The Annals of Probability
1981, Vol. 9, No. 3, 405–412

A LAW OF LARGE NUMBERS FOR IDENTICALLY DISTRIBUTED MARTINGALE DIFFERENCES[1]

BY JOHN ELTON

University of Texas

The averages of an identically distributed martingale difference sequence converge in mean to zero, but the almost sure convergence of the averages characterizes $L \log L$ in the following sense: if the terms of an identically distributed martingale difference sequence are in $L \log L$, the averages converge to zero almost surely; but if f is any integrable random variable with zero expectation which is not in $L \log L$, there is a martingale difference sequence whose terms have the same distribution as f and whose averages diverge almost surely. The maximal function of the averages of an identically distributed martingale difference sequence is integrable if its terms are in $L \log L$; the converse is false.

1. Introduction. Let $(\Omega, \mathscr{F}, P)$ be a probability space, and $(\mathscr{F}_n)$ an increasing sequence of sub-σ-algebras of $\mathscr{F}$. For $f \in L^1(=L^1(P))$, we use $E_n(f)$ to denote $E(f \mid \mathscr{F}_n)$, the conditional expectation of f given $\mathscr{F}_n$. A sequence $f_n \in L^1(\mathscr{F}_n)$ will be called a *martingale difference sequence* (mds) if $E_n(f_{n+1}) = 0$, $n \in N$ (N is the set of positive integers); in other words, the sequence $s_n = \sum_{k=1}^n f_k$ of partial sums is a martingale.

If the f_n are independent and *identically distributed* (id), the sequence $a_n = (1/n)s_n$ of averages converges almost surely (strong law of large numbers) and in L^1-mean to zero (see, e.g., Chow and Teicher (1978), page 131). In Section 2, we show that $a_n \to_{L^1} 0$ without the hypothesis of independence. In Section 3, we show that $a_n \to 0$ almost surely without the hypothesis of independence if we require that $f_1 \in L \log L$, where

$$L \log L = \{f \in L^1 : E(|f| \log^+ |f|) < \infty\}.$$

In Section 4, we show that if $f \in L^1$ with $E(f) = 0$ but $f \notin L \log L$, we can construct an id mds (f_n) with f_1 having the same distribution as f such that (a_n) diverges almost surely. This is our main result in this article. In Section 5, we show that the maximal function of the averages

$$M(\omega) = \sup_n (1/n) \left| \sum_{k=1}^n f_k(\omega) \right|$$

is in L^1 if $f_1 \in L \log L$, which generalizes a result of Marcinkiewicz and Zygmund (1937) for the independent case. However, unlike the independent case, the converse is false. In fact, if f is any symmetric random variable in L^1, there is an id mds (f_n) with f_1 having the same distribution as f such that $M \in L^1$. This is probably true without the hypothesis of symmetry, but we don't know how to show it in general, for f having mean zero.

We introduce some notation. If g is a real-valued function and $c \geq 0$, define

$$\begin{aligned} {}^c g(x) &= g(x) \qquad \text{if } |g(x)| \leq c, \\ &= 0 \qquad \text{otherwise}; \\ {}^{\sim c} g(x) &= g(x) - {}^c g(x). \end{aligned}$$

Received April 23, 1979; revised October, 1979.

[1] Research partially supported by the National Science Foundation Grant No. MCS 8002393.

AMS 1970 *subject classifications*. Primary 60F15, 60G45.

Key words and phrases. Martingale, law of large numbers, maximal function, almost sure convergence.

With a sequence $f_n \in L^1(\mathscr{F}_n)$ we associate the sequence

$$d_n = {}^n f_n - E_{n-1}({}^n f_n)$$

which is a mds whose terms are in L^2.

2. Convergence of the averages to zero in mean.

THEOREM 1. *If* (f_n) *is an* id mds, *then* $E(|a_n|) \to 0$.

PROOF. Write

$$\begin{aligned} f_n &= {}^n f_n + {}^{\sim n} f_n = d_n + E_{n-1}({}^n f_n) + {}^{\sim n} f_n \\ &= d_n + {}^{\sim n} f_n - E_{n-1}({}^{\sim n} f_n), \end{aligned}$$

observing that $E_{n-1}({}^{\sim n} f_n) = -E_{n-1}({}^n f_n)$ since $E_{n-1}(f_n) = 0$. Thus

$$\begin{aligned} E(|a_n|) &\le E(|(1/n)\textstyle\sum_{k=1}^n d_k|) + (1/n)\sum_{k=1}^n E(|{}^{\sim k} f_k| + |E_{k-1}({}^{\sim k} f_k)|) \\ &\le (1/n)\,(\textstyle\sum_{k=1}^n E(d_k^2))^{1/2} + (2/n)\sum_{k=1}^n E(|{}^{\sim k} f_1|), \end{aligned}$$

since the d_k are orthogonal elements of L^2, E_{k-1} is a contraction on L^1, and the f_k are id.

Now $E(|{}^{\sim n} f_1|) \to 0$ as $n \to \infty$ since f_1 is integrable, so the averages $(1/n)\sum_{k=1}^n E(|{}^{\sim k} f_1|) \to 0$ also.

Next, Lemma 1, which follows, shows that

$$\textstyle\sum_{n=1}^\infty (1/n^2) E(d_n^2) < \infty,$$

so $(1/n^2)\sum_{k=1}^n E(d_k^2) \to 0$ as $n \to \infty$ by Kronecker's lemma, and the proof is complete.

LEMMA 1. *If* (f_n) *is an* id *sequence with* $f_1 \in L^1$, *then there is* $K < \infty$ *such that*

$$\textstyle\sum_{n=1}^\infty (1/n^2) E(d_n^2) < KE(|f_1|).$$

PROOF. $I - E_{n-1}$ is a contraction on L^2, so

$$E(d_n^2) = E(({}^n f_n - E_{n-1}({}^n f_n))^2) \le E(({}^n f_n)^2).$$

The rest of the proof is the same as in the classical proof of Kolomogorov's strong law of large numbers:

$$\textstyle\sum_{n=1}^\infty (1/n^2) E(({}^n f_n)^2) = \sum_{m=1}^\infty E(f_1^2 \chi_{\{m-1<|f_1|\le m\}}) \sum_{n=m}^\infty (1/n^2) \le KE(|f_1|),$$

where $K < \infty$ is such that $\sum_{n=m}^\infty (1/n^2) \le K/m$ for all $m \in \mathbb{N}$.

3. Convergence of the averages almost surely to zero when f_1 is in $L \log L$.

LEMMA 2. *Let* $f \in L^1$, $f \ge 0$. *Then*

$$f \in L \log L \text{ iff } \textstyle\sum_{n=1}^\infty (1/n) E(f \chi_{\{f>n\}}) < \infty.$$

PROOF.

$$\begin{aligned} E(f \log^+ f) &\ge \textstyle\sum_{n=1}^\infty (\log n) E(f \chi_{\{n<f\le n+1\}}) \ge \sum_{n=2}^\infty (\sum_{j=2}^n (1/j)) E(f \chi_{\{n<f\le n+1\}}) \\ &= \textstyle\sum_{j=2}^\infty (1/j) \sum_{n=j}^\infty E(f \chi_{\{n<f\le n+1\}}) = \sum_{j=2}^\infty (1/j) E(f \chi_{\{f>j\}}). \end{aligned}$$

The other direction follows similarly.

THEOREM 2. *Let (f_n) be an* id mds *with* $f_1 \in L \log L$. *Then* $a_n \to 0$ *almost surely.*

PROOF. Write

$$f_n = d_n + {}^{\sim n}f_n - E_{n-1}({}^{\sim n}f_n)$$

as in the proof of Theorem 1. Then

$$\begin{aligned} E(|\textstyle\sum_{k=1}^n (1/k)f_k|) &\le E(|\textstyle\sum_{k=1}^n (1/k)d_k|) + 2\textstyle\sum_{k=1}^n (1/k)E(|{}^{\sim k}k_k|) \\ &\le (\textstyle\sum_{k=1}^n (1/k^2)E(d_k^2))^{1/2} + 2\textstyle\sum_{k=1}^n (1/k)E(|f_1|\chi_{\{k<|f_1|\}}). \end{aligned}$$

By Lemmas 1 and 2, this is bounded. So the martingale $\sum_{k=1}^n (1/k)f_k$ converges almost surely. Thus by Kronecker's lemma, $a_n = (1/n)\sum_{k=1}^n f_k \to 0$ almost surely.

4. Existence of an id mds with averages diverging almost surely when f_1 is not in $L \log L$.

LEMMA 3. *Let* $f_n \in L^1(\mathscr{F}_n)$ *be an* id *sequence and* $n_0 \in \mathbb{N}$. *Then for almost all* $\omega \in \Omega$, $\lim_{n\to\infty} (1/n)\sum_{k=1}^n f_k(\omega)$ *exists* iff $\lim_{n\to\infty} (1/n)\sum_{k=1}^n E_{k-1}({}^{n_0+k}f_k)(\omega)$ *exists (and the limits are equal if they exist).*

PROOF. Let $d_n = {}^{n_0+n}f_n - E_{n-1}({}^{n_0+n}f_n)$, $n \in \mathbb{N}$ (a slight change from the way d_n was defined before). There exists $C < \infty$ such that

$$E((\textstyle\sum_{k=1}^n (1/k)d_k)^2) = \textstyle\sum_{k=1}^n (1/k^2)E(d_k^2) \le CE(|f_1|)$$

just as in Lemma 1. So the martingale $\sum_{k=1}^n (1/k)d_k$ converges almost surely, so $(1/n)\sum_{k=1}^n d_k \to 0$ almost surely by Kronoecker's lemma.

Now $P(|f_n| > n + n_0$ infinitely often$) = 0$ by the Borel-Cantelli lemma, since

$$\textstyle\sum_{n=1}^\infty P(|f_n| > n + n_0) \le E(|f_1|).$$

So $\lim_{n\to\infty} (1/n)\sum_{k=1}^n f_k(\omega)$ exists iff $\lim_{n\to\infty} (1/n)\sum_{k=1}^n {}^{n_0+k}f_k(\omega)$ exists (and they are equal when they do), for almost all $\omega \in \Omega$.

LEMMA 4. *Let (a_n) be a nonincreasing sequence of positive numbers. Then*

$$\textstyle\sum_{n=1}^\infty (1/n)a_n = \infty \text{ iff } \sum_{n=1}^\infty a_{k^n} = \infty \qquad \textit{for all} \quad k \ge 2, k \in \mathbb{N}.$$

PROOF. This is a version of the Cauchy condensation test for series.

Construction of the example. Let $f \in L^1$ with $E(f) = 0$ but $f \notin L \log L$. Without loss of generality, we may assume f is a nondecreasing function on $(0, 1)$ and P is Lebesgue measure (the function $g(x) = \sup\{t: F(t) \le x\}$, $x \in (0, 1)$, where F is the distribution function of f, gives such a function with the same distribution as f).

Define intervals $I_i^n \subset (0, 1)$, $i = -3, -2, -1, 0, 1, 2, 3$, $n \in \mathbb{N}$, which partition $(0, 1)$ for each n, by

$$I_{-3}^n = (f < -(n_0 + n))$$

$$I_{-2}^n = (-(n_0 + n) \le f < -n_0)$$

$$I_{-1}^n = I_{-1} = (-n_0 \le f < 0)$$

$$I_0^n = I_0 = (f = 0)$$

$$I_1^n = I_1 = (0 < f \le n_0)$$

$$I_2^n = (n_0 < f \le n_0 + n)$$

$$I_3^n = (n_0 + n < f).$$

Let $\beta_i^n = \int_{n} |f|\, dP$, $n \in \mathbb{N}$, $i = -3, \cdots, 3$; and $\beta_i = \beta_i^n$ for $i = \pm 1$. Since $E(f) = 0$,

$$\beta_{-3}^n + \beta_{-2}^n + \beta_{-1}^n = \beta_1^n + \beta_2^n + \beta_3^n. \tag{1}$$

So we can (and do) choose n_0 so large that

$$\beta_{-1} \geq \beta_2^n + \beta_3^n \qquad \text{and} \qquad \beta_1 \geq \beta_{-2}^n + \beta_{-3}^n.$$

For convenience let $S = \{-3, -2, 0, 1, 2, 3\}$ (note the absence of -1). Define p_i^n, $i \in S$, $n \in \mathbb{N}$, by

$$p_{\varepsilon i}^n = P(I_{\varepsilon i}^n) + (\beta_{\varepsilon i}^n/\beta_{-\varepsilon})P(I_{-\varepsilon}^n), \qquad i = 2, 3; \quad \varepsilon = \pm 1.$$

(2) $$p_1 = p_1^n = (1 - [\beta_2^n + \beta_3^n)/\beta_{-1})P(I_{-1}) + (1 - (\beta_{-2}^n + \beta_{-3}^n)/\beta_1)P(I_1).$$

$$p_0 = p_0^n = P(I_0).$$

Observe that $\sum_{i \in S} p_i^n = 1$, and also that $p_{-3}^n \downarrow 0$, $p_3^n \downarrow 0$.

Next define intervals J_{ij}^n, $i \in S$, $n \in \mathbb{N}$, $j = 1, 2$, when $p_i^n \neq 0$, by

$$J_{i1}^n = (0, (1/p_i^n)P(I_i^n)), \qquad i = \pm 2, \pm 3;$$

(3) $$J_{11}^n = (0, (1/p_1)(1 - (\beta_2^n + \beta_3^n)/\beta_{-1})P(I_{-1}));$$

$$J_{01}^n = (0, 1);$$

$$J_{i2}^n = (0, 1) - G_{i1}^n, \qquad i \in S.$$

Then define functions φ_i^n on $(0, 1)$, $i \in S$, such that φ_i^n is identity if $p_i^n = 0$, and if $p_i^n \neq 0$,

$$\begin{aligned}
\varphi_{\epsilon i}^n(J_{\epsilon i,1}^n) &= I_{\epsilon i}^n && \text{except possibly for endpoints of the interval,}\\
\varphi_{\epsilon i}^n(J_{\epsilon i,2}^n) &= I_{-\epsilon} && i = 2, 3;\ \epsilon = \pm 1;\\
\varphi_1^n(J_{11}^n) &= I_{-1} && \text{except for endpoints;}\\
\varphi_1^n(J_{12}^n) &= I_1 && \text{except for endpoints;}\\
\varphi_0^n(J_{01}^n) &= I_0 && \text{except for endpoints;}
\end{aligned}$$

and φ_i^n is linear and increasing on J_{ij}^n. Observe (using (2) and (3)) that

$$P(J_{\epsilon i,1}^n) = (1/p_{\epsilon i}^n)P(I_{\epsilon i}^n),$$

$$P(J_{\epsilon i,2}^n) = (1/p_{\epsilon i}^n)(\beta_{\epsilon i}^n/\beta_{-\epsilon})P(I_{-\epsilon}), \qquad i = 2, 3;\ \epsilon = \pm 1.$$

$$P(J_{11}^n) = (1/p_1)(1 - (\beta_2^n + \beta_3^n)/\beta_{-1})P(I_{-1}),$$

$$P(J_{12}^n) = (1/p_1)(1 - (\beta_{-2}^n + \beta_{-3}^n)/\beta_1)P(I_1),$$

whenever these are defined. So we have for $i = 2, 3$; $\epsilon = \pm 1$, that

$$\begin{aligned}
\int_0^1 f(\varphi_{\epsilon i}^n(x))\, dx &= \int_{J_{\epsilon i,1}^n} f(\varphi_{\epsilon i}^n(x))\, dx + \int_{J_{\epsilon i,2}^n} f(\varphi_{\epsilon i}^n(x))\, dx\\
&= 1/p_{\epsilon i}^n \int_{I_{\epsilon i}^n} f(x)\, dx + (1/p_{\epsilon i}^n)(\beta_{\epsilon i}^n/\beta_{-\epsilon}) \int_{I_{-\epsilon}} f(x)\, dx\\
&= (1/p_{\epsilon i}^n)(\epsilon\beta_{\epsilon i}^n + (\beta_{\epsilon i}^n/\beta_{-\epsilon})(-\epsilon)\beta_{-\epsilon}) = 0.
\end{aligned}$$

Similarly,

$$\int_0^1 f(\varphi_i^n(x))\, dx = 0 \qquad \text{for all} \quad i \in S \tag{4}$$

(for $i = 1$, use (1)).

We also have for $\epsilon = \pm 1, P^n_{\epsilon 3} \neq 0$,

$$(5)\quad \int_0^1 {}^{n_0+n}f(\varphi^n_{\epsilon 3}(x))\,dx$$
$$= (1/p^n_{\epsilon 3}) \int_{I^n_{\epsilon 3}} {}^{n_0+n}f(x)\,dx + (1/p^n_{\epsilon 3})(\beta^n_{\epsilon 3}/\beta_{-\epsilon}) \int_{I_{-\epsilon}} {}^{n_0+n}f(x)\,dx = (-\epsilon)(\beta^n_{\epsilon 3}/p^n_{\epsilon 3}),$$

but

$$(6)\qquad \int_0^1 {}^{n_0+n}f(\varphi^n_i(x))\,dx = 0 \qquad \text{for} \quad i = 0, 1, \pm 2.$$

Now either $f^+ \notin L \log L$ or $f^- \notin L \log L$, so without loss of generality, assume

$$(7)\qquad f^- \notin L \log L.$$

Let $b = \beta_1/(1 + \beta_1)$ and

$$B = \max\{\beta_1/P(I_1), \beta_{-1}/P(I_{-1})\}.$$

Observe that $0 < b < B$. We have for $n \in \mathbb{N}$,

$$P(I^n_{-3}) \le \beta^n_{-3}, \quad \text{so } p^n_{-3} \le \beta^n_{-3} + (\beta^n_{-3}/\beta_1)P(I_1) \le (1 + (1/\beta_1))\beta^n_{-3},$$

so

$$(8)\qquad \beta^n_{-3}/p^n_{-3} \ge b.$$

Also, $p^n_{-3} \ge (\beta^n_{-3}/\beta_1)P(I_1)$, so

$$(9)\qquad \beta^n_{-3}/p^n_{-3} \le B; \qquad \text{and similarly,} \qquad \beta^n_3/p^n_3 \le B.$$

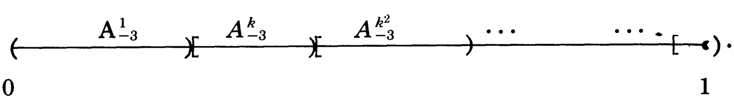
Choose $k \in \mathbb{N}$ so large that

$$(10)\qquad k > 2B/b + 1.$$

For $l = 0, 1, 2, \cdots$ and $k^{l-1} < n \le k^l$, let

$$A^n_{-3} = (0, 1) \cap \operatorname{mod}([\textstyle\sum_{i=0}^{l-1} p^{k^i}_{-3}, \sum_{i=0}^{l-1} p^{k^i}_{-3} + p^n_{-3})),$$

where $\operatorname{mod} x = x -$ greatest integer in x, $x \in \mathbb{R}$.



The motivation will follow. Note that $P(A^n_{-3}) = p^n_{-3}$ for all $n \in \mathbb{N}$. Then define $A^n_i \subset (0, 1)$ for $i = 3, \pm 2, 1, 0$, $n \in \mathbb{N}$ so that $\{A^n_i: i \in S\}$ is a partition of (0, 1) for each n, and $P(A^n_i) = p^n_i$.

We are finally ready to define our mds (f_n). Let

$$\Omega = \textstyle\prod_{i=0}^{\infty} (0, 1)_i, \qquad \text{where} \quad (0, 1)_i = (0, 1) \qquad \text{for all } i,$$

and let μ be product Lebesgue measure on Ω. For $\omega \in \Omega$, we write $\omega = (\omega_0, \omega_1, \cdots)$. Let

$$f_n(\omega) = \textstyle\sum_{i \in S} f(\varphi^n_i(\omega_n))\chi_{A^n_i}(\omega_0), \qquad \omega \in \Omega, n \in \mathbb{N}.$$

Illustration.

$(0,1)_0$		
	$f_n(\omega) = f(\varphi_3^n(\omega_n))$	A_3^n
	$f_n(\omega) = f(\varphi_2^n(\omega_n))$	A_2^n
	$f_n(\omega) = f(\varphi_1^n(\omega_n))$	A_1^n
	$f_n(\omega) = 0$	A_0^n
	$f_n(\omega) = f(\varphi_{-2}^n(\omega_n))$	A_{-2}^n
	$f_n(\omega) = f(\varphi_{-3}^n(\omega_n))$	A_{-3}^n
	$(0,1)_n$	

Let

$$\mathscr{F}_n = \{C \times \textstyle\prod_{i=n+1}^{\infty} (0,1)_i \colon C \quad \text{a Borel set in} \quad \prod_{i=0}^{n} (0,1)_i\},$$

and $\mathscr{F}$ be the Borel sets in Ω. So $(\mathscr{F}_n)$ is an increasing sequence of sub-σ-algebras of $\mathscr{F}$, and f_n is $\mathscr{F}_n$-measurable for all n.

LEMMA 5. *f_n has the same distribution as f for all $n \in \mathbb{N}$, and (f_n) is an* mds.

PROOF. If C is a Borel set in $\mathbb{R}$,

$$\begin{aligned}\mu(f_n \in C) &= \int_\Omega \chi_C(f_n(\omega))\,d\omega = \int_0^1 \left(\int_0^1 \chi_C(f_n(\omega))\,d\omega_n\right) d\omega_0 \\ &= \textstyle\sum_{i\in S} P(A_i^n) \displaystyle\int_0^1 \chi_C(f(\varphi_i^n(\omega_n))\,d\omega_n = \textstyle\sum_{i\in S; j=1,2} P(A_i) \displaystyle\int_{J_{ij}^n} \chi_C(f(\varphi_i^n(\omega_n))\,d\omega_n \\ &= \textstyle\sum_{i=2,3;\epsilon=\pm1} p_{\epsilon i}^n[(1/p_{\epsilon i}^n)P(\{f\in C\}\cap I_{\epsilon i}^n) + (1/p_{\epsilon i}^n)(\beta_{\epsilon i}^n/\beta_{-\epsilon})P(\{f\in C\}\cap I_{-\epsilon})] \\ &\quad + p_1[(1/p_1)(1-(\beta_2^n+\beta_3^n)/\beta_{-1})P(\{f\in C\}\cap I_{-1}) \\ &\quad + (1/p_1)(1-(\beta_{-2}^n+\beta_{-3}^n)/\beta_1)P(\{f\in C\}\cap I_1)] + P(\{f\in C\}\cap I_0) \\ &= \textstyle\sum_{i\in S} P(\{f\in C\}\cap I_i^n) = P(f\in C),\end{aligned}$$

which verifies the first part. Next, let $\tilde{C} = C \times \prod_{i=n}^{\infty} (0,1)_i \in \mathscr{F}_{n-1}$, where C is a Borel set in $\prod_{i=0}^{n-1} (0,1)_i$. Then

$$\begin{aligned}\int_{\tilde{C}} f_n(\omega)\,d\omega &= \textstyle\sum_{i\in S} \displaystyle\int_0^1 \cdots \int_0^1 \cdots \int_0^1 \chi_C(\omega_0,\cdots,\omega_{n-1})\chi_{A_i^n}(\omega_0) f(\varphi_i^n(\omega_n))\,d\omega_n \cdots d\omega_0 \\ &= 0, \qquad \text{by (4) above.}\end{aligned}$$

So $E_{n-1}(f_n) = 0$ for all n.

LEMMA 6. *The averages of the sequence $E_{n-1}({}^{n_0+n}f_n)$ diverge almost surely.*

PROOF. Let $\tilde{A}_i^n = \{\omega \in \Omega \colon \omega_0 \in A_i^n\}$, $i \in S$, $n \in \mathbb{N}$.
Let $\tilde{C} = C \times \prod_{i=n}^{\infty} (0,1)_i \in \mathscr{F}_{n-1}$ as above. Then by (5) and (6),

$$\int_{\tilde{C}} {}^{n_0+n}f_n(\omega)\,d\omega = \textstyle\sum_{i\in S} \displaystyle\int_0^1 \cdots \int_0^1 \chi_C(\omega_0,\cdots,\omega_{n-1})\chi_{A_i^n}(\omega_0)^{n_0+n} f(\varphi_i^n(\omega_n))\,d\omega_n \cdots d\omega_0$$

$$= \int_0^1 \cdots \int_0^1 \chi_C(\omega_0, \cdots, \omega_{n-1})(\textstyle\sum_{\epsilon=\pm1}(-\epsilon)(\beta^n_{\epsilon 3}/p^n_{\epsilon 3})\chi_{A^n_{\epsilon 3}}(\omega_0))\, d\omega_{n-1} \cdots d\omega_0$$

$$= \int_{\tilde{C}} \textstyle\sum_{\epsilon=\pm1}(-\epsilon)(\beta^n_{\epsilon 3}/p^n_{\epsilon 3})\chi_{\tilde{A}^n_{\epsilon 3}}(\omega)\, d\omega.$$

So

$$E_{n-1}({}^{n_0+n}f_n) = \textstyle\sum_{\epsilon=\pm1}(-\epsilon)(\beta^n_{\epsilon 3}/p^n_{\epsilon 3})\chi_{\tilde{A}^n_{\epsilon 3}}.$$

We assumed in (7) that f^- is not in $L \log L$, so we have by Lemma 2 that

$$\textstyle\sum_{n=1}^{\infty} (1/(n_0 + n))\beta^n_{-3} = \infty,$$

and so

$$\textstyle\sum_{n=0}^{\infty} \beta^{k^n}_{-3} = \infty$$

$$\textstyle\sum_{n=0}^{\infty} p^{k^n}_{-3} = \infty \qquad \text{also.}$$

Observe that if $\omega \in \tilde{A}^{k^l}_{-3}$, then $\omega \in \tilde{A}^n_{-3}$ for $k^{l-1} < n \le k^l$ (recall $p^n_{-3} \downarrow$), so using (8) and (9),

$$(1/k^l)\textstyle\sum_{n=1}^{k^l} E_{n-1}({}^{n_0+n}f_n)(\omega) \ge (1/k^l)[(k^l - k^{l-1})b - k^{l-1}B] = ((k-1)/k)b - (1/k)B.$$

But if $\omega \notin \tilde{A}^{k^{l-1}+1}_{-3}$, then $\omega \notin \tilde{A}^n_{-3}$ for $k^{l-1} < n \le k^l$, so

$$(1/k^l) \textstyle\sum_{n=1}^{k^l} E_{n-1}({}^{n_0+n}f_n)(\omega) \le (k^{l-1}/k^l)B = (1/k)B.$$

Since $\sum_{n=0}^{\infty} p^{k^n}_{-3} = \infty$, we have that for each $\omega \in \Omega$, $\omega \in \tilde{A}^{k^l}_{-3}$ occurs for *infinitely* many l (just note that the intervals $[\sum_{i=0}^{l-1} p^{k^i}_{-3}, \sum_{i=0}^{l} p^{k^i}_{-3})$ are adjoining and cover all of R^+, so applying the mod function, we see that the $\tilde{A}^{k^l}_{-3}$ cover (0, 1) infinitely many times). Similarly, for each $\omega \in \Omega$, $\omega \notin \tilde{A}^{k^{l-1}+1}_{-3}$ occurs for infinitely many l (note that if $\omega \in \tilde{A}^{k^{l-1}}_{-3}$, then $\omega \notin \tilde{A}^{k^{l-1}+1}_{-3}$ if $p^{k^{l-1}}_{-3} + p^{k^{l-1}+1}_{-3} \le 1$, which holds for large enough l). Hence we have $\lim \sup_N (1/N) \sum_{n=1}^{N} E_{n-1}({}^{n_0+n}f_n)(\omega) \ge ((k-1)/k)b - (1/k)B$ almost surely, and

$$\lim \inf_N (1/N) \textstyle\sum_{n=1}^{N} E_{n-1}\,({}^{n_0+n}f_n)(\omega) \le (1/k)B \quad \text{a.s.}$$

But $((k-1)/k)b - (1/k)B > (1/k)B$ by our choice in (10) of k, so the proof of the lemma is complete.

THEOREM 3. *If $f \in L^1$ with $Ef = 0$ and $f_1 \notin L \log L$, there is an* id mds (f_n) *with f_1 having the same distribution as f such that the averages of (f_n) diverge almost surely.*

PROOF. This follows from Lemmas 3 and 6.

5. Integrability of the maximal function.

LEMMA 7. *If (a_n) is any sequence of real numbers, then*

$$\sup_n (1/n) | \textstyle\sum_{k=1}^{n} a_k | \le 2 \sup_n | \sum_{k=1}^{n} (1/k)a_k |.$$

PROOF.

$$\begin{aligned} |(1/n) \textstyle\sum_{k=1}^{n} a_k | &= | \textstyle\sum_{k=1}^{n} (1/k)a_k(1-(n-k)/n)| \\ &= |\textstyle\sum_{k=1}^{n} (1/k)a_k - (1/n)\sum_{j=1}^{n-1} \sum_{k=1}^{n-j} (1/k)a_k | \\ &\le (1 + (n-1)/n) \sup_m | \textstyle\sum_{k=1}^{m} (1/k)a_k |, \end{aligned}$$

for all n.

REMARK. This is observed in Marcinkiewicz and Zygmund (1937).

THEOREM 4. *If* (f_n) *is an* id mds *with* $f_1 \in L \log L$, *then* $M \in L^1$.

PROOF. Write

$$f_n = d_n + {}^{\sim n}f_n - E_{n-1}({}^{\sim n}f_n),$$

as in Section 2. Then

$$\begin{aligned} M &= \sup_n (1/n) \left| \textstyle\sum_{k=1}^n f_k \right| \\ &\le \sup_n (1/n) \left| \textstyle\sum_{k=1}^n d_k \right| + \sup_n (1/n) \textstyle\sum_{k=1}^n \left| {}^{\sim k}f_k - E_{k-1}({}^{\sim k}f_k) \right| \\ &\le 2 \sup_n \left| \textstyle\sum_{k=1}^n (1/k) d_k \right| + 2 \sup_n \textstyle\sum_{k=1}^n (1/k) \left| {}^{\sim k}f_k - E_{k-1}({}^{\sim k}f_k) \right|, \end{aligned}$$

by Lemma 7. By an inequality of B. Davis (1970), there is a constant $B < \infty$ such that

$$\begin{aligned} E(\sup_n \left| \textstyle\sum_{k=1}^n (1/k) d_k \right|) &\le BE((\textstyle\sum_{k=1}^\infty (1/k^2) d_k^2)^{1/2}) \\ &\le B(E(\textstyle\sum_{k=1}^\infty (1/k^2) d_k^2))^{1/2} \le B(KE(|f_1|))^{1/2}, \end{aligned}$$

using Lemma 1 for the last step. And

$$E(\sup_n \textstyle\sum_{k=1}^n (1/k) \left| {}^{\sim k}f_k - E_{k-1}({}^{\sim k}f_k) \right|) < \infty$$

since $f_1 \in L \log L$, just as in the proof of Theorem 2.

PROPOSITION. *If* $f \in L^1$ *and* f *is symmetric, then there is an* id mds (f_n), *with* f_1 *having the same distribution as* f, *such that* $M \in L^1$.

PROOF. Let (r_n) be a sequence of independent random variables on [0, 1] for which $m(r_n = 1) = m(r_n = -1) = ½$, where m is Lebesgue measure. Since f is symmetric, the functions

$$f_n = |f| \otimes r_n \text{ on } \Omega \times [0, 1]$$

have the same distribution as f. And $M(\omega, t) = \sup_n (1/n) \left| \sum_{k=1}^n |f(\omega)| r_k(t) \right| \le |f(\omega)|$, so $M \in L^1$. It is easy to see that (f_n) is a mds with respect to the σ-algebras $\mathscr{F} \times \mathscr{D}_n$ where $\mathscr{D}_n$ is the σ-algebra generated in [0, 1] by $\{r_1, \cdots, r_n\}$, since the r_n are independent with mean 0.

REMARK. A similar method works if f is not too asymmetric, but we don't know a method which will work for arbitrary $f \in L^1$.

Acknowledgments. We wish to thank Professor Haskell Rosenthal for suggesting the problem of the law of large numbers for an id mds. It was his insight that a counterexample to the strong law should be possible by consideration of $E_{n-1}({}^{n}f_n)$. We also thank Professor Don Burkholder for suggesting the problem of the integrability of the maximal function.

REFERENCES

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory*: *Independence, Interchangeability, Martingales.* Springer, New York.

DAVIS, B. (1970). On the integrability of the martingale square function. *Israel J. Math.* **8** 187–190.

MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Sur les functions indépendantes. *Fund. Math.* **29** 60–90.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712