Z. Wahrscheinlichkeitstheorie verw. Gebiete
52, 1–7 (1980)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete
© by Springer-Verlag 1980

A Strong Approximation of Partial Sums of i.i.d. Random Variables with Infinite Variance

Joop Mijnheer
Department of Mathematics, University of Leiden,
Wassenaarseweg 80, 2300 RA Leiden, The Netherlands

1. Summary and Introduction

Let $X_1, X_2, \ldots$ be a sequence of symmetric i.i.d. random variables with infinite variance in the domain of attraction of the normal distribution, i.e. their common distribution function F is such that

$$L(x) = \int_{|y| \leqq x} y^2\, dF(y) \tag{1.1}$$

is slowly varying at infinity with $\lim_{x\to\infty} L(x) = \infty$. Let the sequence a_n, $n = 1, 2, \ldots$, satisfy

$$n\, a_n^{-2} L(a_n) \to 1 \qquad \text{for } n \to \infty, \tag{1.2}$$

then it is well-known that the distribution of

$$a_n^{-1}(X_1 + \ldots X_n) \tag{1.3}$$

converges weakly to the normal distribution. See for example Feller (1971), p. 579.

In the case $L(x) \to 1$ for $x \to \infty$, i.e. X_1 is in the domain of *normal* attraction of the normal distribution, Major (1979) has shown the existence of a sequence of independent normally distributed random variables Y_n, $n = 1, 2, \ldots$, such that

$$\left| \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i \right| = o(n^{\frac{1}{2}}) \quad \text{a.s.} \qquad \text{for } n \to \infty. \tag{1.4}$$

Strassen's strong invariance principle is an easy consequence of this result.

In the case $L(x) \to \infty$ for $x \to \infty$ we shall prove a similar assertion. Suppose that for some positive non-decreasing slowly varying function L_2, with $\lim_{x\to\infty} \{L(x)\}^{-1} L_2(x) = \infty$,

0044-3719/80/0052/0001/$01.40

$$\int_1^{\infty} x^2\{L_2(x)\}^{-1}\,dF(x) \tag{1.5}$$

converges. Then we can construct a sequence of independent normally distributed random variables $Y_n, n=1,2,\ldots$, and a sequence b_n, such that

$$\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i\right| = o(b_n) \quad \text{a.s.} \qquad \text{for } n\to\infty, \tag{1.6}$$

with

$$n b_n^{-2} L_2(b_n) \to 1 \qquad \text{for } n\to\infty. \tag{1.7}$$

Note that the case where the random variables have an infinite variance is different from the case where the random variables have a finite variance. It is well-known that the condition $EX^2<\infty$ is equivalent with $\sum P(|X_n|>\varepsilon\sqrt{n})<\infty$, for some $\varepsilon>0$. If we define $X'_n = X_n \cdot 1_{[-\varepsilon\sqrt{n},\varepsilon\sqrt{n}]}$ then, w.p. 1, we have

$$|X_1+\ldots+X_n-(X'_1+\ldots+X'_n)| = o(\sqrt{n}) \qquad \text{for } n\to\infty.$$

In the case $EX^2=\infty$ we want to state a similar assertion. It follows from Feller (1968) Lemmas 3.2 and 3.3 that $P\left(\left|\sum_{i=1}^{n} X_i\right| > a_n \text{ i.o.}\right)=0$ or 1 according as $\sum P(|X_n|>a_n)$ converges or diverges, where a_n satisfies (1.2). However, the last series diverges as follows from the divergence of the integral

$$\int_a^{\infty} x^2\{L(x)\}^{-1}\,dF(x), \tag{1.8}$$

where $a=\inf\{x: x>0 \text{ and } L(x)>0\}$.

Divergence of the integral (1.8) easily follows from the fact that $\lim_{x\to\infty} L(x)=\infty$ and in the proof of the assertion we do not use the assumption that L is slowly varying at infinity. Thus, with probability one, the relation $\sum_{i=1}^{n} X_i > a_n$ will occur infinitely often. This implies that we cannot improve the upperbound in the right-hand side of (1.6) with a_n in stead of b_n. In other words, the bound in (1.6) is sharp.

The symmetry assumption is only used at one place to guarantee that the truncated random variables have zero expectation. I can only prove an assertion as given in (1.6) for non-symmetric random variables with some additional assumptions. In order to avoid those details about the question when the sum of the expectations of the truncated random variables is negligible, we shall only consider the symmetric case.

Our method to prove (1.6) can also be used to prove Major's result. He applies a strengthened version of the Berry-Esseen inequality. We shall apply the Skorohod representation for the truncated random variables.

In Sect. 2 we state and prove the main result. In Sect. 3 we discuss the result and give some examples.

2. Main Result

The main result of this paper is the following theorem.

Theorem 2.1. *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. symmetric random variables with $EX_1=0$ and common distribution function F. Let the function L defined by* (1.1) *be slowly varying at infinity with $\lim_{x\to\infty} L(x)=\infty$ and suppose that for some positive non-decreasing slowly varying function L_2, with $\lim_{x\to\infty} \{L(x)\}^{-1} L_2(x)=\infty$ the integral* (1.5) *converges. Let $\{b_n\}$ satisfy* (1.7). *Then we can define a sequence of positive real numbers τ_n, $n=1,2,\ldots$, and a probability space on which we can define $\{X_n\}$ and a sequence of normally distributed independent random variables Y_n, $n=1,2,\ldots$, with $EY_n=0$, for all n, $EY_n^2=\tau_n$ and their partial sums satisfy* (1.6).

The proof of Theorem 2.1 follows from some lemmas proved by Feller (1968) and the following lemmas. The lemmas of Feller imply that dominating values of $X_1+\ldots+X_n$ are as rule due to the influence of one big observation X_k. This is a well-known result for non-normal stable distributions. See, for example, Feller (1946) or Mijnheer (1975). Here we notice a different behaviour for partial sums in the case where the random variables have a finite variance and the case with an infinite variance.

Define, for $n=1,2,\ldots$, the random variables $X'_n=X_n\cdot 1_{[-b_n,b_n]}$. From the Lemmas 3.1 and 3.2 of Feller (1968) it follows that convergence of the integral (1.5) implies

$$|X_1+\ldots+X_n-(X'_1+\ldots+X'_n)|=o(b_n) \qquad \text{a.s.} \qquad \text{for } n\to\infty. \tag{2.1}$$

The next lemma is essentially the same as Theorem 2.1 of Kostka (1972). We omit the proof.

Lemma 2.1. *Let $\{T_n\}$ be a sequence of Skorohod stopping times with $ET_n=\tau_n$, $n=1,2,\ldots$, and let $\{W(t)\colon 0\leq t<\infty\}$ be a Wiener process. Suppose*

$$\limsup c_n^{-1}\left|\sum_{i=1}^{n} T_i-\sum_{i=1}^{n}\tau_i\right|\leqq k \quad \text{a.s.} \tag{2.2}$$

for some $k>0$, where $\{c_n\}$ is a sequence of positive numbers. Then

$$\limsup (c_n\log n)^{-\frac{1}{2}}\left|W\left(\sum_{i=1}^{n} T_i\right)-W\left(\sum_{i=1}^{n}\tau_i\right)\right|<\infty \quad \text{a.s.} \tag{2.3}$$

Note that in Theorem 2.1 of Kostka (1972) we have $\tau_n=1$, $n=1,2,\ldots$.

Define a sequence of positive real numbers d_n satisfying

$$d_n^2 L(d_n)\sim n(\log n)^{-3}(\log_2 n)^{-1-\delta} \qquad \text{for } n\to\infty, \tag{2.4}$$

where δ is some positive constant. We define τ_n by

$$\tau_n=\int_{|x|\leqq d_n} x^2\, dF(x)=L(d_n). \tag{2.5}$$

Lemma 2.2. *Let $\{T_n\}$ be a sequence of Skorohod stopping times such that, for all n, $W(T_n)$ has the same distribution as $X_n \cdot 1_{[-d_n, d_n]}$, where d_n satisfies* (2.4). *Let $\{\tau_n\}$ be defined by* (2.5). *Then, for $n \to \infty$,*

$$\left|W\left(\sum_{i=1}^{n} T_i\right) - W\left(\sum_{i=1}^{n} \tau_i\right)\right| = o(b_n) \quad a.s., \tag{2.6}$$

where b_n satisfies (1.7).

Proof. The Skorohod representation (see, for example, Breiman (1968), p. 276) yields the existence of a sequence of stopping times $\{T_n\}$ such that, for all n, $\sum_{i=1}^{n} X_i \cdot 1_{[-d_i, d_i]}$ has the same distribution as $W\left(\sum_{i=1}^{n} T_i\right)$ and $ET_n = \sigma^2(X_n \cdot 1_{[-d_n, d_n]}) = \tau_n$.

Take $c_n = n(\log n)^{-1}$. We shall prove.

$$\left|\sum_{i=1}^{n} T_i - \sum_{i=1}^{n} \tau_i\right| = o(c_n) \qquad \text{a.s.} \qquad \text{for } n \to \infty. \tag{2.7}$$

The stopping times T_n satisfy the inequalities, for $p > \frac{1}{2}$,

$$a_p E|W(T_n)|^{2p} \leqq ET_n^p \leqq A_p E|W(T_n)|^{2p}.$$

(See Sawyer (1974).) The inequality on the right-hand side implies

$$ET_n^2 \leqq A \int_{-d_n}^{d_n} x^4\, dF(x) \leqq A d_n^2 L(d_n).$$

Then $\sum c_n^{-2} \operatorname{var}(T_n) \leqq A \sum c_n^{-2} d_n^2 L(d_n) \leqq A \sum n^{-1} (\log n)^{-1} (\log_2 n)^{-1-\delta} < \infty$. Lemma 3.27 of Breiman (1968) implies (2.7). Assertion (2.6) follows from (2.7) and Lemma 2.1.

Define $X_n'' = X_n \cdot 1_{[-\varepsilon b_n, -d_n] \cup [d_n, \varepsilon b_n]}$.

Lemma 2.3. *Assume the convergence of the integral* (1.5). *Let the sequences b_n and d_n satisfy* (1.7) *and* (2.4) *and let X_n'' be defined as above then, with probability one,*

$$|X_1'' + \ldots + X_n''| = o(b_n) \qquad \text{for } n \to \infty.$$

Proof. Let $p > 0$. From the Lemmas 4.1 and 5.1 of Feller (1968) it follows that, w.p. 1,

$$\left|\sum_{k=1}^{n} X_k \cdot 1_{[-\varepsilon b_k, -b_k v^{-p}(b_k)] \cup [b_k v^{-p}(b_k), \varepsilon b_k]}\right| = o(b_n)$$

for $n \to \infty$. Kesten (1972) has discussed some of the arguments. In our case the proofs are simpler because both L and L_2 are non-decreasing and slowly varying at infinity. Thus we may restrict our attention to the random variables $\tilde{X}_n$ that are symmetric truncated at d_n and $f_n = b_n v^{-p}(b_n)$, with arbitrary $p > 0$.

Take a subsequence n_k, $k = 1, 2, \ldots$, satisfying $n_k \sim n_{k-1}(1 + \phi^{-1}(k))$ for $k \to \infty$, where $\lim \phi(k) = \infty$ and properly chosen as we shall see below. Define the sequence of events

$$A_k\colon \bigcup_{\substack{n_{k-1}<i,j\leqq n_k\\ i\neq j}} \{\tilde{X}_i \neq 0 \text{ and } \tilde{X}_j \neq 0\}.$$

Then

$$\begin{aligned} P(A_k) &= \sum_{\substack{n_{k-1}<i,j\leqq n_k\\ i\neq j}} P(\tilde{X}_i \neq 0 \text{ and } \tilde{X}_j \neq 0)\\ &\leqq \{(n_k - n_{k-1}) \max_{n_{k-1}<n\leqq n_k} (c\, d_n^{-2} L_2(d_n)\, I(d_n, f_n])\}^2, \end{aligned}$$

where $I($ $]$ is the measure attributed by the integral (1.8) to the given interval.

By our choice of n_k we have

$$P(A_k) \leqq c_1 (n_k - n_{k-1})^2\, d_{n_k}^{-4}\, L_2^2(d_{n_k})\, I^2(d_{n_k}, f_{n_k}].$$

Now we can choose the function ϕ such that

$$P(A_k) \leqq c_2\, I(f_{n_{k-1}}, f_{n_k}].$$

This implies $\sum P(A_k) < \infty$ and therefore, by the Borel-Cantelli lemma $P(A_k$ i.o.) $=0$. Thus, w.p.1, ultimately each block $n_{k-1} < n \leqq n_k$ contains at most one component $\tilde{X}_n \neq 0$.

Define the random variables Z_k, $k=1,2,\ldots$, by

$$Z_k = \begin{cases} \tilde{X}_n & \text{if } \tilde{X}_n \neq 0 \text{ for some } n \in (n_{k-1}, n_k] \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} \sigma^2(Z_k) &\leqq (n_k - n_{k-1}) \max \int_{d_n}^{f_n} x^2\, dF(x)\\ &\leqq c(n_k - n_{k-1})\, L_2(f_{n_k})\, I(d_{n_k}, f_{n_k}]. \end{aligned}$$

Thus

$$\begin{aligned} \sum b_{n_k}^{-2}\, \sigma^2(Z_k) &\leqq \sum (n_k - n_{k-1})\, L_2(f_{n_k})\, n_k^{-1}\, L_2^{-1}(b_{n_k})\, I(d_{n_k}, f_{n_k}]\\ &\leqq \sum (n_k - n_{k-1})\, n_k^{-1}\, I(d_{n_k}, f_{n_k}], \quad L_2 \text{ is non-decreasing}\\ &\leqq \sum I(f_{n_{k-1}}, f_{n_k}] \end{aligned}$$

by the choice of ϕ in the definition of the sequence n_k.

Now a well-known a.s. convergence criterion implies the assertion of the lemma.

3. Examples and Discussion of Our Results

Feller (1968) has proved that the law of the iterated logarithm only holds if the integral

$$\int_0^\infty \{L(x) \log_2 x\}^{-1}\, x^2\, dF(x) \tag{3.1}$$

converges. (He did not assume that L is slowly varying at infinity.) In our examples we take $L(x)$ equal to $c \log_2 x$, $\exp(\log x / \log_4 x)$ and $c \log_3 x$.

Example 1. Let f be the density of X_1 given by

$$f(x)=\begin{cases} c|x|^{-3}(\log|x|)^{-1} & \text{for } |x|>e \\ 0 & \text{otherwise.} \end{cases}$$

Then, for $x\to\infty$, $L(x)\sim 2c\log_2 x$ and $L_2(x)\sim(\log_2 x)^{1+\varepsilon}$ with $\varepsilon>0$. Therefore

$$a_n\sim\{2cn\log_2 n\}^{\frac{1}{2}},\qquad b_n\sim\{c_1 n(\log_2 n)^{1+\varepsilon}\}^{\frac{1}{2}},$$

and

$$\tau_n=\sigma^2(Y_n)\sim 2c\log_2 n.$$

From the theory of slowly varying functions (c.f. Feller (1971)) we have

$$\sigma^2(Y_1+\ldots+Y_n)\sim 2cn\log_2 n.$$

The law of the iterated logarithm for normally distributed random variables and our main result imply, for $n\to\infty$,

$$\limsup n^{-\frac{1}{2}}(\log_2 n)^{-1}(X_1+\ldots+X_n)=k \quad \text{a.s.},$$

where k is a (known) constant.

Example 2. Let the function L defined by (1.1) for sufficiently large x be given by $L(x)=\exp(\log x/\log_4 x)$. We can take $L_2(x)=\log x\log_2 x\log_3 x(\log_4 x)^{\varepsilon} L(x)$ for some $\varepsilon>0$. Simple calculations show that, for $n\to\infty$,

$$a_n^{-1}b_n\to\infty,\qquad \{L(a_n)\}^{-1}L(b_n)\to\infty,\qquad \{L(d_n)\}^{-1}L(a_n)\to\infty.$$

Obviously the integral (3.1) diverges. Our main result shows the existence of the random variables, $Y_1, Y_2\ldots$ such that (1.6) holds. The L.I.L. for normally distributed random variables yields, for $n\to\infty$,

$$\limsup(2nL(d_n)\log_2 n)^{-\frac{1}{2}}(Y_1+\ldots+Y_n)=k \quad \text{a.s.},$$

where k is a (known) constant.

Note that $\sigma^2(Y_1+\ldots+Y_n)\sim cnL(d_n)$ for $n\to\infty$. On the other hand convergence of the integral (1.5) implies

$$\limsup b_n^{-1}(X_1+\ldots+X_n)=0 \quad \text{a.s.} \qquad \text{for } n\to\infty$$

and divergence of the integral (1.8) implies

$$\limsup a_n^{-1}(X_1+\ldots+X_n)=\infty \quad \text{a.s.} \qquad \text{for } n\to\infty.$$

Divergence of (3.1) implies

$$\limsup(2a_n^2\log_2 a_n)^{-\frac{1}{2}}(X_1+\ldots+X_n)=\infty \quad \text{a.s.} \qquad \text{for } n\to\infty.$$

For this particular choise of L the proof of Lemma 2.3 is simpler than the one we gave in Sect. 2. It also follows from our main result and Lemma 2.3 that we can find normally distributed random variables Y_i, $i=1,2,\ldots$, with $EY_i=0$ and $\sigma^2(Y_1+\ldots+Y_n)\sim a_n^2$ such that (1.6) holds. In view of the remarks made at the end of Section 1, it is clear that the estimate is sharp.

Example 3. Let the density f of X_1 be given by

$$f(x)=\begin{cases} c|x|^{-3}(\log|x|)^{-1}(\log_2|x|)^{-1} & \text{for } |x|>e^e \\ 0 & \text{otherwise,} \end{cases}$$

with c choosen such that $\int_{-\infty}^{+\infty} f(x)\,dx=1$. Then, for large x, $L(x)=2c\log_3 x$ and $L_2(x)\sim(\log_3 x)^{1+\varepsilon}$ with $\varepsilon>0$. For $n\to\infty, a_n^2\sim 2cn\log_3 n$, $b_n^2\sim n(\log_3 n)^{1+\varepsilon}$ and $L(d_n)\sim L(a_n)$.

Thus, as in example 1, we have, for $n\to\infty$,

$$\limsup(n\log_2 n\log_3 n)^{-\frac{1}{2}}(X_1+\ldots+X_n)=k \quad \text{a.s.,}$$

for some constant k. Here we also have convergence of integral (3.1).

As explained in Sect. 1 we need convergence of integral (1.5) to obtain a strong approximation. The original Condition (1.1) is appropriate for a weak approximation, i.e.

$$a_n^{-1}\sup_{1\leq k\leq n}\left(\sum_{i=1}^{k} X_i-\sum_{i=1}^{k} Y_i\right)\to 0 \qquad \text{in probability,}$$

where Y_k is normally distributed with $EY_k=0$ and $\sigma^2(Y_k)=L(a_k)$ and a_k satisfies (1.2). From the theory of slowly varying functions we have

$$\sigma^2(Y_1+\ldots+Y_k)\sim kL(a_k)\sim a_k^2 \qquad \text{for } k\to\infty.$$

In the case where the random variables have a finite variance a similar assertion easily follows from Donsker's theorem. See for example Billingsley (1968). For random variables with an infinite variance the proof of assertion (3.2) follows the same lines.

I would like to thank Péter Major for his comments on the first draft of this paper.

References

Billingsley, P.: Convergence of probability measures. New York: Wiley 1968

Breiman, L.: Probability. Reading (Mass): Addison-Wesley 1968

Feller, W.: A limit theorem for random variables with infinite moments. Amer. J. Math. **68**, 257–262 (1946)

Feller, W.: An extension of the law of the iterated logarithm to variables without variance. J. Math. Mech. **18**, 343–355 (1968)

Feller, W.: An introduction to probability theory and its applications. Vol. II. Second edition. New York: Wiley 1971

Kostka, D.G.: Deviations in the Skorohod-Strassen Approximation Scheme. Z. Wahrscheinlichkeitstheorie verw. Gebiete **24**, 139–153 (1972)

Major, P.: An improvement of Strassen's invariance principle. Ann. Probability **7**, 55–61 (1979)

Mijnheer, J.L.: Sample path properties of stable processes. Math. Centre tracts **59.** Amsterdam: Mathematisch Centrum 1975

Sawyer, S.: The Skorokhod representation. The Rocky Mountain Journal of Mathematics **4**, 579–596 (1974)

Received June 13, 1978; in revised form September 24, 1979