*An Extension of the Law of the Iterated Logarithm to Variables without Variance**

WILLIAM FELLER

Dedicated to SALOMON BOCHNER

1. Introduction. Throughout this paper $X, X_1, X_2, \dots$ will denote a sequence of mutually independent random variables with a common distribution F . In order not to obscure the main point we shall assume F to be *symmetric*, that is $F(-x) = 1 - F(x)$ at points of continuity. The results carry over to the general situation by the standard symmetrization procedure. As usual, we put $S_n = X_1 + \dots + X_n$.

When $E(X^2) = 1$ the central limit theorem asserts that the distribution of $S_n/\sqrt{n}$ tends to the standard normal distribution, whereas by the law of the iterated logarithm

$$(1.1) \quad \limsup \frac{S_n}{(2n \log \log n)^{1/2}} = 1$$

with probability one. From the simple lemma 1 below it is obvious that in the absence of a finite variance this lim sup is infinite. [This observation was made by V. Strassen and proved by deeper tools]. However, the central limit theorem holds also when $E(X^2) = \infty$ provided only that

$$(1.2) \quad H(x) = \int_{|t| \leq x} t^2 F(dt)$$

varies slowly as $x \rightarrow \infty$. If ρ_n is defined as the largest root of

$$(1.3) \quad \rho_n^2 = nH(\rho_n),$$

then the distribution of S_n/ρ_n tends to the standard normal. In this situation it is natural to ask whether there exist any norming constants a_n such that

$$(1.4) \quad \limsup S_n/a_n = 1$$

with probability one and, in particular, whether this is true with

$$(1.5) \quad a_n = \rho_n(2 \log \log \rho_n)^{1/2}.$$

* Research connected with a Project for research in probability at Princeton University, supported by the Army Research Office.

This problem has an unexpected solution because it turns out that in the absence of variances the applicability of the *central limit theorem* is neither necessary nor sufficient for the validity of the relations (1.4)–(1.5).

Our main results are stated in section 7. Neglecting freak phenomena that can occur when the distribution F and the sequence $\{a_n\}$ exhibit extremely wild fluctuations the typical situation may be described roughly as follows.

If there exists any sequence $\{a_n\}$ for which (1.4) holds, then (1.5) will do. Necessary and sufficient for this is that

$$(1.6) \quad \int_0^\infty \frac{H(dx)}{H(x) \log \log x} < \infty.$$

This condition is rather stringent, but it does not imply the slow variation of H . When it fails, the $\limsup$ in (1.4) necessarily equals either zero or infinity for any reasonably regular sequence $\{a_n\}$. In fact, if the integral (1.6) diverges the two events

$$(1.7) \quad \{|S_n| > a_n \text{ i.o.}\} \quad \text{and} \quad \{|X_n| > a_n \text{ i.o.}\}$$

have the same probability (zero and one being the only possible values by virtue of the Hewitt–Savage theorem). This equivalence may be taken as an indication that dominating values of S_n are as a rule due to the influence of one single term X_k . To make this observation more precise, put

$$(1.8) \quad X_n^M = \max [X_1, \dots, X_n] \quad \text{and} \quad {}^{(1)}S_n = S_n - X_n^M.$$

Very roughly speaking, it will be shown that when the integral (1.6) diverges and the constants a_n are chosen so that the series $\sum P|X_n| > a_n$ barely diverges, then with probability one

$$(1.9) \quad \limsup S_n/a_n = \infty \quad \text{but} \quad \limsup {}^{(1)}S_n/a_n = 0.$$

Theorem 2 shows in particular that even within the domain of attraction of the normal distribution the generalized *law of the iterated logarithm* as expressed by (1.4)–(1.5) may fail due to the influence one single component X_k . With the norming constants (1.5) the $\limsup$ in (1.4) is then infinite, but the relation (1.4) holds if S_n is replaced by ${}^{(1)}S_n$.

[The equivalence of the events in (1.7) was discovered in Feller (1946) assuming that F has moments of all orders $< \alpha < 2$ but none of order $> \alpha$. Feller had to impose stronger regularity conditions on $\{a_n\}$. The same phenomenon was discussed by P. Lévy (1952 and 1953) with relaxed conditions on the a_n but additional restrictions on F . The present paper generalizes these results, but its main purpose is to explore the analogue to the law of the iterated logarithm and the influence of the large components when the distribution F is within the domain of attraction of the normal distribution or close to it.]

2. The basic assumptions. The sequence $\{a_n\}$ occurring in the limit relation (1.4) is not uniquely determined since the relation remains valid when the

values a_{n_k} are changed for a sufficiently sparse subsequence $\{n_k\}$. A natural class of norming constants is defined by the conditions that

$$(2.1) \quad a_n / \sqrt{n} \uparrow \infty$$

and there exist a constant M such that

$$(2.2) \quad \frac{a_{2n}}{a_n} < M.$$

Condition (2.1) serves merely to avoid trivialities. Condition (2.2) is used mainly in section 6 to prove that if (1.4) is true for some sequence $\{a_n\}$ then also for the special sequence defined by (1.5). Without (2.2) this assertion is false and $\{a_n\}$ must be replaced by the less pleasing sequence defined in (6.1).

Trite complications can arise when the distribution F varies extremely erratically. (Example *b* shows what can happen if the growth of F is concentrated on a very rapidly increasing sequence $\{x_n\}$). In order not to obscure the main point by side issues we shall assume that there exists a constant M such that for large x

$$(2.3) \quad \frac{H(2x)}{H(x)} < M.$$

This is obviously equivalent to

$$(2.4) \quad \frac{H(tx)}{H(x)} < Mt^\mu$$

for all x and t sufficiently large. As is well known, for F to belong to some domain of attraction it is necessary that as $x \rightarrow \infty$ the left side in (2.4) actually tends to t^μ with $\mu < 2$. In (2.4) values $\mu > 2$ are permitted, and so this condition is mild indeed.

It is not necessary to impose any condition on the distribution F if one assumes instead that the behaviour of $\{a_n\}$ is tuned in a natural way to that of H . For example, it is natural to assume that the ratios a_n/ρ_n increase, and this condition is sufficient. As a matter of fact, a perusal of the arguments shows that (2.3) *can be replaced by the condition that there exists a power σ such that the ratios $a_n/H^\sigma(a_n)$ increase*. All naturally defined norming constants satisfy this condition with $\sigma = 1$ and so in practice our results apply to arbitrary distributions. The condition (2.3) serves to simplify formulations.

3. Preparatory lemmas. We shall say that the sequence $\{X'_n\}$ is obtained from $\{X_n\}$ by (symmetric) truncation at τ_n and t_n if

$$(3.1) \quad \begin{aligned} X'_n &= X_n && \text{when } \tau_n < |X_n| \leq t_n \\ &0 && \text{otherwise.} \end{aligned}$$

The partial sums will be denoted by S'_n .

Lemma 3.1. For arbitrary (symmetric) truncation and arbitrary constants a_n , if

$$(3.2) \quad P\{S'_n > a_n \text{ i.o.}\} = 1$$

then also

$$(3.3) \quad P\{S_n > a_n \text{ i.o.}\} = 1.$$

Proof. Because of symmetry the two events

$$(3.4) \quad \{S'_n > a_n, S_n - S'_n \geq 0 \text{ i.o.}\} \quad \text{and} \quad \{S'_n > a_n, S_n - S'_n \leq 0 \text{ i.o.}\}$$

have the same probability, which equals either zero or one by the Hewitt-Savage theorem. Now the event in (3.2) is the union of the two events in (3.4), whereas the event in (3.3) contains the first event in (3.4), and so (3.2) implies (3.3).

An immediate consequence of this lemma is that *the certainty of the second relation in (1.7) is necessary for the first.*

Lemma 3.2. If (2.1) holds, the series

$$(3.5) \quad \sum P\{X_n > \lambda a_n\}$$

either diverges for all $\lambda > 0$ or else it converges for all $\lambda > 0$.

Proof. Assume divergence when $\lambda = \lambda_0$. The divergence for $\lambda < \lambda_0$ follows from the obvious monotonicity. If $\lambda > \lambda_0$ choose an integer $r > \lambda^2/\lambda_0^2$. Then $\lambda_0 a_{nr} > \lambda a_n$ and hence

$$\sum P\{X_n > \lambda a_n\} \geq \sum P\{X_{nr} \geq \lambda_0 a_{nr}\} \geq r^{-1} \sum P\{X_n > \lambda_0 a_n\},$$

and so the series on the left diverges also for $\lambda > \lambda_0$.

For any sequence $\{a_n\}$ we define a function a_t for all $t > 0$ by monotone interpolation between the integers, and put

$$(3.6) \quad a_t^2 = tA(t).$$

Thus (2.1) implies that $A \geq 0$ and that A is monotone.

Lemma 3.3. The series (3.5) converges if, and only if,

$$(3.7) \quad I = \int_0^\infty \frac{x^2 F(dx)}{A(x)} < \infty.$$

Proof. When $\lambda = 1$ the series (3.5) converges obviously if, and only if,

$$(3.8) \quad \sum n \int_{a_{n-1}}^{a_n} F(dx) < \infty.$$

Between a_{n-1} and a_n the integrand in (3.7) varies monotonically from $n-1$ to n , and so (3.7) and (3.8) are equivalent.

4. The first truncation. We saw that the convergence of the integral (3.7)

implies the convergence of the series (3.5). This in turn implies that with probability one the relation $|X_n| > a_n$ will take place only for finitely many subscripts n , and thus we can limit our attention to the variables obtained by truncation between 0 and a_n . Such a variable will be decomposed into a sum of three variables obtained by further truncation. In this section we dispose of the largest values. For this purpose we introduce the abbreviations

$$(4.1) \quad v(t) = A(t)H^{-1}(t)$$

$$(4.2) \quad b_n = a_n v^{-\eta}(a_n)$$

where $\eta > 0$ is a constant. In the sequel we denote by X'_n the variables obtained by truncation at b_n and ϵa_n . The partial sums will be denoted by S'_n while $^{(1)}S'_n$ denotes (as in (1.8)) the random variable obtained by subtracting from S'_n the largest among the variables $X'_1, \dots, X'_n$.

We shall be interested only in situations where

$$(4.3) \quad v(a_n) \rightarrow \infty.$$

Lemma 4.1. *Assume the convergence of the integral (3.7) and that the conditions (2.1), (2.3), and (4.3) hold. Then there exists a constant η_0 such that for $\eta < \eta_0$ with probability one*

$$(4.4) \quad S'_n = o(a_n) \quad \text{and} \quad ^{(1)}S'_n = o(b_n).$$

Proof. Clearly

$$(4.5) \quad b_n^2 = nH^{2\eta}(a_n)A^{1-2\eta}(a_n);$$

since both H and A are monotone it follows for $\eta \leq \frac{1}{2}$ that $b_n \uparrow \infty$. We can therefore define recursively an increasing sequence of integers n_k such that

$$(4.6) \quad b_{n_k} < a_{n_{k-1}} \leq b_{n_{k+1}}.$$

To avoid double subscripts we put

$$(4.7) \quad a_{n_k} = \alpha_k, \quad b_{n_k} = \beta_k.$$

Let

$$(4.8) \quad \tau_k = \sum_{n_{k-1} < n \leq n_k} P\{X'_n \neq 0\}.$$

By Chebyshev's inequality

$$(4.9) \quad \begin{aligned} \tau_k &< \frac{n_k}{\beta_{k-1}^2} [H(\alpha_k) - H(\beta_k)] \\ &= \frac{\alpha_k^2 v^{2\eta}(\alpha_{k-1})}{\alpha_{k-1}^2} \cdot \frac{H(\alpha_k) - H(\beta_k)}{A(\alpha_k)}. \end{aligned}$$

Now v need not be monotone, but from (2.4) we get

$$(4.10) \quad \frac{v(\alpha_{k-1})}{v(\alpha_k)} \leq \frac{H(\alpha_k)}{H(\alpha_{k-1})} \leq M \cdot \left(\frac{\alpha_k}{\alpha_{k-1}} \right)^\mu.$$

Furthermore, by (4.6)

$$(4.11) \quad \alpha_{k-1} \geq \beta_k = \alpha_k v^{-\eta}(\alpha_k).$$

Putting for abbreviation $C = M^{2\eta}$ and $\gamma = 2\eta(1 + \mu\eta)$ we conclude from (4.9) that

$$(4.12) \quad \tau_k \leq C \cdot v^\gamma(\alpha_k) I(\beta_{k-1}, \alpha_k],$$

where $I(\]$ represents the measure attributed by the integral I to the given interval.

We need also the rougher estimate

$$(4.13) \quad \tau_k \leq C v^{\gamma-1}(\alpha_k)$$

obtained from (4.9) by neglecting $H(\beta_k)$. In view of (4.3) these two estimates imply that when $\gamma < \frac{1}{2}$ and k is sufficiently large

$$(4.14) \quad \tau_k^2 < I(\alpha_{k-2}, \alpha_k].$$

Since the integral I converges this means that

$$(4.15) \quad \sum \tau_k^2 < \infty \quad \text{and hence} \quad \sum \tau_k \tau_{k-1} < \infty.$$

It follows now from the Borel–Cantelli lemma that with probability one ultimately each block $n_{k-2} < n \leq n_k$ contains at most one component $X'_n \neq 0$, and this is by definition $< \epsilon a_n$. For given n let k be the integer such that $n_{k-1} < n \leq n_k$. Omitting, if necessary, the largest variable with a subscript $> n_{k-2}$ we find that with probability one for n sufficiently large

$$(4.16) \quad {}^{(1)}S'_n = S'_{n_{k-2}} < \epsilon(\alpha_{k-2} + \alpha_{k-4} + \dots).$$

But by (4.6)

$$\alpha_{k+1} = v(\alpha_{k+1})\beta_{k+1} \geq v(\alpha_{k+1})\alpha_{k-1}$$

and hence ultimately the right side in (4.16) is dominated by a geometric series with arbitrarily large ratio. The sum is therefore $O(\alpha_{k-2}) = o(b_n)$. Since $X'_n < \epsilon a_n$ and ϵ is arbitrary this implies (4.4).

Remark. If one assumes (as is natural) that v is monotone the condition (2.3) used in the estimate (4.10) becomes unnecessary. In this case the lemma is valid with $\eta_0 = \frac{1}{4}$.

5. The second truncation. In this section we investigate the contribution of the variables X''_n obtained by truncation at $c_n = a_n v^{-p}(a_n)$ and $b_n = a_n v^{-\eta}(a_n)$; here we preserved the notation of the preceding section and $p > \eta$ is an arbitrary constant.

Lemma 5.1. *Suppose $I < \infty$ and that the conditions (2.1)–(2.3) and (4.3) are satisfied. Then with probability one*

$$(5.1) \quad S''_n = o(a_n v^{-\delta}(a_n))$$

for all $\delta > 0$ sufficiently small. If v is montone $\delta < \eta/2$ suffices.

Proof. In view of (2.2) we can choose a sequence of integers n_k such that

$$(5.2) \quad M^{2k-1} < a_{n_k}^2 \leq M^{2k}.$$

We adhere to the abbreviations (4.7) and put similarly $\gamma_k = c_{n_k}$. Put

$$(5.3) \quad U_k = \sum_{n_{k-1} < n \leq n_k} X_n''$$

and denote by U_k^* the maximum of the partial sums on the right. Because of symmetry

$$(5.4) \quad P\{U_k > t\} \geq \frac{1}{2}P\{U_k^* > t\}.$$

On the other hand, the monotonicity of A and (2.3) imply that for $n < m$

$$(5.5) \quad \frac{v(a_n)}{v(a_m)} \leq \frac{H(a_m)}{H(a_n)} \leq M \left(\frac{a_m}{a_n} \right)^\mu$$

and hence for $n_{k-1} < n \leq n_k$ and $\mu\delta < 1$

$$(5.6) \quad \frac{a_n v^{-\delta}(a_n)}{\alpha_k v^{-\delta}(\alpha_k)} \geq M^{-\delta} \left(\frac{\alpha_{k-1}}{\alpha_k} \right)^{1-\mu\delta}$$

The right side remains bounded away from 0 and ∞ , and to prove the assertion (5.1) it suffices therefore to show that for any $\delta > 0$

$$(5.7) \quad \sum P\{U_k > \alpha_k v^{-\delta}(\alpha_k)\} < \infty.$$

This, in turn, is certainly true if

$$(5.8) \quad \sum \alpha_k^{-4} v^{4\delta}(\alpha_k) E(U_k^4) < \infty.$$

Put for abbreviation

$$(5.9) \quad V = \text{Var}(U_k) = \sum_{n_{k-1} < n \leq n_k} E(X_n''^2).$$

The variables occurring in the sum being symmetric and bounded by $\beta_k = \alpha_k v^{-\eta}(\alpha_k)$ one gets easily

$$(5.10) \quad E(U_k^4) \leq [\alpha_k^2 v^{-2\eta}(\alpha_k) + 3V]V.$$

Now trivially

$$(5.11) \quad V \leq n_k H(\alpha_k) = \alpha_k^2 v^{-1}(\alpha_k).$$

Using the relation (5.6) with δ replaced by p we get the alternative estimate

$$(5.12) \quad V \leq n_k A(\alpha_k) I(\epsilon \gamma_k, \beta_k) = \alpha_k^2 I(\epsilon \gamma_k, \beta_k),$$

valid when $\epsilon > 0$ is sufficiently small. Combining the two estimates we find that

$$(5.13) \quad E(U_k^4) \leq 4\alpha_k^4 v^{-2\eta}(\alpha_k) I(\epsilon \alpha_k v^{-p}(\alpha_k), \alpha_k).$$

Since $v(\alpha_k) \rightarrow \infty$ the factor ϵ can be dropped if it is compensated by increasing the exponent p , which is arbitrary. The convergence of the series (5.8) for $\delta < \eta/2$ is therefore guaranteed by the following lemma, which will conclude the proof.

Remark. The condition (2.3) was used in (5.5), and a sharper estimate is obtained by supposing instead that v is monotone.

Lemma 5.2. For arbitrary $p > 0$ and $\sigma > 0$

$$(5.14) \quad \sum v^{-\sigma}(\alpha_k) I(\alpha_k v^{-p}(\alpha_k), \alpha_k] < \infty.$$

Proof. The series may be rewritten in the form

$$(5.15) \quad \sum_k v^{-\sigma}(\alpha_k) \sum_j I(\alpha_{j-1}, \alpha_j]$$

where (for fixed k) the inner sum ranges over all j such that

$$(5.16) \quad \alpha_k v^{-p}(\alpha_k) \leq \alpha_j \leq \alpha_k.$$

But by (5.2) the ratio of two consecutive α_j lies between M and M^3 , and hence for any j satisfying (5.16)

$$(5.17) \quad 0 \leq k - j < C \log v(\alpha_k)$$

where $C = p/\log M$. Now

$$(5.18) \quad \frac{v(\alpha_j)}{v(\alpha_k)} \geq \frac{A(\alpha_j)}{A(\alpha_k)} \geq \frac{\alpha_j^2}{\alpha_k^2} \geq v^{-2p}(\alpha_k)$$

whereas (5.5) shows that

$$(5.19) \quad \frac{v(\alpha_j)}{v(\alpha_k)} \leq M \cdot \left(\frac{\alpha_k}{\alpha_j} \right)^\mu \leq M \cdot v^{p\mu}(\alpha_k).$$

We now change the order of summation in (5.15) to obtain

$$(5.20) \quad \sum_j I(\alpha_{j-1}, \alpha_j] \sum v^{-\sigma}(\alpha_k)$$

the inner summation extending over all k for which (5.16) holds. The preceding inequalities imply that the number of such terms does not exceed a fixed multiple of $\log v(\alpha_j)$ and that each term is dominated by a fixed negative power of $v(\alpha_j)$. The inner sum therefore tends to 0 as $j \rightarrow \infty$, and hence the series is dominated by a multiple of the integral I .

6. The central part. We turn our attention to the variables X_n^T obtained by truncation at 0 and $c_n = a_n v^{-p}(a_n)$. The condition (2.2) guarantees that for fixed $\delta > 0$ the ratios $a_n/a_{\delta n}$ remain bounded. Omitting the first δn variables it follows easily that

$$(6.1) \quad s_n^2 = \text{Var}(S_n^T) \geq (1 - \delta)nH(a_n v^{-q}(a_n))$$

for any $q > p$ and n sufficiently large. On the other hand

$$(6.2) \quad s_n^2 \leq nH(a_nv^{-p}(a_n)).$$

We proceed to show that with the particular choice

$$(6.3) \quad v(x) = \log \log x$$

the fluctuations of S_n^T are subject to the classical law of the iterated logarithm. This asserts that if

$$(6.4) \quad c_n = o(s_n(\log \log s_n)^{-1/2})$$

then with probability one

$$(6.5) \quad \limsup \frac{S_n^T}{s_n(\log \log s_n)^{1/2}} = \sqrt{2}.$$

It is not difficult to see that the condition (6.4) is implied by (6.1) if H is of slow variation (that is, if F belongs to the domain of attraction of the normal distribution). Under our present more general conditions (6.4) need not be true. The relation (6.5) is therefore not assured, and even when it is true it is desirable to replace the clumsy norming constants s_n by simpler ones. This is achieved by

Lemma 6.1. *Assume that the integral (1.6) converges and the conditions (2.1)–(2.2) hold. Then with probability one*

$$(6.6) \quad \limsup S_n^T/a_n = \sqrt{2}.$$

Proof. Assume for the moment that we could sharpen (6.1)–(6.2) to

$$(6.7) \quad s_n^2 \sim nH(a_n).$$

The right side is the same as $a_n^2 v^{-1}(a_n)$ or $c_n^2 v^{2p-1}(a_n)$, and so (6.4) would be equivalent to (6.6). Unfortunately (6.7) need not be true, but we shall show that this relation becomes true if $n \rightarrow \infty$ omitting a set of integers that is so sparse as not to affect the conclusion (6.5). For this purpose we use again the sequence $\{n_k\}$ defined by (5.2) and put again $a_{n_k} = \alpha_k$. From Lemma 5.2 we see that

$$(6.8) \quad \sum \frac{1}{\log^{1+\sigma} k} \cdot \frac{H(\alpha_k) - H(\alpha_k v^{-q}(\alpha_k))}{H(\alpha_k)} < \infty.$$

For arbitrary, but fixed, $\delta > 0$ denote by K the set of all integers k such that

$$(6.9) \quad H(\alpha_k v^{-q}(\alpha_k)) < (1 - \delta)H(\alpha_k).$$

It is clear from (6.8) that

$$(6.10) \quad \sum_{k \in K} \log^{-1-\sigma} k < \infty$$

where $\sigma > 0$ is arbitrary. Now the central limit theorem applies to the truncated

variables X_n^T , and (6.10) implies trivially that

$$(6.11) \quad \sum_{k \in K} P\{S_{n_k}^T > \epsilon \alpha_k\} < \infty$$

for every $\epsilon > 0$. Hence

$$(6.12) \quad \limsup_{k \in K'} S_{n_k}^T / \alpha_k = 0.$$

The standard proof of the iterated logarithm applies when k is restricted to the complementary set K' , and so

$$(6.13) \quad \limsup_{k \in K} \frac{S_{n_k}^T}{s_{n_k}(\log \log s_{n_k})^{1/2}} = \sqrt{2}.$$

But on K' we have

$$(6.14) \quad s_{n_k}^2 > (1 - \delta)^2 n_k H(\alpha_k) = (1 - \delta)^2 \alpha_k \log \log \alpha_k$$

and the inequality is reversed when the factor $(1 - \delta)^2$ is omitted. Since ϵ is arbitrary we conclude that

$$(6.15) \quad \limsup S_{n_k}^T / \alpha_k = \sqrt{2}.$$

Thus (6.6) holds for arbitrary geometric subsequences, and from this (6.6) follows by the standard argument (used in section 5.).

7. Summary. In the preceding section a_n was defined as the largest root of

$$(7.1) \quad a_n^2 = nH(a_n) \log \log a_n.$$

With this choice the condition (2.1) is satisfied, and the two conditions (2.2) and (2.3) are equivalent.

In the final formulation it is preferable to use the canonical norming constants ρ_n defined in (1.3). If the tail $1 - F$ varies slowly then $a_n^2 \sim \rho_n^2 \log \log \rho_n$, but this relation may be false even for distributions belonging to the domain of attraction of the normal distribution. It is therefore remarkable that we have nevertheless

Theorem 1. *Assume (2.1) and (2.2). If the integral (1.6) converges then with probability one*

$$(7.2) \quad \limsup \frac{S_n}{a_n} = \limsup \frac{S_n}{\rho_n(\log \log \rho_n)^{1/2}} = \sqrt{2}.$$

In case of divergence with probability one

$$(7.3) \quad \limsup \frac{S_n}{a_n} = \limsup \frac{S_n}{\rho_n(\log \log \rho_n)^{1/2}} = \infty.$$

Proof. We apply the preceding results with $v(x) = \log \log x$. This function being monotone the assumption (2.3) is not necessary for lemmas 4.1 and 5.1.

Together with Lemma 3.3 these show immediately that in the case of convergence

$$(7.4) \quad S_n = S_n^T + o(a_n).$$

The first relation in (7.2) now follows from (6.6).

Comparing the definitions (7.1) and (1.3) one sees immediately that $a_n \geq \rho_n$, and hence

$$(7.5) \quad a_n^2 \geq \rho_n^2 \log \log \rho_n.$$

As already remarked, the two sides are not necessarily asymptotically equivalent, but we have seen that the integers n for which

$$(7.6) \quad H(a_n/\log \log a_n) < (1 - \delta)H(a_n)$$

are negligible in the sense that if $n \rightarrow \infty$ through values for which (7.6) is true then $S_n^T = o(a_n)$, while on the complementary set the limit relations (6.5) and (6.6) apply. But when (7.6) is false the quantity $x_n^2 = (1 - \delta)a_n^2/\log \log a_n$ satisfies $x_n^2 < nH(x_n)$ and is therefore $\leq \rho_n^2$. Thus $a_n^2 \sim \rho_n^2 \log \log \rho_n$ except on a negligible set, and so (7.2) is true in case of convergence.

In case of divergence the lemmas of section 3 imply that for arbitrary $\lambda > 0$ with probability one the relation $S_n > \lambda a_n$ takes place for infinitely many n . In view of (7.5) this implies (7.3).

The following theorem illustrates the role of the maximal term in the case of divergence. For the notation see (1.8).

Theorem 2. *Suppose that (2.3) holds and*

$$(7.7) \quad \int_0^\infty \frac{x^2 F(dx)}{H(x) \log \log^{1+\gamma} x} < \infty$$

for some $\gamma < \frac{1}{4}$, but that the integral diverges when $\gamma = 0$. Then (7.3) holds but

$$(7.8) \quad \limsup \frac{{}^{(1)}S_n}{\rho_n(\log \log \rho_n)^{1/2}} = \sqrt{2}.$$

Proof. As remarked at the end of section 4, in view of the monotone character of $\log \log$ Lemma 4.1 applies with $\eta_0 = \frac{1}{4}$, and Lemma 5.1 with $\delta < \eta/2$. The proof of the preceding theorem now applies with the sole change that the iterated logarithm in (7.1) is raised to the power $1 + \gamma$.

We saw that a limit relation of the form (7.2) can hold simultaneously for distinct sequences $\{a_n\}$, but it is easily seen that two such sequences can be essentially different only on subsequences that are rather sparse. Example (b) of the concluding section shows that even in the case of a divergent integral (1.6) a relation analogous to (7.2) may hold for a sequence of norming constants obtained by redefining $\{a_n\}$ on a certain sparse set. In order to avoid such freaks and trivialities we now limit our attention to sequences $\{a_n\}$ such that

$$(7.9) \quad \frac{a_n}{\rho_n(\log \log \rho_n)^{1/2}} \rightarrow \infty.$$

Although this is not really essential we simplify the proof by assuming that (2.3) holds for some value $M < 4$; in this case

$$(7.10) \quad \frac{H(tx)}{H(x)} < t^\gamma, \quad \gamma < 2$$

for all t and x sufficiently large.

Theorem 3. Assume the conditions (2.1)–(2.2) and (7.10). For a sequence satisfying (7.9) we have $\limsup S_n a_n^{-1} = 0$ or ∞ according as $\sum P\{X_n > a_n\}$ converges or diverges.

Proof. With v defined by (4.1) we have

$$(7.11) \quad v(a_n) = \frac{a_n^2}{nH(a_n)} = \frac{a_n^2}{\rho_n^2} \cdot \frac{H(\rho_n)}{H(a_n)} \geq \left(\frac{a_n}{\rho_n}\right)^{2-\gamma} \geq (\log \log \rho_n)^{1-\gamma/2}$$

for n sufficiently large. The proof of Theorem 1 now applies in a simplified form, since we have only to prove that $S_n^T = o(a_n)$.

Theorem 4. Assume (2.3). Whenever the integral (1.6) diverges there exists a sequence a'_n such that

$$(7.12) \quad \limsup S_n/a'_n = \infty \quad \text{but} \quad {}^{(1)}S_n = o(a'_n).$$

Proof. It is possible to find a monotone function w such that the integral

$$(7.13) \quad \int_0^\infty \frac{H(dx)}{H(x)w^{1+\gamma}(x)}$$

diverges for $\gamma = 0$ but converges for all $\gamma > 0$. Since (1.6) diverges we may also suppose $w(x) > \log \log x$. If we then define a_n and a'_n respectively as the largest roots of

$$a_n^2 = nH(a_n)w^{1+\gamma}(a_n), \quad a_n'^2 = nH(a'_n)w(a'_n)$$

the first relation in (7.11) follows from Lemma 3.3 and the divergence of (7.12) when $\gamma = 0$. The second relation is true because in the notations of sections 4–6 each of the partial sums S_n^T , ${}^{(1)}S_n'$, and S_n'' is $o(a'_n)$.

8. Examples. (a) Let the distribution F be concentrated on $|x| > 1$ with density $|x|^{-3}$. Then $H(x) = \log x$ for $x > 1$, and $2\rho_n^2 \sim n \log n$. The integral (7.7) converges only for $\gamma > 0$, and by Theorem 2 therefore

$$\limsup \frac{S_n}{(n \log n \log \log n)^{1/2}} = \infty, \quad \limsup \frac{{}^{(1)}S_n}{(n \log n \log \log n)^{1/2}} = 1.$$

In this case the analogue to the classical law of the iterated logarithm has been secured by the expedient of removing the offending maximal term. The same result may be achieved by a different method. With r standing for an arbitrary

integer define

$$a_n^2 = \begin{cases} n \log n \log \log n & \text{if } r^2 < \log \log n < r^2 + 1 \\ n \log n (\log \log n)^2 & \text{otherwise.} \end{cases}$$

It is easily verified that the series (3.5) converges, and the argument used in Theorem 1 shows directly that

$$(8.1) \quad \limsup S_n/a_n = \limsup S_n^r/a_n = 1.$$

To assure that the first part is true we had to modify the sequence $\{a_n\}$, but the second part is true only because it holds on the subsequence on which our new a_n agree with those used in Theorem 1.

(b) As a last example let F be concentrated on the points $\pm x_k$ of a very rapidly increasing sequence, say $x_k = (2k!)^k$. Let $P\{X = x_k\} \sim 1/(k+1)!$. Put

$$a_n = x_k \text{ when } (k-1)! < n \leq k!.$$

In order that $S_n > (1 + \epsilon)a_n$ it is necessary that at least two components X_i exceed x_k . The probability for this is the order of magnitude k^{-2} , and so with probability one this event occurs only finitely often. But with probability one it will happen infinitely often that one component equals x_k , and one concludes easily that $\limsup S_n/a_n = 1$. Such a relation is therefore possible with norming constants in no way related to those of Theorem 1. This however is due only to the erratic behavior of F .

REFERENCES

- [1] W. FELLER, A limit theorem for random variables with infinite moments, *American J. Math.*, **68** (1946) 257-262.
- [2] P. LÉVY, Loi faible et loi forte des grands nombres, *C. R. Acad. Sci. Paris*, **235** (1952) 1186-1188.
- [3] ———, Loi faible et loi forte des grands nombres, *Bull. Sci. Math.*, **77** (1953) 9-40.
- [4] V. STRASSEN, A converse to the law of the iterated logarithm, *Z. Wahrscheinlichkeits theorie*, **4** (1966) 265-268.

Princeton University
and
Rockefeller University
Date communicated: JANUARY 20, 1968.