Z. Wahrscheinlichkeitstheorie verw. Gebiete
61, 453–458 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On Strassen's Law of the Iterated Logarithm

P. Szüsz[1] and B. Volkmann[2]

[1] Dept. of Mathematics, State University of New York, Stony Brook, NY 11794, USA
[2] Mathematisches Institut A, Universität, Pfaffenwaldring 57, D 7000 Stuttgart 80
Federal Republic of Germany

Summary. Kolmogorov's law of the iterated logarithm has been sharpened by Strassen who proved a more refined theorem by using tools from functional analysis. The present paper gives a "classical" proof of Strassen's theorem, using a method along the lines of Kolmogorov's original approach. At the same time the result proved here is more general since a) the random variables involved need not have the same distributions, b) the condition of independence is weakened and c) instead of Kolmogorov's growth condition on the random variables, only a mild restriction on their moments of order $l \geqq 3$ is needed.

1. Introduction

Kolmogorov's classical law of the iterated logarithm [1] states that for any sequence $\xi_1, \xi_2, \ldots$ of mutually independent random variables with expectation 0 and variance $1+o(1)$, the sequence

$$T_n = \frac{S_n}{\sqrt{2n \log\log n}} \quad \text{with} \quad S_n = \sum_{i=1}^{n} \xi_i$$

satisfies, under certain conditions, the equations

$$\overline{\lim_{n\to\infty}}\, T_n = 1 \quad \text{and} \quad \underline{\lim_{n\to\infty}}\, T_n = -1$$

with probability 1. This theorem has been sharpened by Strassen [3] who investigated the behaviour of the functions $\varphi_n(t)$ obtained by linear interpolation of the values

$$\varphi_n\left(\frac{i}{n}\right) = T_n(i), \quad \text{where} \quad T_n(i) = \frac{S_i}{\sqrt{2n \log\log n}} \tag{1.1}$$

(n fixed, $i = 1, \ldots, n$; $\varphi_n(0) = 0$).

0044-3719/82/0061/0453/$01.20

Strassen showed that under certain restrictions imposed upon the r.v.'s ξ_i, the set of limit functions of the sequence $\varphi_3, \varphi_4, \varphi_5, \ldots$ under uniform convergence is the set $\mathfrak{S}$ of absolutely continuous functions χ on $[0,1]$ with $\chi(0)=0$ whose (almost everywhere existing) derivative χ satisfies

$$\int_0^1 \dot{\chi}(t)^2\, dt \leqq 1. \tag{1.2}$$

The present paper was motivated by the attempt to find a "classical" proof of Strassen's theorem (in the sense of avoiding the tools of functional analysis). It turned out indeed that the result can be obtained along the lines of Kolmogorov's original proof, using as additional idea a decomposition of the summation $S_n = \sum_{i=1}^{n} \xi_i$ into suitable blocks. Furthermore we could weaken Kolmogorov's condition

$$|\xi_n| = o\left(\sqrt{\frac{n}{\log\log n}}\right), \tag{1.3}$$

assuming only that for any given $\varepsilon > 0$, there exists an $n_0 = n_0(\varepsilon)$ such that for all $n \geqq n_0$,

$$\sum_{i=1}^{n} E(|\xi_i|^l) \leqq \varepsilon^l \frac{l!\, D(S_n)^l}{(\log\log D^2(S_n))^{\frac{l}{2}-1}} \quad (l=3,4,\ldots). \tag{1.4}$$

Assuming (1.3), the condition (1.4) follows by induction. Indeed, for $l=3$, letting $\varepsilon_1 = \varepsilon^3$ and using that $D(\xi_i) = 1 + o(1)$, we have for all sufficiently large n,

$$\sum_{i=1}^{n} E(|\xi_i|^3) \leqq \varepsilon_1 \sqrt{\frac{n}{\log\log n}} \sum_{i=1}^{n} D^2(\xi_i) \leqq \varepsilon^3 \frac{D(S_n)^3}{\sqrt{\log\log D^2(S_n)}}.$$

If this implication has been established up to some exponent $l \geqq 3$, it follows that

$$\begin{aligned}\sum_{i=1}^{n} E(|\xi_i|^{l+1}) &\leqq \varepsilon \sqrt{\frac{n}{\log\log n}} \sum_{i=1}^{n} E(|\xi_i|^l)\\ &\leqq \varepsilon \sqrt{\frac{n}{\log\log n}}\, \varepsilon^l \frac{D(S_n)^l}{(\log\log D^2(S_n))^{\frac{l}{2}-1}}\\ &= \varepsilon^{l+1} \frac{D(S_n)^{l+1}}{(\log\log D^2(S_n))^{\frac{l+1}{2}-1}},\end{aligned}$$

and thus (1.4) is also true for $l+1$.

Furthermore, our result is more general than Strassen's theorem also in the sense that we do not need to assume that the ξ_i's are identically distributed.

We shall first prove Strassen's theorem under the condition (1.4), but the final form of our result as stated below replaces (1.4) by an even weaker condition which is less restrictive than assuming the existence of all "fractional" moments $E(|\xi_i|^{2+\varepsilon})$, $i \geqq i_0(\varepsilon)$, $\varepsilon > 0$.

2. Lemmas

The following lemmas will be needed, maintaining the notation, the definitions and the assumptions stated in Sect. 1. Probability will always be denoted by P, expectation by E and standard deviation by D.

Lemma 1. *Let $\xi_1, \xi_2, \ldots$ be r.v.'s with $E(\xi_i)=0$ and $D^2(\xi_i)=1$. Suppose that all moments $E(|\xi_i|^l)$ exist and satisfy the inequalities* (1.4). *Then*

$$E(e^{aS_n}) < e^{\frac{a^2}{2} D^2(S_n)(1+\delta_n)}, \qquad (n=1,2,\ldots) \tag{2.1}$$

for all a with $0<a<C_2 \cdot \frac{\sqrt{\log\log D^2(S_n)}}{D(S_n)}$, where δ_n depends on the sequence $\varepsilon_1, \varepsilon_2, \ldots$ of (1.4) *only and $\delta_n \to 0$. Similarly, we have for the same range of a,*

$$E(e^{aS_n}) > e^{\frac{a^2}{2} D^2(S_n)(1-\delta_n)} \qquad (n=1,2,\ldots). \tag{2.2}$$

Proof. We shall only prove (2.1), the proof of (2.2) being analogous.

One has

$$E(e^{a\xi_i}) = \sum_{l=0}^{\infty} \frac{a^l}{l!} E(\xi_i^l) = 1 + \frac{a^2}{2} D^2(\xi_i) + \sum_{l=3}^{\infty} \frac{a^l}{l!} E(\xi_i^l).$$

Therefore, using a well-known inequality, we obtain

$$E(e^{aS_n}) = \prod_{i=1}^{n} E(e^{a\xi_i}) < \exp\left(\frac{a^2}{2} D^2(S_n) + \sum_{l=3}^{\infty} \frac{a^l}{l!} \sum_{i=1}^{n} E(|\xi_i|^l)\right), \tag{2.3}$$

from which the assertion follows by (1.4).

Lemma 2 (Bernstein-Kolmogorov Inequality). *For any $t < C_3 \log\log D^2(S_n)$, one has*

$$P(S_n > \sqrt{2t\, D^2(S_n)}) < e^{-t(1+o(1))}. \tag{2.4}$$

Proof. According to a well-known inequality (see, e.g. Renyi [2], p. 322 (1)) we have for any $a>0$,

$$P\left(S_n > \frac{1}{a}(t + \log E(e^{aS_n}))\right) < e^{-t}. \tag{2.5}$$

Now let

$$a = \sqrt{\frac{2t}{D^2(S_n)}}. \tag{2.6}$$

Because of $t < C_3 \log\log D^2(S_n)$ we can apply Lemma 1 with this number a, thus obtaining

$$\log E(e^{aS_n}) < \frac{a^2}{2} D^2(S_n)(1+o(1)),$$

which proves the assertion.

Lemma 3. *For any t satisfying $C_4 < t < C_3 \log\log D^2(S_n)$, one has*

$$P(S_n > \sqrt{2t\, D^2(S_n)}) > e^{-t(1+o(1))}. \tag{2.7}$$

Proof. See Kolmogorov [1].

Lemma 4. *For any $c>1$ and $\varepsilon>0$ there is an $n_0=n_0(c,\varepsilon)$ such that, for all y satisfying $C_4<|y|<\sqrt{C_5 \log\log D^2(S_n)}$, the inequality*

$$e^{-\frac{1}{2}y^2(1+\varepsilon)}<P(S_n\in(yD(S_n), cyD(S_n))) \qquad (n=1,2,\dots) \tag{2.8}$$

holds. In other words, the probability that $S_n\in(yD(S_n),\ cyD(S_n))$ "dominates" in the term $P(yD(S_n)<S_n)$.

Proof. The assertion follows by combining Lemmas 2 and 3; we only have to choose n so large that

$$c^2(1-\varepsilon)>1+2\varepsilon,$$

where ε is the larger of the two $o(1)$-terms in (2.4) and (2.7).

3. Main Result

Theorem. *Let $\xi_1,\xi_2,\dots$ be mutually independent r.v.'s with $E(\xi_n)=0$ and $D^2(\xi_n)\to 1$ as $n\to\infty$, satisfying the following condition for some sequence $\varepsilon_n>\dfrac{1}{1+\log n}$; $\varepsilon_n\to 0$:*

$$\sum_{i=1}^{\infty} P(\xi_i>\varepsilon_i\sqrt{i})<\infty, \tag{3.1}$$

$$\lim_{i\to\infty} \int_{|x|>\varepsilon_i\sqrt{i}} x^2\, dF_{\xi_i}(x)=0. \tag{3.2}$$

Then the set of limit functions of the sequence $\varphi_3,\varphi_4,\varphi_5,\dots$ under uniform convergence is almost certainly the set $\mathfrak{S}$.

Proof. 1) First we assume that all moments $E(|\xi_i|^l)$ exist and satisfy the condition (1.4). Let $\chi(t)$ be any function of the class $\mathfrak{S}$. Then we shall show that there exists a sequence $n_1,n_2,\dots$ such that almost certainly

$$\lim_{l\to\infty} \varphi_{n_l}(t)=\chi(t),$$

uniformly in $[0,1]$.

Let k be a sufficiently large natural number. We restrict our attention to indices n which are divisible by k, and we split the sum $S_n=\xi_1+\dots+\xi_n$ into k sums of $\frac{n}{k}$ terms each, letting

$$\zeta_1=\xi_1+\dots+\xi_{\frac{n}{k}},\ \zeta_2=\xi_{\frac{n}{k}+1}+\dots+\xi_{\frac{2n}{k}},\ \dots.$$

Then $\zeta_1,\dots,\zeta_k$ are independent r.v.'s.

Let $t\in[0,1]$ and $\delta>0$ be given. Lemma 4 with $\frac{n}{k}$, ζ_l, $\sqrt{\frac{n}{k}}+o(1)$ instead of n, S_n, $D(S_n)$ and with $y=t_l\sqrt{2k}\log\log n$ yields

$$P\left(\zeta_l \in \left(y\sqrt{\frac{n}{k}}, cy\sqrt{\frac{n}{k}}\right)\right) = P\left(\frac{\zeta_l}{\sqrt{2n\log\log n}} \in (t_l, ct_l)\right) \tag{3.3}$$
$$> e^{-k\log\log n(1+\varepsilon)t_l^2}.$$

Because of mutual independence of the ζ_l's, the probability that these conditions are satisfied simultaneously is greater than the product of the right-hand sides of (3.3), i.e. greater than

$$(\log n)^{-(1+\varepsilon)k\sum_{l=1}^{k} t_l^2}.$$

Now we choose

$$t_l = \chi\left(\frac{l+1}{k}\right) - \chi\left(\frac{l}{k}\right) \qquad (l=0,1,\ldots,k-1).$$

Let $\varepsilon' > 0$ be given. Then, writing $\int_0^1 \dot{\chi}(t)^2\,dt = I$, we have

$$\left|k\sum_{l=0}^{k-1} t_l^2 - I\right| < \varepsilon'$$

for all sufficiently large k. If $I<1$, the series

$$\sum_{m=1}^{\infty} (\log q^m)^{-(1+\varepsilon)I}, \; q>1,$$

diverges for sufficiently small ε, and hence the proof can be finished by applying the Borel-Cantelli Lemma, repeating the argument of Kolmogorov [1]. If $I=1$, we apply this argument to some function $\chi^* \in \mathfrak{S}$ whose corresponding integral is $I^*<1$ and which is sufficiently close to χ under the metric of uniform convergence. This finishes the proof of the theorem under the additional assumption made above.

2) In order to prove the general version of the theorem we apply the following truncation procedure: Let

$$\xi_i^0 = \begin{cases} \xi_i & \text{if } |\xi_i| < \varepsilon_i \sqrt{i}, \\ 0 & \text{otherwise} \end{cases}$$

and define $\xi_i^* = \xi_i^0 - E(\xi_i)$ $(i=1,2,\ldots)$. Then it can be shown that the r.v.'s ξ_i^* satisfy all conditions of Step 1): they are mutually independent, $E(\xi_i^*)=0$ and

$$D^2(\xi_i^*) = D^2(\xi_i) + o(1) = 1 + o(1)$$

as an immediate application of Chebyshev's inequality shows, using (3.2).

It remains to show that condition (1.4) is also satisfied. Obviously, one has

$$E(\xi_i^{*l}) = O(\varepsilon_i^l\, i^{\frac{l}{2}}) < c_i\, l!\, i^{\frac{l}{2}} (l=3,4,\ldots), \tag{3.4}$$

where $\lim_{i\to\infty} c_i = 0$. Letting $S_n^* = \sum_{i=1}^{n} \xi_i^*$, we have

$$D(S_n^*) = D(S_n)(1+o(1)) = \sqrt{n}(1+o(1)),$$

and thus (3.4) implies (1.4) for the sequence ξ_i^*. Hence the assertion holds for these r.v.'s as shown in Step 1). But then it is also valid for the initial sequence ξ_i in as much as

$$P(|\xi_i^* - \xi_i| > E(\xi_i)) = P(|\xi_i| > \varepsilon_i \sqrt{i}),$$

and therefore the Borel-Cantelli Lemma implies, because of (3.1), that the relation

$$|\xi_i^* - \xi_i| \leqq E(\xi_i) = o\left(\frac{1}{\varepsilon_i \sqrt{i}}\right) = o(1)$$

holds almost certainly.

Remark. It may be shown by a "shortening" technique, omitting certain terms at the end of each block ζ_1, ζ_2 etc., that the theorem remains true if the condition of mutual independence of the ξ_i's is replaced by a special form of weak dependence.

References

1. Kolmogorov, A.: Das Gesetz des iterierten Logarithmus. Math. Ann. **101**, 126–135 (1929)
2. Rényi, Wahrscheinlichkeitsrechnung. Berlin: Deutscher Verlag der Wiss. 1962
3. Strassen, V.: An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie verw. Gebiete **3**, 211–226 (1964)

Received November 26, 1979; in revised form October 15, 1981