SOME RESULTS GIVING RATES OF CONVERGENCE IN THE LAW OF LARGE NUMBERS FOR WEIGHTED SUMS OF INDEPENDENT RANDOM VARIABLES(1)

BY

W. E. FRANCK AND D. L. HANSON(2)

1. **Introduction and summary.** Let X_N for $N = 1,2,\cdots$ be an independent sequence of random variables with finite first absolute moments; let $a_{N,k}$ for $N, k = 1,2,\cdots$ be real numbers; let

$$A_N = \frac{1}{N} \sum_{k=1}^{N} (X_k - EX_k);$$

and let

$$S_N = \sum_{k=1}^{\infty} a_{N,k}(X_k - EX_k) .$$

A great deal of effort has gone into the investigation of the convergence of the sequence A_N and consequently into the convergence of the sequence $P\{|A_n| > \varepsilon\}$. This paper deals with the convergence of the sequence $P\{|S_N| > \varepsilon\}$.

The intuitive idea behind this work is that at least as much averaging should occur in the sum $\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$ as occurs in the sum $\frac{1}{2}X_1 + \frac{1}{2}X_2$ and thus if one requires $\sum_k |a_{N,k}| \leqq 1$ and $\max_k |a_{N,k}| \leqq 1/N$ one should get at least as rapid convergence of $P(|S_N| > \varepsilon)$ to zero as one gets for $P(|A_N| > \varepsilon)$. One might hope to obtain sharper bounds on the rate of convergence of $P\{|S_N| > \varepsilon\}$ to zero if he could in some way quantitatively measure how much averaging occurs in the sum S_N.

In [1] an exponential rate of convergence of $P(|S_N| > \varepsilon)$ to zero was obtained when moment generating functions of the X_k's exist; this extends the well known result (see [2]) giving an exponential convergence rate of $P(|A_N| > \varepsilon)$ to zero. Jamison, Orey, and Pruitt [3] and Pruitt [4] deal with the convergence (in probability and almost everywhere) of S_N to zero without worrying about specific

Received by the editors January 7, 1966.

(1) This research was reported in an announcement (bearing the same title) which appeared in Bull. Amer. Math. Soc. 72 (1966), 266–268.

(2) Research supported by the Air Force Office of Scientific Research.

rates of convergence. The only other results (of which the authors are aware) dealing with $a_{N,k}$'s more general than $a_{N,k} = 1/N$ for $k = 1, \cdots, N$ and zero otherwise are the results obtained in [5], and results in ergodic theory (see [6] for an example) which do not give convergence rates.

The results obtained here are in the spirit of Theorems 1 and 2a of [7], parts of Theorems 1 and 3 of [8], and parts of Theorems 1, 3 and 4 of [9]. In §2 the main theorems are stated; they are proved in §3. The relations between these results and previously published results are examined in §4. §5 consists of miscellaneous concluding remarks.

2. **Statement of results.** Throughout the remainder of this paper C will denote various positive constants whose exact values do not matter. For example, the expression $1 + C \leqq C$ is a valid inequality using this notation.

Let X_N, $a_{N,k}$, and S_N be as defined in §1. Suppose C, α, β, γ, and t are constants such that

$$\sum_k |a_{N,k}| \leqq CN^{\alpha}; \tag{1}$$

$$\max_k |a_{N,k}| \leqq CN^{-\beta}, \tag{2}$$

$$\sum_k |a_{N,k}|^t \leqq CN^{-\rho}. \tag{3}$$

Notice that if $t \geqq 1$ then

$$\sum_k |a_{N,k}|^t \leqq [\max_k |a_{N,k}|]^{t-1} \sum_k |a_{N,k}| \leqq CN^{-[\beta(t-1)-\alpha]}$$

so that we may assume

$$\rho \geqq \beta(t-1) - \alpha. \tag{4}$$

By noting that $\sum_k |a_{N,k}|^\lambda \geqq [\max_k |a_{N,k}|]^\lambda$ for $\lambda = 1$ and $\lambda = t$ we find that we may also assume

$$\beta \geqq -\alpha \text{ and } \beta \geqq \frac{\rho}{t}. \tag{5}$$

We will assume both (4) and (5) in that which follows. For notational convenience we define

$$F(y) = \sup_k P\{|X_k - EX_k| \geqq y\}. \tag{6}$$

We will prove the following five theorems.

THEOREM 1. *If $\rho > 0$, $\alpha < \beta$, $t > 1$, and $y^t F(y) \leqq M < \infty$ for all $y > 0$, then for every $\varepsilon > 0$*

$$P\{|S_N| > \varepsilon\} \leqq O(N^{-\rho}). \tag{7}$$

THEOREM 2. *If $\rho > 0$, $\alpha < \beta$, $t > 1$, and $y^t F(y) \to 0$ as $y \to \infty$, then for every $\varepsilon > 0$*

$$P\{|S_N| > \varepsilon\} = o(N^{-\rho}). \tag{8}$$

THEOREM 3. *If $\beta(t-1) - \alpha > 0$, $\beta < 0$, $\alpha > \beta$, $t > 1$, and F satisfies*

$$\lim_{y\to\infty} F(y) = 0 \text{ and } \int_0^\infty y^t |dF(y)| < \infty, \tag{9}$$

then for every $\varepsilon > 0$

$$\sum_N N^{\beta(t-1)-\alpha-1} P\{|S_N| > \varepsilon\} < \infty. \tag{10}$$

THEOREM 4. *If $\rho > 0$, $\alpha < \beta$, $t \geqq 1$, and there exists a nonnegative and nonincreasing real valued function G satisfying* (9) *and such that $G(x) \geqq F(x)$ and*

$$\sup_{x\geqq 1} \sup_{y \geqq x} \frac{y^t F(y)}{x^t G(x)} = \gamma < \infty, \tag{11}$$

then for every $\varepsilon > 0$

$$\sum_N N^{\rho-1} P\{|S_N| > \varepsilon\} < \infty. \tag{12}$$

THEOREM 5. *If $\rho > 0$, $\alpha < \beta$, $t \geqq 1$, and F satisfies*

$$\lim_{y\to\infty} F(y) = 0 \text{ and } \int_0^\infty y^t \log^+ y |dF(y)| < \infty, \tag{13}$$

then (12) *holds for every $\varepsilon > 0$.*

3. **Proofs.** The method of proof used here was apparently first used by Erdös [10]. The method was generalized and improved by Katz [7] and modified still more by Pruitt [4]. Thus the proofs given here are at least third order modifications of that given in [10]. In spite of the fact that the general method of proof used here has been used before, the authors consider the results obtained here to be sufficiently interesting to warrant comprehensible (and therefore necessarily complete) proofs.

Throughout the proofs we assume that summations are taken only over those values of k for which $a_{N,k} \neq 0$.

We first show that

$$\begin{aligned} &P\{|S_N| > 3\varepsilon\} \\ &\leqq \sum_k P\{|a_{N,k}(X_k - EX_k)| > \varepsilon\} \end{aligned} \tag{14}$$

$$(15) \qquad + \sum_{j \neq k} P\{|a_{N,k}(X_k - EX_k)| > N^{-\delta}\} P\{|a_{N,j}(X_j - EX_j)| > N^{-\delta}\}$$

$$(16) \qquad + P\{|\sum_k a_{N,k} EY_{N,k}| > \varepsilon\}$$

$$(17) \qquad + P\{|\sum_k a_{N,k}(Y_{N,k} - EY_{N,k})| > \varepsilon\}$$

where

$$Y_{N,k} = \begin{cases} X_k - EX_k & \text{if } |a_{N,k}(X_k - EX_k)| \leqq N^{-\delta}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the probability in (16) is either zero or one since it is the probability of one constant exceeding another. Let $B_N = \{|S_N| > 3\varepsilon\}$, $D_N = \{|a_{N,k}(X_k - EX_k)| > \varepsilon$ for at least one $k\}$, $E_N = \{|a_{N,k}(X_k - EX_k)| > N^{-\delta}$ for at least two values of $k\}$, and $F_N = \{|\sum_k a_{N,k} Y_{N,k}| > 2\varepsilon\}$. We claim that $B_N \subset D_N \cup E_N \cup F_N$ or equivalently that $B_N^c \supset D_N^c \cap E_N^c \cap F_N^c$. If $\omega \in E_N^c$ then $Y_{N,k}(\omega) \neq X_k(\omega) - EX_k$ for at most one value of k. If ω is also in D_N^c that for all values of k we have $|a_{N,k}(X_k(\omega) - EX_k)| \leqq \varepsilon$; this holds in particular for that value of k (if any) for which $Y_{N,k} \neq X_k(\omega) - EX_k$. Thus for $\omega \in D_N^c \cap E_N^c$ the sums $S_N(\omega)$ and $\sum_k a_{N,k} Y_{N,k}(\omega)$ differ by at most ε. If $\omega \in F_N^c$ also, (i.e. $\omega \in D_N^c \cap E_N^c \cap F_N^c$) then $|\sum_k a_{N,k} Y_{N,k}(\omega)| \leqq 2\varepsilon$ so $|S_N(\omega)| \leqq 3\varepsilon$ and $\omega \in B_N^c$. Thus $P(B_N) \leqq P(D_N) + P(E_N) + P(F_N)$. We see that (14) is an upper bound on $P(D_N)$, (15) is an upper bound of $P(E_N)$, and the sum of (16) and (17) is an upper bound on $P(F_N)$.

Our method of proof involves showing that each of the expressions (14), (15), (16), and (17) tends to zero at a rate appropriate for the theorem in question.

Expression (13). For Theorems 1 and 2 we note that (14) is bounded by

$$\sum_k \frac{|a_{N,k}|^t}{\varepsilon^t} \sup_{y \geqq \varepsilon/|a_{N,k}|} (y^t P\{|X_k - EX_k| > y\})$$
$$\leqq CN^{-\rho}\left[\sup_k \sup_{y \geqq CN^\beta} (y^t P\{|X_k - EX_k| > y\})\right].$$

We remind the reader of our convention on the use of the constant C as stated at the beginning of §2. As the reader will notice, C is used twice in the last expression above and takes on different values in the two places. Under an assumption of Theorem 1 the quantity in square brackets in the last expression above is bounded so (14) is $O(N^{-\rho})$. Under the assumptions of Theorem 2 and the assumption $\beta \geqq \rho/t$, this same quantity in square brackets is $o(1)$ so (14) is $o(N^{-\rho})$.

For Theorem 3 we note that

$$\sum_N N^{\beta(t-1)-\alpha-1} \sum_k P\{|a_{N,k}(X_k - EX_k)| > \varepsilon\} \tag{18}$$

$$\leqq \sum_N N^{\beta(t-1)-\alpha-1} \sum_k F\left(\frac{\varepsilon}{|a_{N,k}|}\right)$$

$$\leqq \sum_M F(M-1) \sum_{\{(N,k)|M-1<\varepsilon/|a_{N,k}|\leqq M\}} N^{\beta(t-1)-\alpha-1}$$

$$= \sum_M [(F_M - 1) - F(M)] \sum_{\{(N,k)|\varepsilon\leqq|a_{N,k}|\leqq M\}} N^{\beta(t-1)-\alpha-1}. \tag{19}$$

Now since $\varepsilon/M \leqq |a_{N,k}|/CN^{-\beta}$ we see that there can be terms in the second summation in (19) only if $N \leqq [CM/\varepsilon]^{1/\beta}$. Since $\varepsilon/M \leqq |a_{N,k}|$ and $\sum_k |a_{N,k}| \leqq CN^{\alpha}$, for fixed values of N and M there are at most $CN^{\alpha}M/\varepsilon$ pairs (N,k) over which the second summation in (19) is taken. Thus (19), and therefore (18), is bounded by

$$\sum_M [F(M-1) - F(M)] \sum_{N=1}^{[CM/\varepsilon]^{1/\beta}} \frac{CN^{\alpha}M}{\varepsilon} N^{\beta(t-1)-\alpha-1}$$

$$= C \sum_M M^t[F(M-1) - F(M)] \left[M^{-t+1} \sum_{N=1}^{[CM/\varepsilon]^{1/\beta}} N^{\beta(t-1)-1}\right]$$

$$\leqq C \sum_M M^t[F(M-1) - F(M)] < \infty.$$

Note that if $t = 1$ then $\sum_{N=1}^{[CM/\varepsilon]^{1/\beta}} N^{\beta(t-1)-1}$ is approximately $\log[CM/\varepsilon]^{1/\beta}$ which is not bounded.

For Theorem 4 we use (11) and have

$$\sum_N N^{\rho-1} \sum_k P\{|a_{N,k}(X_k - EX_k)| > \varepsilon\} \tag{20}$$

$$\leqq \sum_{\{N|\varepsilon<CN^{-\beta}\}} N^{\rho-1} \sum_k P\{|a_{N,k}(X_k - EX_k)| > \varepsilon\} \tag{21}$$

$$+ \sum_{\{N|\varepsilon\geqq CN^{-\beta}\}} N^{\rho-1} \sum_k \gamma \frac{(\varepsilon/CN^{-\beta})^t G(\varepsilon/CN^{-\beta})}{(\varepsilon/|a_{N,k}|)^t}.$$

Since $\beta > 0$ there are only a finite number of elements of the set $\{N | \varepsilon < CN^{-\beta}\}$. The first term in (21) is then finite if $\sum_k P\{|a_{N,k}(X_k - EX_k)| > \varepsilon\}$ is finite for each N. But this is bounded by

$$\sum_k F\left(\frac{\varepsilon}{|a_{N,k}|}\right) \leqq \sum_M F(M) \sum_{\{k/M<\varepsilon/|a_{N,k}|\leqq M+1\}} 1$$

$$= \sum_M [F(M) - F(M+1)] \sum_{\{k|\varepsilon/|a_{N,k}|\leqq M+1\}} 1$$

$$\leqq \frac{CN^{\alpha}}{\varepsilon} \sum_M [M+1][F(M) - F(M+1)] < \infty.$$

In the above we used the fact that $\sum_k |a_{N,k}| \leqq CN^{\alpha}$ to deduce that the number of members of the set $\{k \mid \varepsilon/|a_{N,k}| < M+1\}$ is bounded by $CN^{\alpha}(M+1)/\varepsilon$. We have shown that the first term in (21) is finite. We note that the second term in (21) is bounded by

$$\begin{aligned} C \sum_{\{N \mid \varepsilon \geqq CN^{-\beta}\}} N^{\rho-1+t\beta} G\left(\frac{\varepsilon}{CN^{-\beta}}\right) \sum_k |a_{N,k}|^t & \\ \leqq C \sum_N N^{t\beta-1} G\left(\frac{\varepsilon}{CN^{-\beta}}\right) & \\ \leqq C \sum_N N^{t\beta}\left[G\left(\frac{\varepsilon}{C}N^{\beta}\right) - G\left(\frac{\varepsilon}{C}(N+1)^{\beta}\right)\right] & \\ \leqq C \int_0^{\infty} x^t |dG(x)| < \infty . & \end{aligned}$$

For Theorem 5 we have (20) bounded by

$$\sum_M [F(M-1) - F(M)] \sum_{\{(N,k) \mid \varepsilon/|a_{N,k}| \leqq M\}} N^{\rho-1} . \tag{22}$$

As we noted before, $\varepsilon/M < |a_{N,k}| \leqq CN^{-\beta}$ so that there can be terms in the second summation above only if $N \leqq [CM/\varepsilon]^{1/\beta}$. Now $(\varepsilon/M)^t < |a_{N,k}|^t$ and $\sum_k |a_{N,k}|^t \leqq CN^{-\rho}$ so that for fixed N and M there are at most $CN^{-\rho}(M/\varepsilon)^t$ terms in $\{(N,k) \mid \varepsilon/|a_{N,k}| < M\}$. Thus (22), and therefore (20), is bounded by

$$\begin{aligned} C \sum_M [F(M-1) - F(M)] \sum_{N=1}^{[CM/\varepsilon]^{1/\beta}} M^t N^{-1} & \\ \leqq C \sum_M F(M-1) - F(M)] M^t \log M & \\ \leqq C \int_0^{\infty} x^t \log^+ x \, |dF(x)| < \infty . & \end{aligned}$$

Expression (15). We note that (15) is bounded by

$$\left\{ \sum_k |a_{N,k}|^t N^{\delta t}\left[\left(\frac{N^{-\delta}}{|a_{N,k}|}\right)^t P\left\{ |X_k - EX_k| > \frac{N^{-\delta}}{|a_{N,k}|}\right\}\right]\right\}^2 . \tag{23}$$

The hypotheses of each theorem are sufficient to guarantee $y^t F(y) \leqq M < \infty$ for all $y > 0$, and this insures a bound on the quantity in square brackets in (23). Thus (23), and therefore (15), is bounded by

$$C\left(\sum_k |a_{N,k}|^t N^{\delta t}\right)^2 . \tag{24}$$

As noted at the beginning of §2, for $t \geqq 1$ we have

$$\sum_k |a_{N,k}|^t \leqq CN^{-\beta(t-1)+\alpha} . \tag{25}$$

We define

$$\rho_0 = \begin{cases} \rho, & \text{in Theorems 1, 2, 4, and 5,} \\ \beta(t-1)-\alpha, & \text{in Theorem 3} \end{cases} \tag{26}$$

and see that (24) is bounded by $CN^{-2(\rho_0-\delta t)}$. If

$$0 < \delta < \frac{\rho_0}{2t} \tag{27}$$

then (24) and (15) are not only $o(N^{-\rho_0})$ but the summation on N of N^{ρ_0-1} times the expression in (15) is finite.

EXPRESSION (16). We will show that (16) converges to zero at a large enough rate by showing that $\sum_k a_{N,k}EY_{N,k}$ converges to zero, so that (16) is zero for sufficiently large values of N.

Theorems 1 and 2 cause the most trouble here. That which follows holds in all five theorems provided $t>1$. It is necessary to introduce λ and use Hölder's Inequality only in Theorems 1 and 2 where nothing stronger than $y^tF(y)\to 0$ as $y\to\infty$ is available. We remind the reader that (i) $y^tF(y)\leqq M<\infty$ for all $y>0$ implies that for each $0<\lambda<t-1$ the expression $E|X_k-EX_k|^{t-\lambda}$ is uniformly bounded in k, and that (ii) for $p>1$ and nonnegative real numbers b_k and c_k we have Hölder's Inequality:

$$\sum b_kc_k \leqq \left(\sum b_k^p\right)^{1/p}\left(\sum c_k^{p/(p-1)}\right)^{(p-1)/p}. \tag{28}$$

Setting $b_k=|a_{N,k}|^{\lambda/t-1}$, $c_k=|a_{N,k}|^{t((t-1-\lambda)/(t-1))}$, and $p=(t-1)/\lambda$ we obtain from (28),

$$\begin{aligned}\sum_k |a_{N,k}|^{t-\lambda} &\leqq [\sum |a_{N,k}|]^{\lambda/(t-1)}[\sum |a_{N,k}|^t]^{1-\lambda/(t-1)} \\ &\leqq CN^{\alpha\lambda/(t-1)-\rho_0(1-\lambda/(t-1))}\end{aligned} \tag{29}$$

where ρ_0 was defined in (26). It follows that

$$\begin{aligned}\left|\sum_k a_{N,k}EY_{N,k}\right| &\leqq \sum_k |a_{N,k}|\left|\int_{|a_{N,k}(X_k-EX_k)|\leqq N^{-\delta}}(X_k-EX_k(dP\right| \\ &= \sum_k |a_{N,k}|\left|\int_{|a_{N,k}(X_k-EX_k)|>N^{-\delta}}(X_k-EX_k)dP\right| \\ &\leqq \sum_k |a_{N,k}|^{t-\lambda}N^{\delta(t-\lambda-1)}\quad |X_k-EX_k|^{t-\lambda}dP \\ &\leqq CN^{\delta(t-\lambda-1)}\sum_k |a_{N,k}|^{t-\lambda} \\ &\leqq CN^{\delta(t-\lambda-1)+\alpha\lambda/(t-1)-\rho_0(1-\lambda(t-1)}.\end{aligned} \tag{30}$$

If we choose δ so that

$$\delta < \rho_0/t - 1 \tag{31}$$

then for sufficiently small values of λ the exponent of N in (30) is negative. Thus (30), and therefore $\sum_k a_{N,k} EY_{N,k}$, converges to zero as $N \to \infty$.

We have yet to treat Theorems 4 and 5 when $t = 1$. In this case

$$\begin{aligned} \Big| \sum_k a_{N,k} EY_{N,k} \Big| &\leqq \sum_k |a_{N,k}| \, E|X_k - EX_k| \\ &\leqq N^{-\rho} \int_0^\infty x \, |dF(x)| \to 0. \end{aligned}$$

EXPRESSION (17). For positive even integers v, Markov's Inequality (p. 158 of [11]) gives

$$\begin{aligned} &P\Big\{\Big| \sum_k a_{N,k}(Y_{N,k} - EY_{N,k}) \Big| > \varepsilon\Big\} \\ &\leqq \frac{1}{\varepsilon^v} E\Big[\sum_k a_{N,k}(Y_{N,k} - EY_{N,k}) \Big]^v \\ &= \frac{1}{\varepsilon^v} \sum_1 \sum_2 C^{(v)}(m_1, \cdots, m_{a+b}) \prod_{k=1}^{a+b} (a_{N,f(k)})^{m_k} \\ &\qquad \cdot E(Y_{N,f(k)} - EY_{N,f(k)})^{m_k}, \end{aligned} \tag{32}$$

where the first sum is taken over all integers $a, b, m_1, \cdots, m_{a+b}$ such that $2 \leqq m_k < t$ for $k = 1, \cdots, a$; $t \leqq m_k$ for $k = a+1, \cdots, a+b$; $\sum_k m_k = v$; and distinct sets of integers $\{m_1, \cdots, m_{a+b}\}$ appear only once in the sum; the second sum is over all one-to-one functions f mapping $\{1, \cdots, a+b\}$ into the positive integers; and $C^{(v)}(m_1, \cdots, m_{a+b})$ is $v!/\prod_k m_k!$ which is the coefficient of the term $\prod_{k=1}^{a+b} (a_{N,f(k)})^{m_k} (Y_{N,f(k)} - EY_{N,f(k)})^{m_k}$ in the expansion of $[\sum_k a_{N,k}(Y_{N,k} - EY_{N,k})]^v$. Note that all the $C^{(v)}(m_1, \cdots, m_{a+b})$ are bounded by a constant depending on v but not on N or the $a_{N,k}$'s. Note also that $E(Y_{N,f(k)} - EY_{N,f(k)})^1 = 0$ so any term involving an exponent one is omitted in the above expansion.

Not having a tth moment available in Theorems 1 and 2 again causes a little trouble. We deal with expression (32) under the assumption $t > 1$. We remember that the assumptions of each theorem guarantee that when $t > 1$, the moments $E|X_k - EX_k|^{t-\lambda}$ are uniformly bounded in k for each fixed λ with $0 < \lambda \leqq t-1$. We note that this implies a uniform bound in both N and k for the moments $E|Y_{N,k} - EY_{N,k}|^{t-\lambda}$ provided λ is fixed and satisfies $0 < \lambda \leqq t-1$.

Let t_0 be the largest integer strictly less than t. If $1 \leqq k \leqq a$ then

$$\begin{aligned} &|(a_{N,f(k)})^{m_k} E(Y_{N,f(k)} - EY_{N,f(k)})^{m_k}| \\ &\leqq |a_{N,f(k)}| \, [\max_k |a_{N,k}|]^{m_k - 1} [\sup_{N,k} E|Y_{N,k} - EY_{N,k}|^{t_0} + 1]. \end{aligned} \tag{33}$$

Remembering the definition of $Y_{N,k}$ and requiring $0 < \lambda \leqq t-1$ we see that if $a+1 \leqq k \leqq a+b$ then

$$
\begin{aligned}
(34)\qquad &\left|(a_{N,f(k)})^{m_k}E(Y_{N,f(k)} - EY_{N,f(k)})^{m_k}\right| \\
&\leqq \left|a_{N,f(k)}\right|^{t-\lambda}[\sup_{N,k} E\left|Y_{N,k} - EY_{N,k}\right|^{t-\lambda} + 1] \\
&\quad \cdot [\sup_{\omega,N,k}\left|a_{N,k}\right| \left|Y_{N,k}(\omega) - EY_{N,k}\right|]^{m_k-t+\lambda} \\
&\leqq C\left|a_{N,f(k)}\right|^{t-\lambda}N^{-\delta(m_k-t+\lambda)}
\end{aligned}
$$

where C depends on v but not on N and k. From (33) and (34) it follows that (32) is bounded by

$$
\begin{aligned}
&C\sum_1 \sum_2 \prod_{k=1}^{a}\left|a_{N,f(k)}\right|[\max_k\left|a_{N,k}\right|]^{m_k-1} \\
&\quad \cdot \prod_{k=a+1}^{a+b}\left|a_{N,f(k)}\right|^{t-\lambda}N^{-\delta(m_k-t+\lambda)} \\
(35)\qquad &\leqq C\sum_1 N^{-\beta(m_1+\dots+m_a-a)-\delta(m_{a+1}+\dots+m_{a+b}-bt+b\lambda)} \\
&\quad \cdot \sum_2 \prod_{k=1}^{a}\left|a_{N,f(k)}\right| \prod_{k=a+1}^{a+b}\left|a_{N,f(k)}\right|^{t-\lambda} \\
&\leqq C\sum_1 N^{-\beta(m_1+\dots+m_a-a)-\delta(v-m_1-\dots-m_1-bt+b\lambda)} \\
&\quad \cdot N^{a\alpha}N^{b[\alpha\lambda/(t-1)-\rho_0(1-\lambda/(t-1))]}
\end{aligned}
$$

where ρ_0 was defined in (26) and the constant C depends on v but not on N. In the last inequality we used

$$
\begin{aligned}
&\sum_2 \prod_{k=1}^{a}\left|a_{N,f(k)}\right| \prod_{k=a+1}^{a+b}\left|a_{N,f(k)}\right|^{t-\lambda} \\
&\quad \leqq \left[\sum_k \left|a_{N,k}\right|\right]^a \left[\sum_k \left|a_{N,k}\right|^{t-\lambda}\right]^b
\end{aligned}
$$

and (29).

There are only a finite number of terms (the number depending on v but not depending on N) in $\sum_1$. We will show that by appropriate choices of v, λ, and δ all terms are less than or equal to $O(N^{-\tau})$ with $\tau > \rho_0$. This will be sufficient to complete the proofs of all five theorems (except for the case $t = 1$ in Theorems 4 and 5) provided the various conditions put on v, λ, and δ here and conditions (27) and (31) can be satisfied simultaneously. We rewrite the exponent of a single term in $\sum_1$ of (35) as

$$
\begin{aligned}
(36)\qquad &-\beta\left(\frac{m_1+\cdots+m_a}{2} - a\right) - \left((\beta - 2\delta)\frac{m_1+\cdots+m_a}{2} - a\alpha\right) \\
&\quad - b\rho_0 - \delta v + b\left(\delta(t-\lambda) + \lambda\frac{(\alpha+\rho_0)}{t-1}\right).
\end{aligned}
$$

Since $m_i \geqq 2$ for all i we see that $(m_1 + \cdots + m_a)/2 \geqq a$. It follows that the first two terms in (36) are nonpositive (remember the hypothesis $\alpha < \beta$) provided

$$\delta \leqq \frac{1}{2} \min \{\beta, \beta - \alpha\}. \tag{37}$$

If $b = 0$, then choosing

$$v > \frac{\rho_0}{\delta} \tag{38}$$

is sufficient to make (36) $< -\rho_0$. Notice that since $m_i \geqq t$ for $i = a+1, \cdots, a+b$ we have $bt \leqq v$. We rewrite the last three terms of (36) as

$$\begin{aligned} &-\rho_0 - (b-1)\left(\rho_0 - \delta(t-\lambda) - \lambda \frac{(\alpha + \rho_0)}{t-1}\right) \\ &\quad - \delta\left((v - t + \lambda) - \frac{\lambda}{\delta}\left[\frac{\alpha + \rho_0}{t-1}\right]\right). \end{aligned} \tag{39}$$

We can make (39) less than $-\rho_0$ by taking

$$0 < \delta \leqq \frac{\rho_0}{2t}, \tag{40}$$

$$\lambda \leqq \min\left\{t, \quad \frac{\rho_0(t-1)}{2(\alpha + \rho_0)}\right\}, \tag{41}$$

$$v \geqq 3t, \tag{42}$$

$$\lambda \leqq \frac{\delta t(t-1)}{\alpha + \rho_0}. \tag{43}$$

Thus we are done (except for the case $t = 1$ in Theorems 4 and 5) provided we can simultaneously satisfy (27), (31), (37), (38), (40), (41), (42), and (43). First choose δ to satisfy (27) and (31), (37), and (40). These four conditions do not involve λ or v. Next choose v to satisfy (38) and (42). These two conditions do not involve λ. Finally choose λ so as to satisfy (41) and (43).

The proof for the case $t = 1$ uses the uniformity of a bound on tth moments and holds for Theorems 4 and 5. In this case we get (32) bounded by

$$\begin{aligned} &C \sum_1 \sum_2 \prod_{k=1}^{a} |a_{N,f(k)}| [\max_k |a_{N,k}|]^{m_k - 1} \\ &\qquad \cdot \prod_{k=a+1}^{a+b} |a_{N,f(k)}|^t N^{-\delta(m_k - t)} \\ &\leqq C \sum_1 N^{-\beta(m_1 + \ldots + m_a - a) - \delta(m_{a+1} + \ldots + m_{a+b} - bt)} \\ &\qquad \cdot \sum_2 \prod_{k=1}^{a} |a_{N,f(k)}| \prod_{k=a+1}^{a+b} |a_{N,f(k)}|^t \\ &\leqq C \sum_1 N^{-\beta(m_1 + \ldots + m_a - a) - \delta(v - m_1 - \ldots - m_a - bt)} N^{a\alpha} N^{-b\rho_0}. \end{aligned} \tag{44}$$

The exponent in (44) can be rewritten as

$$(45)\qquad -\beta\left(\frac{m_1+\cdots+m_a}{2}-a\right)-\left((\beta-2\delta)\frac{m_1+\cdots+m_a}{2}-a\alpha\right) \\ -\delta(v-bt)-b\rho_0.$$

If $0<\delta\leqq\frac{1}{2}\min\{\beta,\beta-\alpha\}$ and $v>\max\{t,\rho_0/\delta\}$, then the first two terms above are always nonpositive, the third term is less than $-\rho_0$ when $b=0$, the third term is negative when $b=1$, and the fourth term is always nonpositive. Thus (45) is always less than $-\rho_0$.

4. **Relationships between these results and results in the literature.** Although Theorems 4 and 5 are, in a sense, considerably stronger than Theorem 3, two known results can be obtained as corollaries to Theorem 3 by specializing the constant t and the constants α and β from (1) and (2). Pruitt [4] has proved for independent and identically distributed random variables.

THEOREM (PRUITT). *If $\Sigma_k|a_{N,k}|\leqq M$, $\max_k|a_{N,k}|\leqq O(n^{-\gamma}$, $\gamma>0$, and $E|X_k|^{1+1/\gamma}<\infty$, then $S_N\to 0$ almost everywhere.*

If we set $t=1+1/\gamma$, $\alpha=0$, and $\beta=\gamma$, then we have the conditions of Pruitt's theorem. In this case the result of Theorem 3 from (10) is (since $\beta(t-1)-\alpha-1=0$)

$$\sum_N P\{|S_N|>\varepsilon\}<\infty \text{ for every } \varepsilon>0$$

which implies $S_N\to 0$ a.e. Thus Theorem 3 contains Pruitt's theorem. For independent and identically distributed random variables Baum and Katz have proved (see part of Theorem 3 of [8] and [9])

THEOREM (BAUM AND KATZ). *If $t>1$, $r>1$, $\frac{1}{2}<r/t\leqq 1$, and $E|X_k|^t<\infty$, then for every $\varepsilon>0$*

$$\sum_N N^{r-2}P\left\{\left|\sum_{k=1}^{N}(X_k-EX_k)\right|>N^{r/t}\varepsilon\right\}<\infty.$$

If we leave t fixed and set $\alpha=1-r/t$ and $\beta=r/t$, then in order to use Theorem 3 we need $\beta(t-1)-\alpha>0$ and $\alpha<\beta$. But these two conditions are equivalent to the conditions $r>1$ and $\frac{1}{2}<r/t$ in the above theorem. Applying Theorem 3 gives the conclusion of the theorem above.

Baum and Katz have the equivalent of the above theorem (see Theorem 1 of [9]). when $1\leqq t<2$ and $r=1$. This does not follow from our Theorem 3 since we would have to leave t fixed and set $\alpha=1-1/t$ and $\beta=1/t$ in order to apply Theorem 3. This however gives $\beta(t-1)-\alpha=0$ contradicting the hypothesis $\beta(t-1)-\alpha>0$ of Theorem 3.

Part of Theorem 4 of Baum and Katz [9] is implied by our Theorem 2. Baum and Katz proved

THEOREM (BAUM AND KATZ). *If $\tau \geqq 0$, then the following two statements are equivalent*:

(a) $$n^{\tau+1} P\{|X_k| > n\} \to 0 \text{ and } \int_{|x|<n} x dP\{X_k \leqq x\} \to 0,$$

(b) $$n^{\tau} P\{|X_1 + \cdots + X_n| > n\varepsilon\} \to 0 \text{ for each } \varepsilon > 0.$$

If we set $\alpha = 0$, $\beta = 1$, and $t = \tau + 1$ in Theorem 2, then we see that Theorem 2 contains "(a) implies (b)" in the above theorem when $\tau > 0$. This follows since $\rho \geqq \beta(t-1) - \alpha$ and since for $\tau > 1$ (a) above implies $EX_k \equiv 0$.

We have not worked on the cases when $t < 1$ as have Baum and Katz in parts of [8] and [9].

The work done here would appear to give a different type of result from that obtained by Jamison, Orey, and Pruitt in [3]. A particular type of weight was used in [3] and the authors restricted themselves to the laws of large numbers without rates of convergence.

5. **Concluding remarks.** It would be interesting to know whether or not the theorems stated here are optimal in some sense. Pruitt [4 shows that his theorem, stated in the preceeding section, is sharp in the sense that for every $\gamma > 0$ there is a set of coefficients $\{a_{N,k}\}$ satisfying $\max_k |a_{N,k}| = O(n^{-\gamma})$ such that $S_N \to 0$ a.e. implies $E|X_k|^{1+1/\gamma} < \infty$. One would hope to obtain the same type of sharpness here, especially for Theorem 3 and 5.

If one guarantees only the existence of some moment $E|X|^t$ with $t < 1$, then EX may not exist and there is no sequence $\{b_N\}$ such that one can guarantee the convergence of

$$P\left\{\left|\frac{X_1 + \cdots + X_N}{N} - b_N\right| > \varepsilon\right\}$$

to zero. Thus, in a sense, the law of large numbers does not hold (so convergence rates for the law of large numbers make no sense) when $t < 1$. However, using different weights Baum and Katz obtain, for values of $t < 1$, results similar to those mentioned in the preceeding section (see Theorem 4 of [8] and the last half of Theorem 3 of [9]). For completeness, the case $t \leqq 1$ should be thoroughly treated for the coefficient sequences $\{a_{N,k}\}$ considered here.

Though considerable machinery was required to obtain the results obtained here (and in most of the cited references for that matter), the results obtained are in one sense very crude. The result of Theorem 1, for example, states that $P\{|S_N| > \varepsilon\} \leqq CN^{-\rho}$. While asymptotically the constant C does not matter, the statistician wishing to use the theorem has at present no way of estimating C. What are needed here are results of the type obtained by Blackwell and Hodges [12] for lattice valued random variables except that a strict upper bound is needed on the absolute value of the error term.

References

1. D. L. Hanson and L. H. Koopmans, *On the convergence rate of the law of large numbers for linear combinations of independent random variables*, Ann. Math. Statist. **36** (1965), 559–564.

2. Hermann Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. **23** (1952), 493–507.

3. B. Jamison, S. Oreye and W. Pruitt, *Convergence of weighted averages of independent random variables*, Z. Wahrscheinlichkeitstheorie **4** (1965), 40–44.

4. William E. Pruitt, *Summability of independent random variables*, J. Math. Mech. (to appear).

5. D. L. Hanson and L. H. Koopmans, *Convergence rates for the law of large numbers for linear combinations of exchangeable and *-mixing stochastic processes*, Ann. Math. Statist. **36** (1965), 1840–1852.

6. R. V. Chacon, *Ordinary means imply recurrent means*, Bull. Amer. Math. Soc. **70** (1964), 796–797.

7. Melvin L. Katz, *The probability in the tail of a distribution*, Ann. Math. Statist. **34** (1963), 312–318.

8. L. E. Baum and Melvin Katz, *Convergence rates in the law of large numbers*, Bull. Amer. Math. Soc. **69** (1963), 771–772.

9. ———, *Convergence rates in the law of large numbers*. II, Tech. Rep. No. 75, Univ. of New Mexico Albuquerque, Dept. of Math., 1964.

10. P. Erdös, *On a theorem of Hsu and Robbins*, Ann. Math. Statist. **20** (1949), 286–291.

11. M. Loève, *Probability theory*, 3rd ed., Van Nostrand, Princeton, N. J., 1963.

12. David Blackewel and J. L. Hodges, Jr., *The probability in the extreme tail of a convolution*, Ann. Math. Statist. **30** (1959), 1113–1120.

University of Missouri,
Columbia, Missouri