of Steinhaus. Acta Math. Acad. Sci. Hungar. **18**, 217–229 (1967)
12. Meyer, P.A.: Probability and potentials. Waltham, Mass.: Blaisdell 1966
13. Parthasarathy, K.R.: Probability measures on metric spaces. New York-London: Academic Press 1967
14. Révész, P.: The laws of large numbers. New York-London: Academic Press 1967

*Received November 22, 1976*

**Note Added in Proof**

A characterisation of sequences satisfying $\mathfrak{P}'$ has been obtained by Berkes and Rosenthal, "Almost exchangeable sequences of random variables", to appear.