THE STRONG LAW OF LARGE NUMBERS

KAI LAI CHUNG
CORNELL UNIVERSITY

1. Introduction

A well known unsolved problem in the theory of probability is to find a set of necessary and sufficient conditions (nasc's) for the validity of the strong law of large numbers (SLLN) for a sequence of independent random variables. This problem will not be solved in the present paper. To avoid a possible misunderstanding it must be stated at once that nasc's have been found, and several sets of them will be given in section 3, but they are all unsatisfactory. Presumably all (or shall we say most) mathematicians will agree on a satisfactory set of such conditions if and when they are exhibited, but before they are it does not seem easy to lay down criteria of satisfactoriness. On the other hand it is safe to rule out certain conditions as unsatisfactory, for example those in which sums of random variables enter; the conditions to be given in section 3 all have this undesirable property.

The purpose of this paper is to give an account of the latest information on this problem, at least in some directions. While undoubtedly much that follows is known to experts in the field or, so to speak, lurks in the corners of their minds, it is hoped that some of the results below are printed here for the first time and not sufficiently known to a wider circle of probabilists. It is to acquaint this latter group with the present status of knowledge of the problem that this paper is written.

The paper is divided into three sections. Section 2 is quite independent of the others and deals with the case of identically distributed, independent random variables (r.v.'s). In this case it is known, after Kolmogorov,[1] that a nasc for the validity of the SLLN is the finiteness of the first absolute moment of the common distribution function (d.f.). For use in certain statistical applications Professor Wald raised the question of the uniformity of the strong convergence with respect to a family of d.f.'s (see section 2). A nasc for this is given in section 2, which includes Kolmogorov's theorem as a special case. The method of proof is classical.

In section 3 several sets of nas, but unsatisfactory, conditions for the validity of the SLLN are given and their interrelations, mostly trivial, are explored. The results of this section includes Kawata's partial result[2] in this direction, and Pro-

This research was prepared at the University of California under the partial support of the Office of Naval Research.

[1] See Kolmogorov [1, p. 67]. As far as the author is aware the proof was never published by him. The proof of the sufficiency part is given in Fréchet [2]. The necessity part has been given without centering at the medians; see Feller [3], for more general results.

[2] Kawata [4]. He stated the theorem with zero expectations, an assumption which he never used.

horov's result recently announced [5].* The proof given here of Prohorov's result is different from and somewhat longer than his,[3] but it is hoped that it brings out the connections more clearly. As an application a simple proof of a sufficient condition which includes Kolmogorov's [7] and Brunk's [8] is given, as also announced by Prohorov.

In section 4 satisfactory nasc's for the SLLN are found for r.v.'s which are individually bounded and whose bounds satisfy certain restrictive order conditions. Such a result was also announced by Prohorov. By using a deep estimate due to Cramér and Feller [9], Prohorov's result is extended to slightly more general cases

In the following $\{X_n\}$, $n = 1, 2, \ldots$ will always denote a sequence of independent real valued r.v.'s, and $S_n = \sum_{k=1}^{n} X_k$. If X is a r.v., $m(X)$ denotes a median[4] of X; X^0 denotes the centered r.v. $X - m(X)$; $E(X)$ the expectation of X. If A is an event, $P(A)$ denotes its probability. The letters "i.o." are an abbreviation of the phrase "infinitely often," namely, "for an infinite number of values of whatever subscript is in question." The symbol ϵ denotes an arbitrarily small positive number, thus a proposition involving ϵ should read: "For every $\epsilon > 0$, etc."

If there exists a sequence of real numbers $\{c_n\}$ such that

$$P\left(\lim_{n\to\infty} \frac{S_n - c_n}{n} = 0\right) = 1 \tag{1.1}$$

we say that the sequence $\{X_n\}$ obeys the SLLN. In this case it is trivial that we can replace c_n by $m(S_n)$. Thus (1.1) is equivalent to

$$P\left(\lim_{n\to\infty} \frac{S_n^0}{n} = 0\right) = 1 \tag{1.1 bis}$$

or to the following:

$$P\left(|S_n^0| > n\epsilon \text{ i.o.}\right) = 0 \tag{1.2}$$

or to

$$P\left(|S_n^0| \leqq n\epsilon \quad \text{for all} \quad n \geqq N\right) = 1\,. \tag{1.3}$$

Note that in (1.3) the N is allowed to depend not only on ϵ, but also on the sample sequence $\{X_n\}$. Thus (1.3) is equivalent to the following: given any $\epsilon > 0$, there exists a fixed N_0 depending on ϵ but no longer on the sample sequence such that

$$P\left(|S_n^0| \leqq n\epsilon \quad \text{for all} \quad n \geqq N_0\right) \geqq 1 - \epsilon\,. \tag{1.4}$$

2. The identically distributed case

Let all X_n have the same d.f. $F(x)$. Kolmogorov (see footnote 2) proved that

* *Added in proof:* Prohorov's complete account has in the meanwhile appeared in *Izvestia Akad. Nauk. USSR*, Vol. 14 (1950), pp. 523–536.

[3] His proof depends on a new inequality of Kolmogorov, the idea of which is very close to one of P. Lévy [6, p. 138].

[4] Throughout this paper we could use instead of the median, any number $\mu(X)$ such that $P[X \geqq \mu(X)] \geqq \lambda$, $P[X \leqq \mu(X)] \geqq \lambda$ for some fixed $\lambda: 0 < \lambda < 1$.

a nasc that $\{X_n\}$ obeys the SLLN is that

$$\int_{-\infty}^{\infty} |x| \, dF(x) < \infty .$$

The proof of the necessity part is trivial; as to the proof of the sufficiency part there are three essentially different methods:

(i) Khintchine-Kolmogorov's method which depends on truncation and Kolmogorov's famous inequality with or without the intervention of infinite series;

(ii) A special case of G. D. Birkhoff's individual ergodic theorem, ever so many proofs of which have been given;[5]

(iii) Doob's [11] very elegant proof using the theory of martingales.

The proof of the following more general theorem uses method (i) and is in essence nothing but a precision of that method. It is not clear whether the other methods will be applicable.

Let a family of d.f.'s $F(x, \theta)$ be given where θ is the parameter of the family. All the r.v.'s X_n have one and the same d.f. $F(x, \theta)$ from the family where θ may be any value of the parameter. The sequence $\{X_n\}$ is said to obey the SLLN uniformly with respect to θ if: given any $\epsilon > 0$, there exists a fixed $N_0 = N_0(\epsilon)$ not depending on θ such that (1.4) holds no matter what θ is.

THEOREM. *A sufficient condition for the sequence $\{X_n\}$ to obey the SLLN uniformly with respect to θ is the following: given any δ there exists a number $A(\delta)$ not depending on θ such that*

$$\int_{|x| > A(\delta)} |x| \, dF(x, \theta) < \delta . \tag{2.1}$$

If so we can replace S_n^0 in (1.4) by $S_n - E(S_n)$. This condition is also necessary if the median $m(\theta)$ of $F(x, \theta)$ is a bounded function of θ.

PROOF. *Sufficiency.* From (2.1) it follows that

$$\int_{-\infty}^{\infty} |x| \, dF(x, \theta) \leqq A(1) + 1 = M .$$

Now choose $N \geqq 2$ and such that

$$\int_{|x| \geqq N} |x| \, dF(x, \theta) + \frac{16}{\epsilon^2} \left(\frac{2}{N^{1/2}} + 6 \int_{|x| \geqq N^{1/4}} |x| \, dF(x, \theta) \right) < \frac{\epsilon}{4} . \tag{2.2}$$

Having chosen N, choose $N_0 > N$ such that

$$\frac{16 N^2}{N_0 \epsilon} \int_{|x| > N_0 \epsilon / 3N} |x| \, dF(x, \theta) + \frac{NM}{N_0} < \frac{\epsilon}{4} . \tag{2.3}$$

We have

$$\sum_{k=N}^{\infty} P(|X_k| \geqq k) \leqq \int_{|x| \geqq N} |x| \, dF(x, \theta) . \tag{2.4}$$

[5] Extensions of methods (i) and (ii) to the case of dependent r.v.'s which includes the case of independence have been announced by M. Loève [10].

Next,

$$(2.5)\quad \sum_{k=N}^{\infty}\frac{1}{k^2}\int_{|x|<k}|x|^2 dF(x,\theta)=\sum_{k=0}^{\infty}\int_{k\le|x|<k+1}|x|^2 dF(x,\theta)\sum_{j=\max(N,k)}^{\infty}\frac{1}{j^2}$$

$$\le\frac{1}{N-1}\int_{0\le|x|<N^{1/4}}|x|^2 dF(x,\theta)+\frac{N+1}{N-1}\int_{|x|\ge N^{1/4}}|x|\,dF(x,\theta)$$

$$+\frac{N+2}{N}\int_{|x|\ge N+1}|x|\,dF(x,\theta)\le\frac{2}{N^{1/2}}+6\int_{|x|\ge N^{1/4}}|x|\,dF(x,\theta).$$

Define

$$X'_k=\begin{cases}X_k & \text{if} \quad |X_k|<k,\\ 0 & \text{if} \quad |X_k|\ge k\end{cases}$$

$$X''_k=X'_k-E(X'_k).$$

Then

$$E(X''^2_k)\le\int_{|x|<k}|x|^2 dF(x,\theta).$$

By Kolmogorov's inequality, and (2.3)-(2.5),

$$P\left(\left|\sum_{k=N}^{n}\frac{X_k-E(X'_k)}{k}\right|>\frac{\epsilon}{4}\text{ for at least one } n\ge N\right)<\frac{}{4}.$$

It follows from Kronecker's lemma that

$$P\left(\left|\frac{1}{n}\sum_{k=N}^{n}[X_k-E(X'_k)]\right|>\frac{\epsilon}{4}\text{ for at least one } n\ge N\right)<\frac{\epsilon}{4}.$$

Moreover, we have

$$P\left(\left|\frac{X_1+\ldots+X_N}{n}\right|>\frac{\epsilon}{4}\right)\le N\int_{|x|>n\epsilon/4N}dF(x,\theta)\le\frac{16N^2}{n\epsilon}$$

$$\times\int_{|x|>n\epsilon/3N}|x|\,dF(x,\theta)<\frac{\epsilon}{4},$$

and if $n\ge N_0$, by (2.2) and (2.3)

$$\left|\frac{1}{n}\sum_{k=N}^{n}E(X'_k)\right|\le\frac{NM}{n}+\int_{|x|\ge N}|x|\,dF(x,\theta)<\frac{\epsilon}{2}.$$

Altogether we conclude that

$$P\left(\left|\frac{S_n-E(S_n)}{n}\right|>\epsilon\text{ for at least one } n\ge N_0\right)<\frac{\epsilon}{2},$$

or

$$P[\,|S_n-E(S_n)|\le n\epsilon\text{ for all } n\ge N_0]\ge 1-\epsilon.$$

If $\epsilon<\frac{1}{2}$ it follows by the definition of a median that for all $n\ge N_0$,

$$(2.6)\qquad |m(S_n)-E(S_n)|\le n\epsilon.$$

Thus

$$P\,[\,|S_n - m\,(S_n)\,| \leqq 2n\epsilon \quad \text{for all} \quad n \geqq N_0] \geqq 1 - \epsilon\,.$$

This implies (1.4), whatever θ is.

Necessity. Suppose that (1.4) holds where N_0 does not depend on θ. Then if $n \geqq N_0$,

$$P\,[\,|X_n - m\,(S_n) + m\,(S_{n-1})\,| \leqq 2n\epsilon] \geqq 1 - \epsilon\,.$$

If $\epsilon < \frac{1}{2}$ this entails

$$|m\,(X_n) + m\,(S_{n-1}) - m\,(S_n)\,| \leqq 2n\epsilon\,.$$

Since by hypothesis $|m\,(X_n)\,| = |\,m\,(\theta)\,| \leqq m$ where m does not depend on θ, there exists a number N_1, not depending on θ, such that if $n \geqq N_1$

$$|m\,(S_{n-1}) - m\,(S_n)\,| < 3n\epsilon\,. \tag{2.7}$$

Now suppose that (2.1) did not hold and we wish to reach a contradiction. If (2.1) did not hold there exists a $\delta > 0$ such that for any N there is a θ_N for which

$$\int_{|x|>N} |x|\,dF(x,\ \theta_N) \geqq \delta\,.$$

Hence

$$\sum_{k=N+1}^{\infty} \int_{|x|>k} dF(x,\ \theta_N) + (N+1) \int_{|x|>N} dF(x,\ \theta) \geqq \delta\,.$$

It follows that one of the following two cases would occur:

Case (i). For a sequence $N_i \uparrow \infty$, there corresponds a sequence θ_i such that if all X_n have the d.f. $F(x, \theta_i)$, then

$$\sum_{n=N_i+1}^{\infty} P(\,|X_n| \geqq n) \geqq \frac{\delta}{2}.$$

Hence for this sequence $\{X_n\}$ we have

$$\sum_{n=N_i+1}^{\infty} P(\,|X_n| < n) \leqq e^{-\delta/2}\,.$$

$$P\,(|X_n| \geqq n \quad \text{for at least one} \quad n \geqq N_i + 1) \geqq 1 - e^{-\delta/2}\,. \tag{2.8}$$

We have $X_n = S_n - m(S_n) - [S_{n-1} - m\,(S_{n-1})] + m(S_n) - m(S_{n-1})$; by (2.7), since $N_i \geqq N_1$, if $\epsilon < 1/6$, (2.8) entails

$$P\,[\,|S_n - m\,(S_n)\,| \geqq \frac{n}{4} \quad \text{for at least one} \quad n \geqq N_i] \geqq 1 - e^{-\delta/2}\,.$$

Since $N_i \uparrow \infty$, (1.4) becomes false for $\epsilon < \min\,(1/4,\ 1 - e^{-\delta/2})$.

Case (ii). For a sequence $N_i' \uparrow \infty$, there corresponds a sequence θ_i' such that if all X_n have the d.f. $F(x, \theta_i')$, then

$$\sum_{n=N_i'}^{2N_i'} P(\,|X_n| > N_i') \geqq \frac{\delta}{2}$$

whence

$$P(|X_n| > N_i' \text{ for at least one } n\colon N_i' \leqq n \leqq 2N_i') \geqq 1 - e^{-\delta/2}.$$

The same argument as in case (i) finishes the proof.

Remark 1. If the family of d.f.'s consists of a single d.f., then the theorem reduces to Kolmogorov's.

Remark 2. Without the assumption of the boundedness of the medians, the condition stated in the theorem is not necessary.

Example. Let θ run over the positive integers and define $F(x, n)$ to be the d.f. which has a single jump at the point $x = n$.

Remark 3. The following simpler version may be more useful for applications; its proof is similar but simpler. Suppose that for every θ,

$$\int_{-\infty}^{\infty} x\,dF(x, \theta) = 0, \qquad \int_{-\infty}^{\infty} |x|\,dF(x, \theta) < \infty.$$

Then the condition stated in the theorem is a nasc that: given any $\epsilon > 0$, there exists a N_0 depending on ϵ but not on θ nor on the sample sequence, such that

$$P(|S_n| \leqq n\epsilon \text{ for all } n \geqq N_0) \geqq 1 - \epsilon.$$

3. Necessary and sufficient conditions and a sufficient condition

We return now to the general case. We shall consider, besides the SLLN embodied in formula (1.1), also a modified form, namely,

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = 0\right) = 1. \tag{3.1}$$

For any given sequence c_n, (1.1) can be reduced to (3.1) by an obvious change of variables: $X_n^* = X_n - c_n + c_{n-1}$. Thus while (1.1) answers the question: does there exist some sequence $\{c_n\}$ such that (1.1) holds; (3.1) answers the question: does (1.1) hold with a given sequence c_n. The second question will of course be answered via the first if we can decide whether $\lim_{n\to\infty} \frac{m(S_n) - c_n}{n} = 0$ or not, but there seems in general no control over $m(S_n)$.

To simplify writing, "convergence in probability" will be denoted by an arrow $\rightarrow$; "convergence with probability one" or "almost sure convergence" by a double arrow $\rightrightarrows$. If A and B are two propositions, $A \supset B$ means "A implies B"; $A \equiv B$ means "A and B are equivalent."

Consider the following:

(1) $\quad \dfrac{S_n}{n} \rightrightarrows 0;$

(2) $\quad \dfrac{S_n}{n} \rightarrow 0;$

(3) $\quad \dfrac{S_{2^n}}{2^n} \rightrightarrows 0;$ [6]

[6] Instead of 2^n we can take any sequence of positive integers such that $0 < A_1 < q_{n+1}/q_n < A_2 < \infty$.

(4) $$\frac{S_{2^{n+1}} - S_{2^n}}{2^n} \to 0 ;$$

(5) $$\sum_n P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) < \infty ;$$

(6) $$\sum_n P(|S_{2^n}| > 2^n\epsilon) < \infty .$$

The following relations obtain:

I. (1) $\supset$ (2): Well known.

II. (1) $\supset$ (3) $\equiv$ (4) $\equiv$ (5): The equivalence of (3) and (4) is a simple analytical fact; that of (4) and (5) is a consequence of the Borel-Cantelli lemma.

III. (5) $\subset$ (6) $\supset$ (3): That (6) implies (3) is a consequence of one half of the Borel-Cantelli lemma; that (6) implies (5) follows from Boole's inequality.

IV. (2) and (3) $\supset$ (1).

PROOF. (3) is equivalent to

$$P(|S_{2^n}| > 2^n\epsilon \text{ i.o.}) = 0 .$$

For every positive integer k define $n(k)$ by $2^{n(k)-1} \leqq k < 2^{n(k)}$. Since

$$P(|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon) \geqq 1 - P(|S_{2^{n(k)}}| > 2^{n(k)-1}\epsilon) - P(|S_k| > 2^{n(k)-1}\epsilon)$$

by (2) we have if $k > k_0 = 2^{n_0}$

$$P(|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon) > \tfrac{1}{2} .$$

Now $|S_k| > 2^{n(k)+1}\epsilon$ and $|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon$ together imply $|S_{2^{n(k)}}| > 2^{n(k)}\epsilon$ and the first two events are independent, hence by a simple argument,

$$P(|S_k| > 2^{n(k)+1}\epsilon \text{ for some } k > k_0) \leqq 2P(|S_{2^{n(k)}}| > 2^{n(k)}\epsilon \text{ for some } k > k_0) .$$

Letting $k_0 \to \infty$ we obtain

$$P(|S_k| > 4k\epsilon \text{ i.o.}) \leqq 2P(|S_{2^n}| > 2^n\epsilon \text{ i.o.}) = 0 .$$

V. (1) $\equiv$ (2) and (3) $\equiv$ (2) and (4) $\equiv$ (2) and (5) $\subset$ (2) and (6): from I–IV.

The implication (2) and (6) $\supset$ (1) was proved by Kawata [4]. Proposition (2) is one form of the weak law of large numbers (WLLN). Thus the relations in V show that the SLLN, in the form (2.1), is equivalent to the corresponding WLLN plus the SLLN for the subsequence S_{2^n}. Now satisfactory nasc for the WLLN have been given by Kolmogorov [12] and Feller [13]. Hence we can, if we prefer, replace (2) everywhere in V by these conditions. The significance of Prohorov's result below lies in the elimination of the WLLN as part of the sufficient conditions for the SLLN, and this is done by centering (at the medians).

Now we consider the relations (1)–(6) with S_n, S_{2^n}, $S_{2^{n+1}} - S_{2^n}$ replaced by S_n^0, $S_{2^n}^0$, $(S_{2^{n+1}} - S_{2^n})^0$ respectively. (Note that $S_{2^{n+1}} - S_{2^n}$ is *not* replaced by $S_{2^{n+1}}^0 - S_{2^n}^0$!). The resulting propositions we call (1^0)–(6^0). We add a new proposition

(7^0) $$\frac{S_{2^n}^0}{2^n} \to 0 .$$

We notice first that (7^0) entails

$$\frac{S_{2^{n+1}} - S_{2^n} - m\,(S_{2^{n+1}}) - m\,(S_{2^n})}{2^n} \to 0\,;$$

hence it follows that

$$(8^0) \qquad \frac{m\,(S_{2^{n+1}}) - m\,(S_{2^n}) - m\,(S_{2^{n+1}} - S_{2^n})}{2^n} \to 0\,.$$

All the relations I–III above carry over for the circled propositions and the corresponding relations will be referred to as I^0–III^0. We have only to consider $X_n - m(S_n) + m(S_{n-1})$ as new r.v.'s and apply I–III; the fact that (7^0) $\supset$ (8^0) has to be used in several places.

IV^0. (3^0) $\supset$ (7^0) $\supset$ (2^0).

PROOF. Only the second implication needs a proof and this is easiest done by resorting to characteristic functions (c.f.'s). Let X_n have the c.f. $f_n(t)$. (7^0) implies

$$\lim_{n\to\infty} \prod_{j=1}^{2^n} \left| f_j\left(\frac{t}{2^n}\right)\right| = 1$$

uniformly in any finite interval $|t| \leqq T$. This immediately implies

$$\lim_{k\to\infty} \prod_{j=1}^{k} \left| f_j\left(\frac{2^{n(k)}}{k}\cdot\frac{t}{2^{n(k)}}\right)\right| = 1$$

uniformly in $|t| \leqq T/2$, where $n(k)$ is defined in IV, whence (2^0).

V^0. (5^0) $\supset$ (1^0).

PROOF. (5^0) $\equiv$ (3^0) $\supset$ (7^0) $\supset$ (8^0). Since (8^0) holds (5^0) is equivalent to

$$(9^0) \qquad \sum P\,(\,|\,S^0_{2^{n+1}} - S^0_{2^n}\,| > 2^n\epsilon) < \infty\,.$$

Moreover by II^0 and IV^0, (5^0) $\supset$ (2^0). But (2^0) and (9^0) imply (1^0), by the third proposition in V.

VI^0. (1^0) $\equiv$ (3^0) $\equiv$ (4^0) $\equiv$ (5^0) $\subset$ (6^0): from II^0, III^0, IV^0, V^0.

The equivalence (1^0) $\equiv$ (5^0) is Prohorov's theorem [5].

We shall now prove the following theorem which gives a sufficient condition for the SLLN and includes Kolmogorov's ($r = 1$) and Brunk's (r integer $\geqq 1$).

THEOREM. *Let $E(X_n) = 0$ for every n, and $E(|X_n|^{2r}) < \infty$ for some real number $r \geqq 1$. If*

$$(3.2) \qquad \sum_n \frac{E\,(\,|\,X_n\,|^{2r})}{n^{r+1}} < \infty\,.$$

Then (2.1) *holds.*

PROOF. We need the following inequality

$$(3.3) \qquad E\left(\left|\sum_{k=1}^{n} X_k\right|^{2r}\right) \leqq A\,n^{r-1} \sum_{k=1}^{n} E\,(\,|\,X_k\,|^{2r})$$

where A depends only on r. This is easily proved if we use an inequality due to Marcinkiewicz and Zygmund [14] (trivial if r is an integer) according to which

$$E\left(\left|\sum_{k=1}^{n} X_k\right|^{2r}\right) \leqq A E\left[\left(\sum_{k=1}^{n} X_k^2\right)^r\right].$$

Now by Hölder's inequality

$$\Big(\sum_{k=1}^{n} X_k^2\Big)^r \leqq n^{r-1}\sum_{k=1}^{n} |X_k|^{2r};$$

hence (3.3) follows.

From (3.2) it follows firstly by Kronecker's lemma

$$\overline{\lim_{n\to\infty}} E\Big(\Big|\frac{S_n}{n}\Big|^{2r}\Big) \leqq \lim_{n\to\infty} \frac{A}{n^{r+1}}\sum_{k=1}^{n} E(|X_k|^{2r}) = 0 .$$

Hence by Tschebicheff's inequality, we have proposition (2) above. Next, again by Tschebicheff's inequality, and (3.3)

$$P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) \leqq \frac{A\, 2^{n(r-1)}}{(2^n\epsilon)^{2r}} \sum_{k=2^n+1}^{2^{n+1}} E(|X_k|^{2r}) \leqq \frac{A\, 2^{r+1}}{\epsilon^{2r}} \sum_{k=2^n+1}^{2^{n+1}} \frac{E(|X_k|^{2r})}{k^{r+1}} .$$

Hence from (3.2) follows proposition (5). Since (2) and (5) $\equiv$ (1) by VI we have (1).

Remark. Using truncated variables the theorem can be stated without assuming any moments. We shall not insist on this, and also other more or less trivial extensions of the theorem [15].

4. Necessary and sufficient conditions for some special cases

For easier reference we shall rewrite some of the previous formulas:

$$\frac{S_n^0}{n} \rightrightarrows 0 , \tag{4.1}$$

$$\frac{S_n}{n} \rightrightarrows 0 . \tag{4.2}$$

If (4.1) holds, it is easy to see that we have

$$\sum P(|X_n^0| \geqq n\epsilon) < \infty . \tag{4.3}$$

Define $X_n' = X_n^0$ if $|X_n^0| \leqq n\epsilon$, and $X_n' = 0$ if $|X_n^0| > n\epsilon$; then under (4.3) the sequences $\{X_n'\}$ and $\{X_n^0\}$ are equivalent in the sense of Khintchine. If the SLLN is valid for $\{X_n\}$ it is by definition also valid for $\{X_n^0\}$ and so for $\{X_n'\}$; conversely if the SLLN is valid for $\{X_n'\}$, and (4.3) is assumed, then it is also valid for $\{X_n\}$. Hence we may confine ourselves to r.v.'s $\{X_n\}$ satisfying the following condition:

$$\sup |X_n| = o(n) . \tag{4.4}$$

Under (4.4) we may, without loss of generality, assume that

$$E(X_n) = 0 , E(X_n^2) = \sigma_n^2 , \sum_{k=1}^{n} \sigma_k^2 = s_n^2 [= s^2(n)], s^2(2^{n+1}) - s^2(2^n) = d_n^2 . \tag{4.5}$$

Lemma. *Under* (4.4) *and* (4.5), (4.1) *and* (4.2) *are equivalent.*

Proof. We need only prove that (4.1) implies (4.2). We shall first prove that (4.1) implies $s_n = o(n)$. Otherwise let $n_k \uparrow \infty$ and $s_{n_k} \geqq \delta n_k$ for all k with $\delta > 0$.

Then by (4.4) we have $\max_{1\leqq j\leqq n_k} \sup |X_j| = o(s_{n_k})$. By a classical theorem of P. Lévy ([6], p. 102) the central limit theorem holds for the sequence S_{n_k}, in particular for every $\eta > 0$

$$\lim_{k\to\infty} P(S_{n_k} \geqq \eta\, s_{n_k}) = \lim_{k\to\infty} P(S_{n_k} \leqq -\eta\, s_{n_k}) = \frac{1}{\sqrt{2\pi}}\int_\eta^\infty e^{-y^2/2} dy < \tfrac{1}{2}\,.$$

It follows that $m(S_{n_k}) = o(s_{n_k})$. Consequently the d.f. of $s_{n_k}^{-1}(S_{n_k}^0)$ tends to that of the normal and (4.1) cannot be true.

Therefore $s_n = o(n)$ and it follows that $n^{-1}S_n \to 0$ (in probability !). This and (4.1) imply that $m(S_m) = o(n)$ and hence (4.2). q.e.d.

Of the results in section 3 we shall use the following which, combined with the lemma above, will be referred to as (P).

(P) Under (4.4) and (4.5): if $s_n = o(n)$ and also

$$\sum P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) < \infty \tag{4.6}$$

then both (4.1) and (4.2) hold; if (4.1) or (4.2) holds, then (4.6) holds.

In the following (4.4) and (4.5) will be assumed.

THEOREM 1. *If the further condition is satisfied:*

$$M_n = \max_{2^n<k\leqq 2^{n+1}} \sup |X_k| = o\left(\frac{d_n^2}{2^n}\right), \tag{4.7}$$

then a nasc for (4.1) *or* (4.2) *is*

$$\sum \exp(-\epsilon 2^{2n} d_n^{-2}) < \infty\,. \tag{4.8}$$

PROOF. *Sufficiency.* If (4.8) holds, then $s_n = o(n)$ because s_n is nondecreasing. Furthermore, by a theorem of Feller and using his notation, see [9],

$$\max_{2^n<k\leqq 2^{n+1}} \sup |X_k| \leqq \lambda_n d_n\,,\quad \lambda_n = 2^{-n} d_n = o(1)\,,\quad x = \epsilon 2^n d_n^{-1}\,,\quad 0 < \lambda_n x = \epsilon$$

we have

$$P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) \sim \frac{C d_n}{2^n \epsilon} \exp\{(-\epsilon^2 2^{2n} d_n^{-2})(1+\delta)\} \tag{4.9}$$

where $\delta \to 0$ as $\epsilon \to 0$, and C is an absolute constant. Thus (4.8) implies (4.6). (4.2) and (4.1) follow by (P).

Necessity. If (4.1) or (4.2) holds, (4.6) holds by (P). Also from the proof of the lemma, we have $s_n = o(n)$. Hence the estimate (4.9) is applicable and (4.6) reduces to (4.8).

Remark. Obviously (4.8) is a nasc for (4.1) or (4.2) if all the X_n have a normal distribution with zero mean.

THEOREM 2. *If instead of* (4.7) *we assume*

$$\sup |X_n| = o\left(\frac{n}{\lg \lg n}\right) \tag{4.10}$$

then (4.8) *is again a nasc for* (4.1) *or* (4.2).

PROOF. It follows from (4.10) that $M_n = o(2^n \lg^{-1} n)$. Let the values of n for which $M_n > 2^{-n}d_n^2$ be n_k, $k = 1, 2, \ldots$. Let $M_n^* = \epsilon^{-1}M_n(> M_n$ if $\epsilon < 1)$. Then

$$\frac{d_{n_k}^2}{\epsilon 2^{n_k}} < M_{n_k}^* = o\left(\frac{2^{n_k}}{\lg n_k}\right).$$

By an inequality of Kolmogorov [16], we have if $n = n_k$

$$(4.11) \qquad P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) \leqq \max\left[\exp\left(\frac{-2^n\epsilon}{4M_n^*}\right), \exp\left(-\frac{2^{2n}\epsilon^2}{4d_n^2}\right)\right] = \exp\left(-\frac{2^n\epsilon}{4M_n^*}\right).$$

It follows from (4.11) that

$$(4.12) \qquad \sum_k P(|S_{2^{n_k+1}} - S_{2^{n_k}}| > 2^{n_k}\epsilon) < \infty .$$

On the other hand it is trivial that

$$(4.13) \qquad \sum_k \exp[-\epsilon 2^{2n_k} d^{-2}(n_k)] < \infty .$$

Now, if (4.1) or (4.2) holds, then (4.6) holds by (P). If $n \neq n_k$, (4.9) is applicable, hence

$$\sum_{n \neq n_k} \exp[-\epsilon 2^{2n} d^{-2}(n)] < \infty .$$

This and (4.13) give (4.8).

Conversely, if (4.8) holds, then by (4.9)

$$\sum_{n \neq n_k} P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) < \infty .$$

This and (4.12) give (4.6). (4.1) and (4.2) follow by (P).

Theorem 2 was announced by Prohorov. If s_n is of a greater order of magnitude than $n(\lg\lg n)^{-1/2}$, theorem 1 provides an extension. Although these two theorems are better than the crude results which can be obtained directly from the law of the iterated logarithm, the domain of their applicability is essentially the same as that of the latter, since we use the estimate (4.9) which leads to it.

The following examples, due to Dr. Erdös, show that in general (4.8) is neither necessary nor sufficient for (4.1) or (4.2) even under (4.4) and (4.5).

Example 1. $X_k \equiv 0$ if $k \neq 2^n$; $X_{2^n} = \pm 2^n(\lg\lg n)^{-1}$ each with probability $\frac{1}{2}$.

Example 2. $X_n = 0$ with probability $1 - 2n^{-1}(\lg\lg n)^{-2}$; $X_n = \pm n(\lg\lg\lg n)^{-1}$ each with probability $n^{-1}(\lg\lg n)^{-2}$.

5. Concluding remarks

An opinion may be ventured in conclusion. It is quite possible that the strong law of large numbers will be solved by an approach entirely different from that sketched here. It is even possible that it will be solved by a stroke of great cunning, circumventing all the difficulties inherent in the present methods. Or it may be solved as a result of obtaining sharp asymptotic estimates for probabilities of the form $P(|S_n| > n\epsilon)$ in the general case.

It would appear from the necessary and sufficient conditions given in section 3 that such estimates may be indispensable, but this is not true in the special case of identically distributed random variables (see section 2). However, one thing should be said: the appraisal of such probabilities *is* one of the fundamental problems of the theory of probability, and any real progress in this direction will be of more importance than the solution of a specific problem.

"One hates that power does not come from oneself,
but one does not care if the task is done by oneself."

—Confucius

REFERENCES

[1] A. N. KOLMOGOROV, "Grundbegriffe der Wahrscheinlichkeitsrechnung," *Ergebnisse der Mathematik*, Vol. 2 (1937); English translation, Chelsea, 1950.
[2] M. FRÉCHET, *Recherches Théoriques Modernes sur la Théorie des Probabilités*, Book 1, Gauthier-Villars, Paris, 1937.
[3] W. FELLER, "A limit theorem for random variables with infinite moments," *Amer. Jour. Math.*, Vol. 58 (1946), pp. 257–262.
[4] T. KAWATA, "On the strong law of large numbers," *Proc. Imperial Acad. Tokyo*, Vol. 16 (1940), pp. 109–112.
[5] U. V. PROHOROV, "On the strong law of large numbers," (in Russian), *Doklady Akad. Nauk.*, Vol. 69 (1949), pp. 607–610.
[6] P. LÉVY, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937.
[7] A. N. KOLMOGOROV, "Sur la loi forte de grandes nombres," *C. R. Acad. Sci., Paris*, Vol. 191 (1930), pp. 910–911.
[8] H. D. BRUNK, "The strong law of large numbers," *Duke Math. Jour.*, Vol. 15 (1948), pp. 181–195.
[9] W. FELLER, "Generalization of a probability limit theorem of Cramér," *Trans. Amer. Math. Soc.*, Vol. 54 (1943), pp. 361–372.
[10] M. LOÈVE, "Remarques sur la convergence presque sure," *C. R. Acad. Sci., Paris*, Vol. 230 (1950), pp. 52–53; see also his Symposium paper in the present volume.
[11] J. L. DOOB, "Application of the theory of martingales," *Le Calcul des Probabilités et Ses Applications, Colloques Internationaux du Centre National de la Recherche Scientifique*, No. 13 (1949), pp. 23–27.
[12] A. N. KOLMOGOROV, "Ueber die Summen durch den Zufall bestimmter unabhängiger Grössen," *Math. Annalen*, Vol. 99 (1928), pp. 300–319; Vol. 102 (1930), pp. 484–488.
[13] W. FELLER, "Ueber das Gesetz der grossen Zahlen," *Acta Litt. ac Sci., Sect. Sci. Math., Szeged*, Vol. 8 (1936–37), pp. 191–201.
[14] J. MARCINKIEWICZ and A. ZYGMUND, "Sur les fonctions indépendantes," *Fund. Math.*, Vol. 29 (1937), pp. 60–90.
[15] K. L. CHUNG, "Note on some strong laws of large numbers," *Amer. Jour. Math.*, Vol. 69 (1947), pp. 189–192.
[16] A. N. KOLMOGOROV, "Ueber das Gesetz des iterierten Logarithms," *Math. Annalen*, Vol. 101 (1929), pp. 126–135.