The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1235–1248

CONVERGENCE PROPERTIES OF S_n UNDER MOMENT RESTRICTIONS[1]

BY R. J. SERFLING

Florida State University

1. Introduction and preliminaries. Let $\{X_i\}_{-\infty}^{\infty}$ be a sequence of rv's having finite variances $\{\sigma_i^2\}$. Assume throughout (without loss of generality) that $E(X_i) \equiv 0$. For each vector $\mathbf{X}_{a,n} = (X_{a+1}, \cdots, X_{a+n})$ of n consecutive X_i's, let $F_{a,n}$ denote the joint df and let

$$S_{a,n} = \textstyle\sum_{a+1}^{a+n} X_i. \tag{1.0}$$

In statements about $\mathbf{X}_{0,n}$ only, the abbreviated notation $\mathbf{X}_n$, F_n, S_n, etc., shall be employed.

This paper concerns stochastic convergence properties of sequences $\{S_{a,n}\}_{n=1}^{\infty}$. It will suffice to prove statements about the sequence $\{S_n\}_{n=1}^{\infty}$, but the more general notation will be of use in formulating some of the restrictions adopted.

Convergence properties of the following types shall be discussed:

$$P[S_n/n \to 0] = 1; \tag{1.1}$$

$$\textstyle\sum_1^\infty a_n P[\sup_{k \geqq n} |S_k/k| > \varepsilon] \quad \text{converges for every } \varepsilon > 0; \tag{1.2}$$

$$P[\limsup_{n\to\infty} |S_n/b_n| \leqq 1] = 1; \tag{1.3}$$

$$\textstyle\sum_1^\infty c_n P[\sup_{k \geqq n} |S_k/b_k| > \varepsilon] \quad \text{converges for every } \varepsilon > 0; \tag{1.4}$$

$$\textstyle\sum_1^\infty P[|S_n/d_n| > \varepsilon] \quad \text{converges for every } \varepsilon > 0. \tag{1.5}$$

In the above, $\{a_i\}$, $\{b_i\}$, $\{c_i\}$ and $\{d_i\}$ are sequences of constants.

Condition (1.1) expresses the strong law of large numbers (SLLN) for the sequence $\{X_i\}_1^{\infty}$ and condition (1.2) represents information regarding the rate of the convergence in (1.1). The larger that the a_n's may be chosen (in asymptotic order of magnitude), the sharper is the result stated by (1.2). Condition (1.3) expresses a form of the law of the iterated logarithm (LIL) (e.g., in typical cases b_n may be taken as small as $O((n \ln\ln n)^{\frac{1}{2}})$.) and condition (1.4), like (1.2), concerns the rate of convergence. Finally, condition (1.5) states that the sequence $\{S_n/d_n\}_{n=1}^{\infty}$ *converges completely* to zero in the sense of Hsu and Robbins [7]. The smaller that the d_n's may be chosen, the sharper is the statement. By the Borel–Cantelli lemma, complete convergence implies strong convergence.

Properties such as (1.1)–(1.5) will be obtained as consequences of restrictions imposed upon the absolute νth moments, for some $\nu \geqq 2$, of sums $\sum_{a+1}^{a+n} w_i X_i$, where the w_i are given constants (e.g., $w_i \equiv 1$, or $w_i = \ln i$). *Thus it is not assumed that the X_i's are mutually independent and, in fact, the only restrictions on the dependence will be what is implied by the restrictions of the type mentioned.* (See [12], [13] for details.)

Received April 23, 1969; revised January 29, 1970.

[1] Research supported in part by NSF Grant GZ-315.

The results below are obtained by adaptation of more or less standard arguments [3], [10], [11] to make use of recent results [13] which state bounds for the νth moment of

$$M_{a,n} = \max\{|S_{a,1}|, \cdots, |S_{a,n}|\} \tag{1.6}$$

in terms of assumed bounds on the νth moment of $|S_{a,n}|$.

Section 2 deals with the SLLN under moment restrictions of second order only. In Section 3, the LIL (which implies the SLLN) is derived under moment restrictions of order higher than the second. The case of bounded random variables is considered in Section 4. Convergence rates associated with these SLLN and LIL are studied in Section 5. Finally, in Section 6, the question of norming S_n suitably for S_n/d_n to converge completely to zero is considered.

2. SLLN under moment restrictions of second order. (Stronger conclusions (LIL) under restrictions on moments of order higher than second appear in subsequent sections.)

The proofs of Theorems 2.1 and 2.2 below each make use of the following lemma, which is a special case of Theorem A of [13]. Here $g(F_{a,n})$ denotes a functional depending on the joint df of $X_{a+1}, \cdots, X_{a+n}$. Examples are $g(F_{a,n}) = \sum_{a+1}^{a+n} \sigma_i^2$ or simply $g(F_{a,n}) = n$.

LEMMA A. *Suppose that* (2.1) *holds, where* $g(F_{a,n})$ *satisfies* (2.2). *Then*

$$E(M_{a,n}^2) \leqq (\log_2 2n)^2 g(F_{a,n}) \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1). \tag{2.0}$$

THEOREM 2.1. *Suppose that*

$$E(S_{a,n}^2) \leqq g(F_{a,n}) \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1), \tag{2.1}$$

where $g(F_{a,n})$ *satisfies*

$$g(F_{a,k}) + g(F_{a+k,l}) \leqq g(F_{a,k+l}), \qquad 1 \leqq k < k+l, \tag{2.2}$$

and, uniformly in $a \geqq 0$,

$$g(F_{a,n}) = O\left(\frac{n^2}{(\ln n)^2(\ln\ln n)^2}\right), \qquad n \to \infty. \tag{2.3}$$

Then, with probability 1,

$$S_n/n \to 0, \qquad n \to \infty. \tag{2.4}$$

PROOF. Let $\varepsilon > 0$. By (2.1), (2.3) and Chebyshev's inequality,

$$P[|S_n| > \varepsilon n] = O\left(\frac{1}{(\ln n)^2(\ln\ln n)^2}\right). \tag{2.5}$$

Put $n_k = [\exp k^{\frac{1}{2}}]$, where $[\cdot]$ denotes integer part. Then

$$P[|S_{n_k}| > \varepsilon n_k] = O\left(\frac{1}{k\ln^2 k}\right), \qquad k \to \infty, \tag{2.6}$$

and by the Borel–Cantelli Lemma, with probability 1,

$$|S_{n_k}| < \varepsilon n_k \qquad \text{for all } k \text{ large enough.} \tag{2.7}$$

Put $\xi_k = M_{n_k, n_{k+1}-n_k}$. Then

$$|S_n| \leqq |S_{n_k}| + \xi_k \qquad \text{if} \quad n_k \leqq n \leqq n_{k+1}. \tag{2.8}$$

By Lemma A

$$E(\xi_k^2) \leqq [\log_2 2(n_{k+1}-n_k)]^2 g(F_{n_k,n_{k+1}-n_k}) \tag{2.9}$$

and by (2.3), uniformly in k,

$$E(\xi_k^2) = O\left(\frac{(n_{k+1}-n_k)^2}{[\ln\ln(n_{k+1}-n_k)]^2}\right), \tag{2.10}$$

so that, uniformly in k,

$$P[\xi_k > \varepsilon n_k] = O\left[\left(\frac{n_{k+1}-n_k}{n_k}\right)^2 \frac{1}{[\ln\ln(n_{k+1}-n_k)]^2}\right]. \tag{2.11}$$

It is easily seen that the right-hand side of (2.11) is $O((k\ln^2 k)^{-1})$. Hence (Borel–Cantelli), with probability 1,

$$\xi_k < \varepsilon n_k \qquad \text{for all } k \text{ large enough.} \tag{2.12}$$

Combining (2.7), (2.8) and (2.12), we have, with probability 1,

$$|S_n/n| \leqq 2\varepsilon \qquad \text{for all } n \text{ large enough.} \tag{2.13}$$

Therefore, (2.4) holds.

COROLLARY 2.1.1. *Suppose that*

$$E(X_i X_j) \leqq \rho_{j-i}\sigma_i\sigma_j \qquad (\textit{all } i \leqq j), \tag{2.14}$$

where $0 \leqq \rho_k \leqq 1$ *for all* $k = 0, 1, \cdots$. *If, uniformly in* $a \geqq 0$,

$$\left(\textstyle\sum_{a+1}^{a+n} \sigma_i^2\right)\left(\textstyle\sum_0^n \rho_i\right) = O\left(\frac{n^2}{(\ln n)^2(\ln\ln n)^2}\right), \qquad n \to \infty, \tag{2.15}$$

then, with probability 1, $S_n/n \to 0$.

The result is obtained from Theorem 2.1 by taking $g(F_{a,n})$ equal to twice the left-hand side of (2.15). Then (2.2) is trivially satisfied and, by Lemma 2.1 below, (2.1) is satisfied.

Note that the uniformity part of condition (2.15) restricts the possible variation in the sequence $\{\sigma_i\}$, but the requirement on the sequence $\{\rho_i\}$ is quite mild. Thus Corollary 2.1.1 has a useful implication in the case of a *weakly stationary* sequence, i.e., when $\sigma_i^2 \equiv \sigma^2$ and $E(X_i X_j) = r_{j-i}$ (all i, all j). Then (2.14) is satisfied with $\sigma_i = \sigma_j = \sigma$ and $\rho_{j-i} = |r_{j-i}|/\sigma^2$ and we have

COROLLARY 2.1.2. *Let $\{X_i\}$ be weakly stationary with $E(X_i X_{i+j}) = r_j$ (all i, all $j = 0, 1, \cdots$). If*

$$\sum_1^n |r_j| = O\left(\frac{n}{(\ln n)^2(\ln \ln n)^2}\right), \tag{2.16}$$

then, with probability 1, $S_n/n \to 0$.

Since $|r_j| \leqq r_0$ by the Cauchy–Schwarz inequality, we have

$$\sum_1^n |r_j| = O(n) \tag{2.17}$$

automatically for a weakly stationary sequence. Hence (2.16) and likewise (2.15) are rather mild dependence restrictions.

In contrast with Corollary 2.1.1, a later result (Corollary 2.2.1) allows much greater variation within the sequence $\{\sigma_i\}$ but at the expense of more stringent conditions on the sequence $\{\rho_i\}$.

The lemma necessary to Corollary 2.1.1 above and to Corollary 2.2.1 later is the following.

LEMMA 2.1. *Suppose that*

$$E(X_i X_j) \leqq \rho_{j-i}\sigma_i\sigma_j \qquad (\text{all } i \leqq j), \tag{2.14}$$

where $0 \leqq \rho_k \leqq 1$ *for all* $k = 0, 1, \cdots$. *Let* $b_i \geqq 0$ *(all i). Then*

$$E(\textstyle\sum_{a+1}^{a+n} b_i X_i)^2 \leqq 2(\sum_{a+1}^{a+n} b_i^2\sigma_i^2)(\sum_0^n \rho_i). \tag{2.18}$$

PROOF. Write $c_i = b_i\sigma_i$. Then, applying (2.14),

$$\begin{aligned} E(\textstyle\sum_{a+1}^{a+n} b_i X_i)^2 &\leqq \textstyle\sum_{a+1}^{a+n} c_i^2 + 2\sum_{i=a+1}^{a+n-1}\sum_{j=i+1}^{a+n} c_i c_j \rho_{j-i} \\ &= \textstyle\sum_{a+1}^{a+n} c_i^2 + 2\sum_{j=1}^{n-1}\rho_j \sum_{i=a+1}^{a+n-j} c_i c_{i+j} \\ &\leqq \textstyle\sum_{a+1}^{a+n} c_i^2 + \sum_{j=1}^{n-1}\rho_j(\sum_{i=a+1}^{a+n-j} c_i^2 + \sum_{i=a+1}^{a+n-j} c_{i+j}^2) \\ &\leqq \textstyle\sum_{a+1}^{a+n} c_i^2 + \sum_{j=1}^{n-1}\rho_j(2\sum_{i=a+1}^{a+n} c_i^2), \end{aligned}$$

which implies (2.18) since $\rho_0 = 1$.

The following theorem, generalizing a result proved by Rademacher and Mensov for mutually orthogonal rv's (see [3] 157, or [11] 86), gives conditions under which the series $\sum_1^\infty X_n$ converges with probability 1. Application of the result to the series $\sum_1^\infty X_n/n$ then yields, by the Kronecker lemma ([15] 238), a criterion for the convergence, with probability 1, of S_n/n to zero.

THEOREM 2.2. *Suppose that*

$$E(S_{a,n}^2) \leqq g(F_{a,n}) \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1) \tag{2.19}$$

where $g(F_{a,n})$ satisfies

$$g(F_{a,k}) + g(F_{a+k,l}) \leqq g(F_{a,k+l}), \qquad 1 \leqq k < k+l. \tag{2.20}$$

Suppose that for k sufficiently large, say $k \geqq k_0 > 1$,

$$g(F_{k,n}) \leqq h(F_{k,n})/\ln^2(k+1), \tag{2.21}$$

where $h(F_{a,n})$ *satisfies*

$$h(F_{a,k}) + h(F_{a+k,l}) \leqq h(F_{a,k+l}), \qquad 1 \leqq k < k+l, \tag{2.22}$$

and

$$h(F_{1,n}) \to A < \infty \qquad (n \to \infty). \tag{2.23}$$

Then, with probability 1,

$$\textstyle\sum_1^\infty X_i \quad \textit{converges.} \tag{2.24}$$

PROOF. Let $n \geqq k_0$. Then, by (2.1), (2.3) and (2.5),

$$E(\textstyle\sum_n^\infty X_i)^2 \leqq A/\ln^2 n. \tag{2.25}$$

Put $n_k = 2^k$. Then, by (2.25), the series

$$\textstyle\sum_{k=1}^\infty E(\sum_{n_k}^\infty X_i)^2 \tag{2.26}$$

converges, implying, with probability 1, that

$$\textstyle\sum_{n_k}^\infty X_i \to 0, \qquad k \to \infty. \tag{2.27}$$

Now let $\xi_k = M_{n_k, n_{k+1}-n_k}$. By Lemma A and (2.19), (2.20) and (2.21), since $n_{k+1} - n_k = n_k$, we have for k sufficiently large

$$E(\xi_k^2) \leqq \frac{(\log_2 n_k)^2}{\ln^2 n_k} h(F_{n_k,n_k}). \tag{2.28}$$

It follows by this, (2.22) and (2.23) that the series

$$\textstyle\sum_{k=1}^\infty E(\xi_k^2) \tag{2.29}$$

converges, implying, with probability 1, that

$$\xi_k \to 0, \qquad k \to \infty. \tag{2.30}$$

Combining (2.27) and (2.30), we obtain, with probability 1,

$$\textstyle\sum_n^\infty X_i \to 0, \qquad n \to \infty. \tag{2.31}$$

COROLLARY 2.2.1. *Let* $\{X_i\}$ *satisfy*

$$E(X_i X_j) \leqq \rho_{j-i}\sigma_i\sigma_j \qquad (\textit{all } i \leqq j), \tag{2.14}$$

where $0 \leqq \rho_k \leqq 1$ *for all* $k = 0, 1, \cdots$. *If* $\{\sigma_i\}$ *and* $\{\rho_i\}$ *satisfy*

$$\sum_1^\infty \left(\frac{\ln i}{i}\right)^2 \sigma_i^2 < \infty \qquad \textit{and} \tag{2.32}$$

$$\textstyle\sum_1^\infty \rho_i < \infty, \tag{2.33}$$

then, with probability 1, $S_n/n \to 0$.

PROOF. Theorem 2.2 is to be applied to the rv's $Y_i = X_i/i$. (As mentioned earlier, convergence of the series $\sum_1^\infty Y_i$ implies that $S_n/n \to 0$.) Note that

$$E(Y_i) \equiv 0, \qquad E(Y_i^2) = \sigma_i^2/i^2. \tag{2.34}$$

Define

$$g(F_{a,n}) = 2(\textstyle\sum_{a+1}^{a+n} \sigma_i^2/i^2)(\sum_0^n \rho_i) \tag{2.35}$$

and

$$h(F_{a,n}) = 2\left[\sum_{a+1}^{a+n} \left(\frac{\ln i}{i}\right)^2 \sigma_i^2\right]\left(\sum_0^n \rho_i\right). \tag{2.36}$$

By Lemma 2.1 and (2.34),

$$E(\textstyle\sum_{a+1}^{a+n} Y_i)^2 \leqq g(F_{a,n}) \qquad \text{(all } a \geqq 0, \text{ all } n \geqq 1). \tag{2.37}$$

Thus the conditions of Theorem 2.2 are satisfied, completing the proof.

A special case of Cor. 2.2.1 is the Rademacher–Mensov result:

COROLLARY 2.2.2. *Let* $\{X_i\}$ *be mutually orthogonal* rv's. *If the series* $\sum_1^\infty (\ln n)^2 \sigma_n^2/n^2$ *converges, then, with probability* 1, $S_n/n \to 0$.

3. LIL under moment restrictions of order higher than second. (The special case of bounded variables is reserved for the subsequent section.) Analogously to the use of Lemma A in the previous section, the following lemma, which is Theorem B of [13], is needed in the present section.

LEMMA B. *Let* $\nu > 2$. *Suppose that* (3.1) *holds, where* $g(n)$ *is nondecreasing,* $2g(n) \leqq g(2n)$ *and* $g(n)/g(n+1) \to 1$ *as* $n \to \infty$. *Then there exists a finite constant* K *such that*

$$E(M_{a,n}^\nu) \leqq K g^{\frac{1}{2}\nu}(n) \qquad (\textit{all } a \geqq 0, \textit{ all } n \geqq 1). \tag{3.0}$$

THEOREM 3.1. *Let* $\nu > 2$. *Suppose that*

$$E|S_{a,n}|^\nu \leqq g^{\frac{1}{2}\nu}(n) \qquad (\textit{all } a \geqq 0, \textit{ all } n \geqq 1), \tag{3.1}$$

where $g(n)$ *is nondecreasing,* $2g(n) \leqq g(2n)$ *and* $g(n)/g(n+1) \to 1$ *as* $n \to \infty$. *Suppose also that* $g(2n)/g(n)$ *is bounded,* $n \to \infty$. *Then, with probability* 1,

$$|S_n| = O(g^{\frac{1}{2}}(n)(\ln n)^{1/\nu}(\ln \ln n)^{2/\nu}), \qquad n \to \infty. \tag{3.2}$$

PROOF. Imitating well-known techniques of argument (e.g., Lamperti [10]), put $a(n) = g^{\frac{1}{2}}(n)(\ln n)^{1/\nu}(\ln\ln n)^{2/\nu}$ and $M_n = M_{0,n}$. By Lemma B,

$$P[M_n > a(n)] \leqq \frac{K}{(\ln n)(\ln \ln n)^2}. \tag{3.3}$$

Put $n_k = 2^k$. Then (Borel–Cantelli), with probability 1,

$$M_{n_k} < a(n_k) \qquad \text{for } k \text{ large enough.} \tag{3.4}$$

Now, for $n_k \leqq n \leqq n_{k+1}$, we have

$$a(n) \geqq a(n_k) \quad \text{and} \quad |S_n| \leqq M_{n_{k+1}} \tag{3.5}$$

and thus, with probability 1

$$\frac{|S_n|}{a(n)} \leqq \frac{M_{n_{k+1}}}{a(n_k)} \leqq \frac{a(n_{k+1})}{a(n_k)} \qquad \text{for } k \text{ large enough.} \tag{3.6}$$

Since the right-hand side of (3.6) is bounded, $k \to \infty$, (3.2) follows.

REMARK. The result does not require boundedness of the X_i's, only uniform boundedness of the νth absolute moments, i.e.,

$$E|X_i|^\nu \leqq g^{\frac{1}{2}\nu}(1) \qquad (\text{all } i). \tag{3.7}$$

The conclusion (3.2) improves as ν increases. This is possible because Lemma B, in contrast with the appropriate generalization of Lemma A to the case $\nu > 2$, does not have a factor $(\ln n)^\nu$ in the bound on $E(M_{a,n}^\nu)$. With such a factor present, the right-hand side of (3.2) cannot be improved beyond "$O(g^{\frac{1}{2}}(n)(\ln n)^{1+\varepsilon})$ for any $\varepsilon > 0$."

The LIL implies the following SLLN.

COROLLARY 3.1. (SLLN). *Let* $\nu > 2$. *Suppose that*

$$E|S_{a,n}|^\nu \leqq g^{\frac{1}{2}\nu}(n) \qquad (\textit{all } a \geqq 0, \textit{ all } n \geqq 1), \tag{3.1}$$

where $g(n)$ *is nondecreasing,* $2g(n) \leqq g(2n)$, $g(n)/g(n+1) \to 1$ *as* $n \to \infty$, *and* $g(2n)/g(n)$ *is bounded. If* $g(n)$ *satisfies*

$$g(n) = o\left(\frac{n^2}{(\ln n)^{2/\nu}(\ln \ln n)^{4/\nu}}\right), \qquad n \to \infty, \tag{3.2}$$

then, with probability 1, $S_n/n \to 0$.

Of particular importance and generality is the case given by

$$g(n) = An^\delta, \tag{3.8}$$

where $A < \infty$ and $1 \leqq \delta \leqq 2$. In (3.8), we require $\delta \geqq 1$ in order that $2g(n) \leqq g(2n)$ hold. The restriction $\delta \leqq 2$ may be assumed without loss of generality, since by Minkowski's inequality (3.7) implies a condition of form (3.1) with the right-hand side of order n^ν. Clearly the $g(n)$ in (3.8) is nondecreasing, $g(n) \sim g(n+1)$ and $g(2n)/g(n)$ is bounded. Thus a condition of type (3.1) for *some* function $g(n)$ implies a similar condition (though possibly milder) with $g(n)$ of the form (3.8). The lower that δ may be taken ($\geqq 1$), the more stringent is (3.1) and the more powerful are the conclusions obtained. Specifically, we have

COROLLARY 3.2. *Let* $\nu > 2$. *Suppose that*

$$E|S_{a,n}|^\nu \leqq An^{\frac{1}{2}\nu\delta} \qquad (\textit{all } a \geqq 0, \textit{ all } n \geqq 1), \tag{3.9}$$

where A $< \infty$ *and* $1 \leqq \delta \leqq 2$. *Then, with probability* 1,

$$|S_n| = O(n^{\frac{1}{2}\delta}(\ln n)^{1/\nu}(\ln \ln n)^{2/\nu}), \qquad n \to \infty. \tag{3.10}$$

If, further, $\delta < 2$, *then* $S_n/n \to 0$ $(n \to \infty)$ *with probability* 1.

The most stringent case, $\delta = 1$, applies under various different dependence restrictions (weaker than independence), given that the moments $E|X_i|^v$ are uniformly bounded. Thus, under the latter assumption, we have, with probability 1,

$$|S_n| = O(n^{\frac{1}{2}}(\ln n)^{1/v}(\ln\ln n)^{2/v}), \qquad n \to \infty, \tag{3.11}$$

if $\{X_i\}$ is any of the following well-known types of sequence:

(i) a sequence of mutually independent rv's;
(ii) a strictly stationary sequence satisfying a certain mixing condition weaker than m-dependence;
(iii) a stationary Markov sequence satisfying Doeblin's condition;
(iv) a sequence multiplicative of order v, where v is even;
(v) a sequence of martingale differences.

In the above, the "conventional" definitions are intended. Detailed discussion and references may be found in [3], [8], [12] and [14].

4. LIL for uniformly bounded variables. Here we assume

$$|X_i| \leqq B \qquad \text{(all } i\text{)}, \tag{4.1}$$

in which case for each $v > 0$ a relation of form

$$E|S_{a,n}|^v \leqq An^{\frac{1}{2}v\delta} \qquad \text{(all } a \geqq 0, \text{ all } n \geqq 1) \tag{4.2}$$

is satisfied, where $A < \infty$ and $1 \leqq \delta \leqq 2$ and A and δ may depend upon v. The immediate implication of the preceding section is that for each $v > 0$ there exists δ, $1 \leqq \delta \leqq 2$, such that, with probability 1,

$$|S_n| = O(n^{\frac{1}{2}\delta}(\ln n)^{1/v}(\ln\ln n)^{2/v}), \qquad n \to \infty. \tag{4.3}$$

If δ does not depend upon v (e.g., in cases that $\delta = 1$), we obtain, by letting $v \to \infty$ in (4.3), that for any $\varepsilon > 0$, with probability 1,

$$|S_n| = O(n^{\frac{1}{2}\delta}(\ln n)^{\varepsilon}), \qquad n \to \infty. \tag{4.4}$$

In particular, (4.4) is true with $\delta = 1$ for any $\varepsilon > 0$ in the case of a bounded sequence satisfying any of (i)–(v) mentioned at the end of Section 3. However, under some of these dependence restrictions, it is possible to obtain a stronger conclusion by the use of an appropriate probability inequality on S_n. For example, in Hoeffding [6] it is shown that (4.6) below holds in the case of mutually independent rv's satisfying (4.1) and it is pointed out that his argument extends to the case of X_i's being martingale differences. In [14], it is shown that, more generally, (4.6) holds if the X_i's are multiplicative rv's (defined below). Inequalities similar to (4.6) can certainly be developed under other kinds of dependence restriction. Thus the following result has a broad scope of application.

THEOREM 4.1. *Let $|X_i| \leqq B$ (all i). Suppose that for any $v > 0$, there exists $A_v < \infty$ such that*

$$E|S_{a,n}|^v \leqq A_v n^{\frac{1}{2}v} \qquad (\textit{all } a \geqq 0, \textit{ all } n \geqq 1). \tag{4.5}$$

Suppose, further, that

$$P[|S_n| > t] \leqq 2e^{-t^2/2B^2n} \qquad (t > 0). \tag{4.6}$$

Then, for any $\theta > 2B^2$, *with probability* 1,

$$|S_n| \leqq (\theta n \ln\ln n)^{\frac{1}{2}}, \qquad \textit{for all } n \textit{ large enough.} \tag{4.7}$$

PROOF. Let $\theta > 2B^2$. By (4.6), we have

$$P[|S_n| > (\theta n \ln\ln n)^{\frac{1}{2}}] \leqq 2(\ln n)^{-\theta/2B^2}. \tag{4.8}$$

Put $n_k = [\exp k^a]$, where $[\cdot]$ denotes integer part. Since $\theta > 2B^2$, we may choose $2B^2/\theta < a < 1$. Then $a\theta/2B^2 > 1$ and (Borel–Cantelli)

$$|S_{n_k}| < (\theta n_k \ln\ln n_k)^{\frac{1}{2}}, \qquad \text{for all } k \text{ large enough,} \tag{4.9}$$

with probability 1.

Since $a < 1$, there exists a value of $v > 2$ such that $a < (v-2)/v$, in which case $(1-a)(\frac{1}{2}v) > 1$. For such a value of v, let us utilize (4.5) in conjunction with Lemma B of Section 3. It follows that, for some $M < \infty$, and each k,

$$E\left[\max_{n_k \leqq n \leqq n_{k+1}} \left|\frac{S_n - S_{n_k}}{(n_k \ln\ln n_k)^{\frac{1}{2}}}\right|^v\right] \leqq M\frac{(n_{k+1}-n_k)^{\frac{1}{2}v}}{n_k^{\frac{1}{2}v}(\ln\ln n_k)^{\frac{1}{2}v}}. \tag{4.10}$$

The right-hand side of (4.10) is $O(k^{-(1-a)(\frac{1}{2}v)}(\ln k)^{\frac{1}{2}v})$, so that the infinite sum of the left-hand sides in (4.10) converges. Therefore, with probability 1,

$$\max_{n_k \leqq n \leqq n_{k+1}} \frac{|S_n - S_{n_k}|}{(n_k \ln\ln n_k)^{\frac{1}{2}}} \to 0, \qquad k \to \infty. \tag{4.11}$$

From (4.9) and (4.11) it follows that, for any $\varepsilon > 0$, with probability 1,

$$|S_n| \leqq (\theta+\varepsilon)^{\frac{1}{2}}(n \ln\ln n)^{\frac{1}{2}}, \qquad \text{for all } n \text{ large enough.} \tag{4.12}$$

Since $\theta + \varepsilon$ may be chosen arbitrarily close ($>$) to $2B^2$, the conclusion of the theorem is proved.

As a particular case, consider a sequence $\{X_i\}$ of *multiplicative* rv's:

$$E(X_{i_1} \cdots X_{i_k}) = 0 \qquad \text{if} \quad k \geqq 1 \quad \text{and} \quad i_1 < \cdots < i_k. \tag{4.13}$$

This condition is stronger than mutual orthogonality but includes the case of a sequence of martingale differences and the case of mutually independent rv's (when the expectations in (4.13) exist). As mentioned above, (4.5) and (4.6) hold for a bounded multiplicative sequence. Hence

COROLLARY 4.1. *Let* $\{X_i\}$ *be a sequence of bounded* ($|X_i| \leqq B$) *multiplicative* rv's. *Then, for any* $\theta > 2B^2$, *with probability* 1,

$$|S_n| \leqq (\theta n \ln\ln n)^{\frac{1}{2}}, \qquad \textit{for all } n \textit{ large enough.} \tag{4.7}$$

5. Convergence rates in the SLLN and LIL. The following theorem is part of Theorem 1 of Katz [9], which generalizes earlier results due to Erdös [4] and Spitzer [15].

THEOREM (Katz–Spitzer–Erdös). *Let* $v \geqq 1$. *Let* $\{X_i\}$ *be a sequence of independent and identically distributed* rv's, *with* $E|X_1|^v < \infty$. *Then, for each* $\varepsilon > 0$,

$$\sum n^{v-2} P[|S_n/n| > \varepsilon] < \infty. \tag{5.1}$$

This result was proved by complicated methods based on those of [4]. However, its implication in the case of bounded variables can be obtained more easily and it will be helpful in the sequel to see how. The implication in question is, of course, that for each $\varepsilon > 0$,

$$\sum n^{\alpha} P[|S_n/n| > \varepsilon] < \infty, \qquad \text{for every choice of } \alpha. \tag{5.2}$$

Now, as mentioned earlier, Hoeffding [6] proves the probability inequality,

$$P[|S_n| > t] \leqq 2e^{-t^2/2B^2n} \qquad (t > 0), \tag{5.3}$$

for independent and uniformly bounded ($|X_i| \leqq B$) rv's. With $t = n\varepsilon$, (5.3) immediately yields (5.2).

Actually, the SLLN concerns $P[\sup_{k \geqq n}|S_k/k| > \varepsilon]$ more than $P[|S_n/n| > \varepsilon]$. It is of interest to obtain results of type (5.1) involving the former probabilities in place of the latter. It also is of interest, as previously in this paper, to obtain the conclusions under dependence restrictions weaker than independence.

We shall assume that $\{X_i\}$ satisfies

$$E|S_{a,n}|^v \leqq A_v n^{\frac{1}{2}v} \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1), \tag{5.4}$$

where $v > 2$ and $A_v < \infty$. By Lemma B, this implies the existence of a constant $K_v < \infty$ such that

$$E(M_{a,n}^v) \leqq K_v n^{\frac{1}{2}v} \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1). \tag{5.5}$$

By Markov's inequality, (5.4) and (5.5) imply, respectively, for $\varepsilon > 0$,

$$P[|S_{a,n}| > n\varepsilon] \leqq A_v \varepsilon^{-v} n^{-\frac{1}{2}v} \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1) \tag{5.6}$$

and

$$P[M_{a,n} > n\varepsilon] \leqq K_v \varepsilon^{-v} n^{-\frac{1}{2}v} \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1). \tag{5.7}$$

THEOREM 5.1. *Let* $v > 2$. *Let* $\{X_i\}$ *satisfy* (5.6) *and* (5.7). *Then for each* $\varepsilon > 0$ *there exists a constant* $C_\varepsilon < \infty$ (*depending on* A_v *and* K_v) *such that*

$$P[\sup_{k \geqq n}|S_k/k| > \varepsilon] \leqq C_\varepsilon n^{-\frac{1}{2}v} \qquad (\textit{all } n \geqq 1). \tag{5.8}$$

PROOF. Trivially we have

$$P[\sup_{k \geqq n}|S_k/k| > \varepsilon] \leqq \sum_{j=0}^{\infty} P[\max_{k_j \leqq k \leqq k_{j+1}} |S_k/k| > \varepsilon], \tag{5.9}$$

where $k_j = n2^j$ $(j = 0, 1, \cdots)$. And

$$\begin{aligned} &P[\max_{k_j \leqq k \leqq k_{j+1}} |S_k/k| > \varepsilon] \\ &\qquad \leqq P[|S_{k_j}/k_j| > \tfrac{1}{2}\varepsilon] + P[\max_{k_j \leqq k \leqq k_{j+1}} |(S_k - S_{k_j})/k_j| > \tfrac{1}{2}\varepsilon]. \end{aligned} \tag{5.10}$$

The use of (5.6) and (5.7) in the right-hand side yields

$$P[\sup_{k \geqq n} |S_k/k| > \varepsilon] \leqq (A_\nu + K_\nu)(\tfrac{1}{2}\varepsilon)^{-\nu} n^{-\frac{1}{2}\nu} \textstyle\sum_{j=0}^{\infty} 2^{-\frac{1}{2}\nu j}, \tag{5.11}$$

so that (5.8) holds with $C_\varepsilon = (A_\nu + K_\nu)(\frac{1}{2}\varepsilon)^{-\nu}(1-2^{-\frac{1}{2}\nu})^{-1}$.

Thus we have

COROLLARY 5.1.1. *Let $\nu > 2$. Let $\{X_i\}$ satisfy* (5.6) *and* (5.7). *Then, for each $\varepsilon > 0$,*

$$\sum \frac{n^{\frac{1}{2}\nu - 1}}{\ln^2 n} P[\sup_{k \geqq n} |S_k/k| > \varepsilon] < \infty. \tag{5.12}$$

The implications of Theorem 5.1 under moment restrictions of type (5.4) are of particular interest. As discussed in Section 3, (5.4) is a dependence restriction satisfied in a variety of stochastic processes.

COROLLARY 5.1.2. *Let $\nu > 2$. Let $\{X_i\}$ satisfy* (5.4). *Then, for each $\varepsilon > 0$,* (5.12) *holds.*

In particular, for bounded variables, we have a conclusion substantially better than (5.2).

COROLLARY 5.1.3. *Let $\{X_i\}$ be uniformly bounded ($|X_i| \leqq B$) and satisfy* (5.4) *for each $\nu > 0$. Then, for each $\varepsilon > 0$,*

$$\sum n^\alpha P[\sup_{k \geqq n} |S_k/k| > \varepsilon] < \infty, \qquad \textit{for every choice of } \alpha. \tag{5.13}$$

It should be noted that under a restriction milder than (5.4), e.g., with $A_\nu n^{\frac{1}{2}\nu\delta}$ for $1 < \delta < 2$ on the right-hand side, analogous conclusions may be derived.

Finally, as regards convergence rates in the SLLN, we consider the improvements in the above which are possible when the probability inequality (5.3) may be used.

THEOREM 5.2. *Let $\{X_i\}$ be uniformly bounded ($|X_i| \leqq B$) and satisfy the probability inequality* (5.3). *Then for each $\varepsilon > 0$ there exist positive constants $C_\varepsilon < \infty$ and $\rho_\varepsilon < 1$ such that*

$$P[\sup_{k \geqq n} |S_k/k| > \varepsilon] \leqq C_\varepsilon \rho_\varepsilon^n \qquad (\textit{all } n \geqq 1). \tag{5.14}$$

PROOF. Trivially we have

$$P[\sup_{k \geqq n} |S_k/k| > \varepsilon] \leqq \textstyle\sum_{k=n}^{\infty} P[|S_k| > k\varepsilon]. \tag{5.15}$$

Applying (5.3) with $t = k\varepsilon$ in the kth term of the above series,

$$P[\sup_{k \geqq n} |S_k/k| > \varepsilon] \leqq 2 \textstyle\sum_{k=n}^{\infty} e^{-\varepsilon^2 k/2B^2}. \tag{5.16}$$

$$= 2e^{-\varepsilon^2 n/2B^2}(1-e^{-\varepsilon^2/2B^2})^{-1}. \tag{5.17}$$

Thus (5.14) holds with $\rho_\varepsilon = e^{-\varepsilon^2/2B^2}$ and $C_\varepsilon = 2/(1-\rho_\varepsilon)$.

COROLLARY 5.2.1. *Let $\{X_i\}$ be uniformly bounded ($|X_i| \leqq B$) and satisfy the probability inequality* (5.3). *Then, for each $\varepsilon > 0$,*

$$\sum \alpha^n P[\sup_{k \geqq n} |S_k/k| > \varepsilon] < \infty, \qquad \textit{for any} \quad \alpha < e^{\varepsilon^2/2B^2}. \tag{5.18}$$

The conclusion (5.18) is, of, course considerably sharper than (5.13), which in turn is sharper than (5.2). The scope of application of Corollary 5.2.1 is noted by recalling, as discussed in the previous section, that (5.3) holds for a sequence of uniformly bounded multiplicative rv's.

Turning now to convergence rates in the LIL, let us note the following result, given as Theorem 2 of Davis [2].

THEOREM (Davis). *Let $\{X_i\}$ be a sequence of independent and identically distributed* rv's, *with $E(X_1^2) = 1$ and $E(X_1^2 \ln\ln |X_1|) < \infty$. Then, for each $\theta > 2$,*

$$\sum \frac{1}{n \ln n} P\left[\sup_{k \geqq n} \frac{|S_k|}{(\theta k \ln\ln k)^{\frac{1}{2}}} > 1\right] < \infty. \tag{5.19}$$

As in foregoing discussions, we are interested in a parallel result to the above under milder dependence and stationarity assumptions but making use of moment restrictions of order higher than second.

THEOREM 5.3. *Let $v > 2$. Let $\{X_i\}$ satisfy* (5.4). *Then, for any choice of α and β satisfying $0 \leqq \beta < \alpha v - 1 \leqq 1$,*

$$\sum \frac{1}{n(\ln n)^{1-\beta}} P\left[\sup_{k \geqq n} \frac{|S_k|}{k^{\frac{1}{2}}(\ln k)^{\alpha}} > 1\right] < \infty. \tag{5.20}$$

In this result, the stationarity assumption is relaxed to that of uniform boundedness of $E|X_i|^v$ and the independence assumption is relaxed to the requirement that $E|S_{a,n}|^v$ be $O(n^{\frac{1}{2}v})$ uniformly in a. Moreover, the factor $(n \ln n)^{-1}$ in the nth term of (5.19) is improved to $n^{-1}(\ln n)^{\beta-1}$, where β is suitably chosen in $[0, 1)$. On the other hand, the factor $(\theta \ln\ln k)^{\frac{1}{2}}$ in the expression (5.19) become replaced by a less sharp factor $(\ln k)^{\alpha}$, where α must be $> 1/v$, and a moment restriction of order higher than 2 is imposed. The best restriction on α, namely that $\alpha > 0$, occurs if v may be chosen arbitrarily large, as in the following

COROLLARY 5.3.1. *Let $\{X_i\}$ be uniformly bounded* ($|X_i| \leqq B$) *and satisfy* (5.4) *for each $v > 0$. Then* (5.20) *holds for each choice of $\alpha > 0$ and $0 < \beta < 1$.*

PROOF OF THEOREM 5.3. It is easily seen that the series in (5.20) is bounded from above by the series

$$\sum \frac{1}{(\ln 2^j)^{1-\beta}} P\left[\sup_{k \geqq 2^j} \frac{|S_k|}{k^{\frac{1}{2}}(\ln k)^{\alpha}} > 1\right] \tag{5.21}$$

and therefore also by

$$C \sum \frac{1}{j^{1-\beta}} \sum_{i=j}^{\infty} P\left[\max_{2^i \leqq k \leqq 2^{i+1}} \frac{|S_k|}{k^{\frac{1}{2}}(\ln k)^{\alpha}} > 1\right], \tag{5.22}$$

for a suitable constant C. By an argument similar to the proof of Theorem 5.1, and for a suitable constant C_1, the series in (5.22) is less than

$$C_1 \sum \frac{1}{j^{1-\beta}} \sum_{i=j}^{\infty} \frac{1}{i^{\alpha v}}. \tag{5.23}$$

If $\beta = 0$, the latter is less than $C_2 \sum (\ln i) i^{-\alpha\nu}$, for some constant C_2. If $0 < \beta < 1$, it is less than $C_2 \sum i^{\beta - \alpha\nu}$, for some constant C_2. In either case, therefore, we have convergence of the series in (5.23) and hence of that in (5.20).

6. Complete convergence of normed S_n and M_n. Following Hsu and Robbins [7], we say that a sequence of rv's $\{\xi_n\}$ *converges completely to zero* if the series $\sum P|\xi_n| > \varepsilon]$ converges for each $\varepsilon > 0$. For further discussion and references concerning the following remarks, see Hsu and Robbins [7], Erdös [4], Chow [1] and Stout [16].

To establish a frame of reference, we state the following result.

THEOREM (Hsu–Robbins–Erdös–Chow–Stout). *Let $\nu \geqq 2$. Let $\{X_i\}$ be a sequence of independent and identically distributed* rv's, *with $E|X_1|^\nu < \infty$. Then*:

(i) *In the case* $2 \leqq \nu < 4$,

$$\sum P\left[\frac{|S_n|}{n^{2/\nu}} > \varepsilon\right] < \infty, \qquad \textit{all} \quad \varepsilon > 0. \tag{6.1}$$

(ii) *In the case* $\nu \geqq 4$,

$$\sum P\left[\frac{|S_n|}{n^{\frac{1}{2}}(\ln n)^{\frac{1}{2}} g(n)} > \varepsilon\right] < \infty, \quad \textit{all}\ \varepsilon > 0, \quad \textit{if}\ g(n) \to \infty. \tag{6.2}$$

The proof is contained in [1] and [16]: statement (i) follows from Theorem 6 in [1], while statement (ii) follows from Corollary 1 and subsequent discussion in [16].

It is noted that statement (ii) is sharp since in the case of X_i's having a common normal distribution the series in (6.2) cannot converge unless $g(n) \to \infty$.

Now let us consider results of the above type under less stringent dependence and stationarity assumptions. As in the previous section, suppose that $E|X_i|^\nu$ is uniformly bounded and, further, that

$$E|S_{a,n}|^\nu \leqq A_\nu n^{\frac{1}{2}\nu} \qquad (\text{all } a \geqq 0, \text{ all } n \geqq 1), \tag{6.3}$$

where $\nu \geqq 2$ and $A_\nu < \infty$. Then immediately we obtain by the Markov inequality that

$$\sum P\left[\frac{|S_n|}{n^{\frac{1}{2}+1/\nu}(\ln n)^{2/\nu}} > \varepsilon\right] < \infty, \tag{6.4}$$

which, of course, is not as sharp as (6.1). However, if $\nu > 2$ is assumed in (6.3), then by Lemma B we have the stronger conclusion

$$\sum P\left[\frac{M_n}{n^{\frac{1}{2}+1/\nu}(\ln n)^{2/\nu}} > \varepsilon\right] < \infty, \tag{6.5}$$

giving information not contained in (6.1), (6.2).

In the case of a sequence $\{X_i\}$ for which $E|X_i|^\nu$ is uniformly bounded for all ν and, further, (6.3) holds for all ν, we may conclude, therefore,

$$\sum P[M_n/n^\alpha > \varepsilon] < \infty, \qquad \text{for each} \quad \alpha > \tfrac{1}{2}, \tag{6.6}$$

a result that compares nicely with (6.2). Thus (6.6) holds for strictly stationary sequences $\{X_i\}$ satisfying a certain mixing condition (see Ibragimov [8]) and satisfying $E|X_1|^\nu < \infty$, all ν.

Confining attention now to a uniformly bounded ($|X_i| \leqq B$) sequence, suppose that the probability inequality

$$P[|S_n| > t] \leqq 2e^{-t^2/2B^2 n} \qquad (t > 0) \tag{6.7}$$

is satisfied. Then it follows easily that the sharp result (6.2) holds. Stout [16] notes that most of his results remain true if the X_i's are not mutually independent but form a sequence of martingale differences. His Corollary 4 implies that (6.2) holds if $\{X_i\}$ is a sequence of uniformly bounded martingale differences. Now, as remarked in Section 4, inequality (6.7) is satisfied if $\{X_i\}$ is a sequence of uniformly bounded multiplicative rv's. Hence (6.2) holds for such a sequence. Obviously, further relaxations of the conditions sufficient for (6.2) to hold can be obtained as a simple consequence of proving the inequality (6.7) or like inequalities for a broader class of sequences.

REFERENCES

[1] CHOW, Y. S. (1966). Some convergence theorems for independent random variables. *Ann. Math. Statist.* **37** 1482–1493.

[2] DAVIS, JAMES AVERY (1968). Convergence rates for the law of the iterated logarithm. *Ann. Math. Statist.* **39** 1479–1485.

[3] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[4] ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20** 286–291.

[5] FELLER, WILLIAM (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

[6] HOEFFDING, WASSILY (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

[7] HSU, P. L. and ROBBINS, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA* **33** 25–31.

[8] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.* **7** 349–382.

[9] KATZ, MELVIN (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34** 312–318.

[10] LAMPERTI, JOHN (1966). *Probability.* Benjamin, New York.

[11] RÉVÉSZ, PÁL (1968). *The Laws of Large Numbers.* Academic Press, New York.

[12] SERFLING, R. J. (1968). Contributions to central limit theory for dependent variables. *Ann. Math. Statist.* **39** 1158–1175.

[13] SERFLING, R. J. (1970). Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* **41** 1227–1234.

[14] SERFLING, R. J. (1969). Probability inequalities and convergence properties for sums of multiplicative random variables. (Submitted.)

[15] SPITZER, FRANK (1956). A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.

[16] STOUT, WILLIAM F. (1968). Some results on the complete and almost sure convergence of linear combinations of independent random variables and martingale differences. *Ann. Math. Statist.* **39** 1549–1562.