Z. Wahrscheinlichkeitstheorie verw. Geb. 13, 321–332 (1969)

On the Deviations in the Skorokhod-Strassen Approximation Scheme

J. KIEFER*

Summary. In deriving his strong invariance principles, Strassen used a construction of Skorokhod: if the univariate d.f. F has first, second, and fourth moments 0, 1, and $\beta<\infty$, respectively, then there is a probability space on which are defined a standard Brownian motion $\{\xi(t), t\geqq 0\}$ and a sequence of nonnegative i.i.d. Skorokhod random variables $\{T_i, i>0\}$ such that

$$\left\{\xi\left(\sum_1^{n+1} T_i\right)-\xi\left(\sum_1^{n} T_i\right), n\geqq 0\right\}$$

are i.i.d. with d.f. F. Let

$$Z=\limsup_{n\to\infty} \pm\left[\xi\left(\sum_1^n T_i\right)-\xi(n)\right]\Big/[n(\log n)^2\log\log n]^{\frac{1}{4}}.$$

Strassen showed $Z=O(1)$ wp 1. We prove $Z=(2\beta)^{\frac{1}{4}}$ wp 1. Consequently $Z=0$ wp 1 implies F is Gaussian, answering a special case of a question of Strassen. Analogous results hold for cases where $\xi\left(\sum_1^n T_i\right)$ is not a sum of independent random variables.

1. Introduction

Let F be a univariate d.f. satisfying

$$\begin{aligned}
&\text{(a)} && \int_{-\infty}^{\infty} x\,dF(x)=0,\\
&\text{(b)} && \int_{-\infty}^{\infty} x^2\,dF(x)=1,\\
&\text{(c)} && \int_{-\infty}^{\infty} x^4\,dF(x)<\infty.
\end{aligned}\tag{1}$$

Skorokhod [18] showed how to construct a probability space on which are defined (i) a sequence $\{X_i, 1\leqq i<\infty\}$ of i.i.d.r.v.'s[1] with common d.f. F, (ii) a Brownian motion $\{\xi(t), t\geqq 0\}$ of standard normalization $(E\xi(t)\equiv 0, E\xi^2(t)=t)$, and (iii) a sequence of non-negative random variables $\{T_i, 1\leqq i<\infty\}$, such that, writing

$$\begin{aligned}
&\beta=\operatorname{var}(T_1),\\
&U_0=0,\quad U_n=\sum_1^n T_i && (n>0),\\
&S_n=\xi(U_n) && (n\geqq 0),\\
&Y_n=S_n-S_{n-1} && (n>0),\\
&\Delta_n=\sigma\text{-field generated by } X_1,\dots,X_n \text{ and}\\
&\qquad\{\xi(t), t\leqq U_n\} && (n>0),
\end{aligned}\tag{2}$$

* Research under ONR Contract 401(50). Reproduction for any purpose of the U.S. Government is permitted.

1. Abbreviations used in this paper are listed in the last paragraph of this section.

the following properties hold:

(a) the ξ-process is independent of $\{X_i\}$; T_n is Δ_n-measurable; $\{\xi(U_n+t)-\xi(U_n), t>0\}$ is independent of Δ_n;
(b) the T_i are i.i.d. with $E\{T_i\}=1$, (3)
(c) $\beta<\infty$,
(d) the sequence $\{Y_i, i>0\}$ has the same law as $\{X_i, i>0\}$.

We will refer to the setup described in the previous paragraph as the *i.i.d. case*, even though the general Skorokhod construction omits (1c) and (3c). Strassen [21], p. 333, has mentioned extensions, and has kindly informed the author that a more complete description is that Dambis and Dubins-Schwarz obtained analogues of Skorokhod's result for continuous parameter martingales, whereas Frank Jonas has carried out in detail the construction suggested by an observation (of Strassen and, independently, of David Freeman) that Skorokhod's construction extends to the case where

$$\left\{\sum_1^n X_i, 1\leqq n<\infty\right\}$$

is a martingale, with appropriate modification in the conclusion (3); subsequently Dubins [5] and Hall [9] obtained other constructions in the martingale case. We need not be more precise at this point, since the only properties we use of this *martingale case* (with restrictions whose consequences parallel those of (1c)) will be listed in (12) and (28) below; in particular, our results apply not just to Skorokhod's construction, but rather to general stopping variables, including those of Dubins and of Hall.

Strassen [19, 21] used the Skorokhod representation and its martingale extension in developing his beautiful strong invariance principles. Theorem 1.5 of [21] states that, in the i.i.d. case, as $n\to\infty$, with S_n obtained as in (2)–(3),

$$\xi(n)-S_n=O\left((n\log\log n)^{\frac{1}{4}}(\log n)^{\frac{1}{2}}\right) \qquad \text{wp 1.} \tag{4}$$

Strassen asks whether the existence of a probability space on which are defined a Brownian notion ξ and i.i.d. sequence $\{S_{n+1}-S_n, n\geqq 0\}$ with d.f. F, for which (4) holds with O replaced by o, implies that F is Gaussian. The answer in this generality, when S_n is not required to be obtained from stopping variables as in (2), is unknown. If S_n is assumed to be obtained as $\xi(U_n)$ from stopping times U_n as in (2)–(3), the question is answered affirmatively by the following result of the present paper:

Theorem 3. *In the i.i.d. case, under the assumptions of* (1), *if* $\beta>0$, *then for either choice of sign*

$$\limsup_{n\to\infty}\pm[\xi(n)-\xi(U_n)]/[2\beta n(\log n)^2\log\log n]^{\frac{1}{4}}=1 \qquad \text{wp 1.} \tag{5}$$

(Of course, when $\beta=0$ we have $T_n=1$ w.p. 1 and hence F is standard Gaussian.) Theorem 3 is obtained as a corollary of the more general Upper Class Theorem 1 and Lower Class Theorem 2, which do not require the Y_i to be i.i.d. (nor even for $\{\xi(U_n), n>0\}$ to be a martingale), but only the validity of certain properties (12) and (28) which are analogous to those of (1) or which are associated with usual proofs of the LIL.

In this paper it is only in the i.i.d. case (Theorem 3) that we verify in detail that such a simple condition as (1c) can be used to invoke Theorems 2 and 3. It is not hard to find sufficient conditions to invoke Theorems 2 and 3 for various cases where the Y_i are not i.i.d., but the known conditions are not sharp in the following sense that (1c) is: if (3c) fails in the i.i.d. case, the U_n do not obey the LIL (see Strassen [20]), and (4) will not generally hold.

We remark that Breiman [2] has considered the more general case of i.i.d. X_i not necessarily satisfying (1c), obtaining in this case bounds which are of a larger order than that of (4), and also necessary moment conditions for such behavior. While we do not consider such results in the present paper, we believe the techniques used herein may be adaptable to such cases where (1c) fails, through the use of analogues of standard LIL developments. To those who may be interested in pursuing such results, we will try to point out, where they occur, our few technical departures from usual patterns of proof of LIL-type theorems or from ideas of [12] and [13] which we have used.

Also, Breiman [3], Section 13.6, obtained a representation of the sample d.f. for uniformly distributed r.v.'s, in terms of the Brownian bridge, by using Skorokhod's construction for exponential r.v.'s $\{X_i\}$. It is easy to see that $n^{\frac{1}{4}}$ times the maximum error in this representation satisfies the right side of (4). Later Brillinger [4], evidently unaware of Breiman's book, obtained this result independently. See [15] for further comments.

The developments of this paper yield certain other results, as the interested reader will find it easy to verify. We now list some of these.

Theorem 4. *If* $\{\xi(t), t \geqq 0\}$ *is standard Brownian motion and* $0<\beta<\infty$, *then, for either choice of sign,*

$$\limsup_{t\to\infty} \Big\{ \sup_{|\tau-t|<[2\beta t \log\log t]^{\frac{1}{2}}} \pm[\xi(t)-\xi(\tau)]/[2\beta t(\log t)^2 \log\log t]^{\frac{1}{4}} \Big\} = 1 \qquad \text{wp } 1$$

and

$$\limsup_{t\downarrow 0} \Big\{ \sup_{|\tau-t|<[2\beta t^3 \log|\log t|]^{\frac{1}{2}}} \pm[\xi(t)-\xi(\tau)]/[2\beta t^3(\log t)^2 \log|\log t|]^{\frac{1}{4}} \Big\} = 1 \qquad \text{wp } 1.$$

The first of these results is obtained easily from the developments of the succeeding sections, and the second follows from the first on using time inversion $(\{t\,\xi(t^{-1}), t>0\}$ is a standard Brownian motion$)$ and elementary estimates. Theorem 4 is related to the domain of Lévy's Hölder condition, the upper bound half of which in fact follows easily from (7) below. (See Ito-McKean [11], pp. 36–38).

The present paper also yields results concerning the *sample quantile process* $\eta^{(2)}$ introduced in Section 6 of [13], where we proved results for sample quantiles corresponding to the main results of [19] for sums of independent r.v.'s (Theorem 5 of [13]) and the *strong form of the LIL for sample quantiles* (Theorem 4 of [13]) corresponding to that of Feller [6] for sums of independent r.v.'s (strengthening the standard form of the LIL for sample quantiles obtained in [1]). Without taking the space here to define $\eta^{(2)}$, we remark that the result of [1] or Theorem 1 of [13] shows immediately that the order (with the right constant) of the deviation of ξ from $\eta^{(2)}$ is exactly the same as that of ξ from the random walk process $\eta^{(1)}$ (defined in Section 6.7 of [13]), and the latter is given at once by Theorem 3 of the present paper. See [15] regarding variation in p of p-tiles.

Finally, a development like that of Lemma 1 yields probability bounds for the difference between the maximum and minimum of partial sums and more general processes, analogous to Lévy-Kolmogorov-Hajek-Renyi bounds for the maximum or maximum norm. We will discuss this elsewhere [14].

It would be interesting to obtain "strong forms" corresponding to Theorems 3 and 4.

The following standard notation is used in this paper: The complement of an event A is denoted $\bar{A}$. The natural numbers and nonnegative reals are denoted by N and R^+, respectively. For real x, the greatest integer $\leqq x$ is denoted by $\operatorname{int}\{x\}$. All "orders" (O or o) or asymptotic relations ($\sim$) refer to behavior as the exhibited dummy variable n or $r \to \infty$. We abbreviate "random variable" by "r.v.", "independent and identically distributed" by "i.i.d.", "distribution function" by "d.f.", "infinitely often" by "i.o.", "almost all n" (i.e., all n in N except for a finite number) by "a.a.n.", "law of the iterated logarithm" by "LIL", "with probability one" by "wp 1". Whenever we write such summation operations as $\sum_r$ it will be understood that the summation is over all r in N which are large enough that expressions like $\log \log n_r$ which appear in the summand are real.

The author is grateful to Volker Strassen for helpful comments.

2. Upper Class Result

In this section we suppose there is given a probability space on which are defined a standard Brownian motion $\{\xi(t), t \geqq 0\}$ and a sequence of nonnegative r.v.'s $\{T_i, i > 0\}$. We define U_n as in (2). The only assumption we require is an upper class LIL-type estimate (12) on the sequence $\{U_n\}$; questions of dependence among the U_n and ξ do not otherwise enter.

As is often the case in LIL-type results, the upper class result is much easier to prove than the lower class result. Thus, Strassen rightfully calls (4) "easy" in the i.i.d. case, and our details are required only to obtain the right constant in (13). At the same time, a little delicacy is needed to obtain that constant. (It amounts to using (21) and (7) rather than an estimate like line 4, p. 22 of [2], and to replacing the n_r of [2] by numbers which yield our (13) rather than the larger order of Theorem 2 of [2].)

We remark also that, although Theorem 1 can be proved using approximately geometric times $n'_r \sim c^r$ as in the lower class proof of Theorem 2, instead of the n_r used below, it is not very enlightening to use such n'_r. This is because the desired result is concerned with the magnitude of oscillations of ξ over relatively short periods of time, which do not persist over long periods. In the lower class result we use such widely spaced n'_r and study what happens between succesive n'_r, because of such developments as the use of (50) to prove (43). (Other reasons for this choice will be found in the proof of Theorem 2 and the remarks which follow it.) But to use such n'_r in the upper class proof and then to subdivide the period between successive n'_r to look like the periods between the n_r below, is an unnecessary complication. Moreover, the direct use of the n_r makes the source of (4) and (5) transparent and points the way toward the right normalization in the cases where (1c) fails (as mentioned in the discussion of [2] in Section 1), although we have not investigated whether proof technicalities dictate the use of analogues of the n_r or the n'_r in such cases.

Before stating Theorem 1, we prove a simple lemma; it is related to the developments on pp. 36–38 of Ito-McKean [11], and it can also be obtained from the results on pp. 329–330 of Feller [7] (or, for large c, which is all that matters, from (3.6) of [8]), or from an application of the results of p. 651 of [12], although it will be shorter to prove it directly by a simple method which has further applications [14].

Lemma 1. *If ξ is standard Brownian motion and T, L, δ, c are positive values with $T<L$, then*

$$P\{\sup_{0\leqq t_1<t_2\leqq T}|\xi(t_1)-\xi(t_2)|\geqq c\}\leqq\frac{8T^{\frac{1}{2}}}{c(2\pi)^{\frac{1}{2}}}e^{-c^2/2T} \tag{6}$$

and

$$P\{\sup_{0\leqq t_1<t_2\leqq L,\,|t_1-t_2|\leqq T}|\xi(t_1)-\xi(t_2)|\geqq c\}\leqq\frac{8(L-T+\delta)(T+2\delta)^{\frac{1}{2}}}{\delta c(2\pi)^{\frac{1}{2}}}e^{-c^2/2(T+2\delta)}. \tag{7}$$

Proof. We first prove (6). Let Γ be the event that for some v and w in $[0,T]$ with $v<w$ we have $|\xi(w)-\xi(v)|\geqq c$. The r.v. W is defined on Γ to be the least such w, and V is the least v corresponding to that w. Let $\Gamma^+=\Gamma\cap\{\xi(W)>\xi(V)\}$ and $\Gamma^-=\Gamma\cap\{\xi(W)<\xi(V)\}$. The event $W=w_1$ depends only on $\{\xi(t), 0\leqq t\leqq w_1\}$; hence, we have (as in Lévy's treatment of $\sup_{0\leqq t\leqq T}\xi(t)$)

$$P\{\xi(T)\geqq\xi(W)\,|\,\Gamma^+;\,W=w_1\}=\tfrac{1}{2} \tag{8}$$

and thus, since $\xi(W)-\xi(V)=c$ wp 1 on Γ^+,

$$\begin{aligned}P\{\Gamma^+\}&\leqq 2P\{\Gamma^+;\,\xi(T)\geqq\xi(W)\}\\&=2P\{\Gamma^+;\,\xi(T)-\xi(V)\geqq c\}\\&\leqq 2P\{\xi(T)-\min_{0\leqq t\leqq T}\xi(t)\geqq c\}\\&=2P\{\max_{0\leqq t\leqq T}[\xi(T)-\xi(T-t)]\geqq c\}.\end{aligned} \tag{9}$$

But the process $\{\xi(T)-\xi(T-t), 0\leqq t\leqq T\}$ is again a standard Brownian motion, so that Lévy's argument yields

$$P\{\Gamma^+\}\leqq 4P\{\xi(T)\geqq c\}. \tag{10}$$

The analogous result for Γ^- and the standard inequality for Gaussian tail probabilities yields (6).

We now turn to (7). Let $J=1+\mathrm{int}\{(L-T)/\delta\}$, and for $1\leqq j\leqq J$ define the event

$$F_j=\{\sup_{(j-1)\delta\leqq t_1,t_2\leqq T+j\delta}|\xi(t_1)-\xi(t_2)|\geqq c\}. \tag{11}$$

For every interval $[t_1,t_2]$ of length $\leqq T$ contained in $[0,L]$ there is at least one $j(1\leqq j\leqq J)$ such that $[t_1,t_2]\subset[(j-1)\delta,\,T+j\delta]$. Hence, the left side of (7) is no greater than $\sum_1^J P\{F_j\}$. Each $P\{F_j\}$ is bounded by (6) on replacing T there by $T+2\delta$. Since $J\leqq(L-T+\delta)/\delta$, we obtain (7).

We now state our upper class theorem.

Theorem 1. *Suppose there is a finite value $\beta > 0$ such that*

$$\limsup_{n\to\infty} |U_n - n|/[2\beta n \log\log n]^{\frac{1}{2}} \leqq 1 \qquad \text{wp } 1. \tag{12}$$

Then

$$\limsup_{n\to\infty} |\xi(n) - \xi(U_n)|/[2\beta n (\log n)^2 \log\log n]^{\frac{1}{4}} \leqq 1 \qquad \text{wp } 1. \tag{13}$$

Proof. Let $\varepsilon > 0$ be given, and write

$$q_n = [2\beta(1+\varepsilon) n \log\log n]^{\frac{1}{2}} \tag{14}$$

and

$$d_n = [2\beta(1+\varepsilon)^3 n (\log n)^2 \log\log n]^{\frac{1}{4}}. \tag{15}$$

Let $\{n_r\}$ be any increasing sequence of natural numbers satisfying

$$n_r = \{2^{-1}\beta(1+\varepsilon) r^2 \log\log r\} [1 + o(r^{-1})]. \tag{16}$$

It is easy to compute that

$$n_{r+1} - n_r \sim \beta(1+\varepsilon) r \log\log r \sim q_{n_r}. \tag{17}$$

Let

$$M_n = \{t\colon t \in R^+;\ |t-n| < q_n\} \tag{18}$$

and

$$\begin{aligned} M_r^* = \{(t, n)\colon t \in R^+;\ n \in N;\ |t-n| < q_{n_{r+1}};\\ t, n \in [n_r - q_{n_{r+1}}, n_{r+1} + q_{n_{r+1}}]\}. \end{aligned} \tag{19}$$

We define the events

$$A_n = \{\sup_{t\in M_n} |\xi(t) - \xi(n)| > d_n\} \tag{20}$$

and

$$A_r^* = \{\sup_{(t,n)\in M_r^*} |\xi(t) - \xi(n)| > d_{n_r}\}. \tag{21}$$

Our theorem will be proved if we show that

$$P\{A_n \text{ occurs i.o.}\} = 0. \tag{22}$$

If A_n occurs for some n satisfying $n_r \leqq n \leqq n_{r+1}$, then clearly A_r^* occurs. Hence, by the Borel-Cantelli lemma, (22) will follow from

$$\sum_r P\{A_r^*\} < \infty. \tag{23}$$

Let A_r^{**} be the event defined by (21) when N is replaced by R^+ in (19). Clearly A_r^* implies A_r^{**}. We bound $P\{A_r^{**}\}$ by applying (7) with

$$\begin{aligned} L &= n_{r+1} - n_r + 2q_{n_{r+1}} \sim 3\beta(1+\varepsilon) r \log\log r,\\ T &= q_{n_{r+1}} \sim \beta(1+\varepsilon) r \log\log r,\\ c &= d_{n_r} \sim [2\beta(1+\varepsilon)^2 r \log r \log\log r]^{\frac{1}{2}},\\ \delta &= r, \end{aligned}$$

and obtain

$$P\{A_r^*\} \leqq P\{A_r^{**}\} \leqq \frac{16\, q_{n_{r+1}}^{\frac{3}{2}}[1+o(1)]}{r\, d_{n_r}(2\pi)^{\frac{1}{2}}} \exp\{-d_{n_r}^2/2(q_{n_{r+1}}+2r)\}$$
$$\sim \frac{8\beta(1+\varepsilon)^{\frac{1}{2}} \log\log r}{(\pi \log r)^{\frac{1}{2}}} \exp\{-(1+\varepsilon)(1+o(1)) \log r\}. \tag{24}$$

This yields (23) and completes the proof of Theorem 1.

3. Lower Class Result

We again suppose there is given a probability space on which are defined a standard Brownian motion $\{\xi(t), t \geqq 0\}$ and a sequence of non-negative r.v.'s $\{T_i, i>0\}$, and define U_n as in (2). Our assumptions will be stated in (28) below in terms of certain positive values β, ε, and $\gamma>1$. For any choice of such values, let $\{n_r\}$ be a specific non-decreasing sequence satisfying

$$n_r \sim \gamma^r \tag{25}$$

(for example $n_r = \operatorname{int}\{\gamma^r\}$), define q_n by (14) and

$$m_r = n_r - n_{r-1}, \tag{26}$$

and define the events

$$D_r = \{|U_{n_r} - n_r| < q_{n_r}\},$$
$$B_r = \{(1-\varepsilon)\, q_{m_r} < (U_{n_r} - n_r) - (U_{n_{r-1}} - n_{r-1}) < q_{m_r}\}, \tag{27}$$
$$F_{r,\delta} = \{\max_{1 \leqq i \leqq \delta n_r} |U_{n_r+i} - U_{n_r} - i| < \varepsilon\, q_{m_r}\} \qquad \text{for } \delta>0.$$

Theorem 2. *Suppose there is a finite value $\beta>0$ such that for every sufficiently small $\varepsilon>0$ and sufficiently large $\gamma>0$ we have*

(a) $P\{D_r \text{ for a.a.} r\} = 1$,

(b) $P\{B_r \text{ i.o.}\} = 1$, (28)

(c) $P\{F_{r,\delta} \text{ for a.a. } r\} = 1$ *for δ sufficiently small and positive.*

Then, for either choice of sign,

$$\limsup_{n\to\infty} \pm [\xi(n) - \xi(U_n)]/[2\beta\, n (\log n)^2 \log\log n]^{\frac{1}{4}} \geqq 1 \qquad \text{wp 1.} \tag{29}$$

Proof. Let ε be given, such that $0<\varepsilon<\frac{1}{400}$, and ε is small enough that (28) holds for all large γ. In particular,

$$(1-4\varepsilon)^2 (1+\varepsilon)^{-\frac{1}{2}} > \tfrac{16}{17}. \tag{30}$$

Choose γ so large that

$$\gamma - 1 > \varepsilon^{-2} \tag{31}$$

and that (28) holds. Choose δ so small that the equation of (28(c)) holds.

If the event

$$B_r \cap F_{r,\delta} \cap D_{r-1} \tag{32}$$

occurs, and if

$$n_r \leqq n \leqq n_r(1+\delta), \tag{33}$$

we have

$$(1-2\varepsilon)\, q_{m_r} - q_{n_{r-1}} < U_n - n < (1+\varepsilon)\, q_{m_r} + q_{n_{r-1}}. \tag{34}$$

Since $q_{n_{r-1}}/q_{m_r} \sim (\gamma-1)^{-\frac{1}{2}}$ and (31) holds, if r is larger than some constant (34) implies

$$(1-3\varepsilon)\, q_{m_r} < U_n - n < (1+2\varepsilon)\, q_{m_r}. \tag{35}$$

We now define

$$d'_n = [2\beta(1-4\varepsilon)^4\, n(\log n)^2 \log\log n]^{\frac{1}{4}}. \tag{36}$$

Let

$$J_r = \operatorname{int}\{\delta\, n_r / 2\, q_{m_r}\}, \tag{37}$$

and for $0 \leqq i < J_r$ define the numbers

$$\begin{aligned} n'_{r,i} &= n_r + \operatorname{int}\{i(1+5\varepsilon)\, q_{m_r}\}, \\ n''_{r,i} &= n_r + \operatorname{int}\{[i(1+5\varepsilon)+(1-4\varepsilon)]\, q_{m_r}\}, \end{aligned} \tag{38}$$

and the events

$$\begin{aligned} C'_{r,i} &= \{\xi(n''_{r,i}) - \xi(n'_{r,i}) > d'_{n_r}\}, \\ C''_{r,i} &= \{\sup_{0 \leqq x \leqq 8\varepsilon} |\xi(n''_{r,i} + x\, q_{m_r}) - \xi(n''_{r,i})| < 3\varepsilon^{\frac{1}{2}} d'_{n_r}\}, \\ Q'_r &= \bigcup_{0 \leqq i < J_r} C'_{r,i}, \\ Q''_r &= \bigcap_{0 \leqq i < J_r} C''_{r,i}. \end{aligned} \tag{39}$$

Suppose (35) holds for $n = n'_{r,i}$, where $0 \leqq i < J_r$ (so that (33) holds for large r). Then

$$n''_{r,i} < n'_{r,i} + (1-3\varepsilon)\, q_{m_r} < U_{n'_{r,i}} < n'_{r,i} + (1+2\varepsilon)\, q_{m_r} < n''_{r,i} + 8\varepsilon\, q_{m_r}, \tag{40}$$

the extreme inequalities requiring only $r >$ some constant. Obviously, (40) together with $C'_{r,i} \cap C''_{r,i}$ entails

$$\xi(U_{n'_{r,i}}) - \xi(n'_{r,i}) > (1-3\varepsilon^{\frac{1}{2}})\, d'_{n_r}. \tag{41}$$

Since ε is arbitrarily small, we conclude from (41) that Theorem 2 will follow from

$$P\{B_r \cap F_{r,\delta} \cap D_{r-1} \cap Q'_r \cap Q''_r \text{ i.o.}\} = 1. \tag{42}$$

We shall show below that

$$P\{Q'_r \text{ for a.a. } r\} = 1 \tag{43}$$

and

$$P\{Q''_r \text{ for a.a. } r\} = 1. \tag{44}$$

In view of (43), (44), (28a), and (28c), the condition (28b) thus will entail (42) and, thus, our theorem.

It remains to prove (44) and (43), which we do in that order. For the complement of $C''_{r,i}$ the familiar Lévy estimate and (30) yield, for all sufficiently large r, and $0 \leqq i < J_r$,

$$\begin{aligned}\log P\{\bar{C}''_{r,i}\} &\leqq -(3\varepsilon^{\frac{1}{2}} d'_{n_r})^2/2(8\varepsilon q_{m_r}) < -(3d'_{n_r})^2/16 q_{n_r} \\ &= -(\tfrac{9}{16})(1-4\varepsilon)^2(1+\varepsilon)^{-\frac{1}{2}} \log n_r < -(\tfrac{9}{17}) \log n_r.\end{aligned} \tag{45}$$

Hence, for the complement of Q''_r we have from (37), (45), and (14), for all large r,

$$\begin{aligned}P\{\bar{Q}''_r\} &\leqq \sum_{i=0}^{J_r-1} P\{\bar{C}''_{r,i}\} < (\delta n_r/2 q_{m_r}) n_r^{-\frac{9}{17}} \\ &\sim (1-\gamma^{-1})^{-\frac{1}{2}} \delta n_r^{-\frac{1}{34}} [8\beta(1+\varepsilon) \log\log n_r]^{-\frac{1}{2}}.\end{aligned} \tag{46}$$

The Borel-Cantelli lemma and (46) yield $P\{\bar{Q}''_r \text{ i.o.}\}=0$, which is equivalent to (44).

Next, since $1-5\varepsilon < [n''_{r,i} - n'_{r,i}]/q_{m_r} < 1-3\varepsilon$ for all sufficiently large r and for $0 \leqq i < J_r$, the standard Gaussian tail estimate then gives

$$P\{C'_{r,i}\} > \frac{[(1-6\varepsilon) q_{m_r}/2\pi]^{\frac{1}{2}}}{d'_{n_r}} \exp\{-(d'_{n_r})^2/2(1-3\varepsilon) q_{m_r}\}. \tag{47}$$

Hence, for all large r, and for $0 \leqq i < J_r$,

$$\begin{aligned}\log P\{C'_{r,i}\} &> -(d'_{n_r})^2/2 q_{m_r}(1-4\varepsilon) \\ &= -\tfrac{1}{2}(1-4\varepsilon)(1+\varepsilon)^{-\frac{1}{2}} \left\{\frac{n_r \log\log n_r}{m_r \log\log m_r}\right\}^{\frac{1}{2}} \log n_r.\end{aligned} \tag{48}$$

Since $n_r/m_r \sim \gamma/(\gamma-1) < [1+\varepsilon^2]$ by (31), and since $\varepsilon < \frac{1}{400}$, we obtain from (48) for all large r and for $0 \leqq i < J_r$,

$$\begin{aligned}\log P\{C'_{r,i}\} &> -\tfrac{1}{2}(1-4\varepsilon)[(1+\varepsilon^2)/(1+\varepsilon)]^{\frac{1}{2}} \log n_r \\ &> -\tfrac{1}{2}(1-\varepsilon) \log n_r.\end{aligned} \tag{49}$$

Since for each sufficiently large r the $C'_{r,i}$ $(0 \leqq i < J_r)$ are independent, we then obtain for the complement of Q'_r,

$$\begin{aligned}\log P\{\bar{Q}'_r\} &= \sum_{i=0}^{J_r-1} \log P\{\bar{C}'_{r,i}\} < (\delta n_r/2 q_{m_r}) \log[1-n_r^{-(1-\varepsilon)/2}] \\ &< -(\delta n_r/2 q_{m_r}) n_r^{-(1-\varepsilon)/2} \\ &\sim -\delta[8(1+\varepsilon)\beta(\gamma-1)\gamma^{-1} \log\log n_r]^{-\frac{1}{2}} n_r^{\varepsilon/2}.\end{aligned} \tag{50}$$

The Borel-Cantelli lemma yields $P\{\bar{Q}'_r \text{ i.o.}\}=0$, which is equivalent to (43). This completes the proof of Theorem 2.

Remarks on Theorem 2.

1. The conclusion of Theorem 2 obviously remains valid under such changes as replacing U_n by $-U_n$ in the definition of B_r, and replacing (28c) by $\lim_{\delta \downarrow 0} P = 1$ (which is pretty artificial in view of $0-1$ laws for $\{U_n\}$ which will usually hold in applications).

2. At first sight it may appear strange that no analogue of the last part of (3a) is assumed in Theorem 2. This is misleading, since the strong assumption (28c) allows us to avoid computing conditional probabilities of certain events, given B_r, which might require an analogue of (3a). Moreover, the verification of (43) by use of (50) makes strong use of the independent increments of ξ, and the use of (43)–(44) in our proof requires (28b) to hold; the verification of the latter in the application of Theorem 2 to Theorem 3 (Section 4) makes use of the structure of the $\{U_n\}$ process, and this can be expected in other applications. Thus, a weakening of (28), especially of (28c), can be given if one assumes conditions including analogues of the last part of (3a), but so far we have not obtained natural and useful conditions of this form.

3. A crucial aspect of the proof is that although Q'_r is concerned only with a small fraction $\delta/(\gamma-1)$ of the interval $[n_r, n_{r+1}]$, the number J_r is still large enough to yield (43). In generalizations and extensions one may want to replace (43)–(44) by an analogue of the weaker statement that $\bigcup_i (C'_{r,i} \cap C''_{r,i})$ occurs for a.a. r wp 1.

4. Proof of Theorem 3

We must verify the three parts of (28) in the i.i.d. case under the assumption (1) with $\beta > 0$.

Condition (28a) is a consequence of the Hartman-Wintner LIL [10].

To prove (28c), suppose $\delta < \gamma - 1$. Write $n_r^* = \operatorname{int}\{n_r(1+\delta)\}$. There is a value ρ such that $n_r + 1 < n_r^* < n_{r+1}$ for $r \geqq \rho$. The r.v.'s T_i, $n_r < i \leqq n_r^*$, $r \geqq \rho$, are then distinct, and we relabel the sequence

$$T_{n_\rho+1}, T_{n_\rho+2}, \ldots, T_{n_\rho^*}, T_{n_{\rho+1}+1}, \ldots, T_{n_{\rho+1}^*}, T_{n_{\rho+2}+1}, \ldots$$

as

$$1+V_1, 1+V_2, 1+V_3, \ldots.$$

The Hartman-Wintner LIL for the V_i says that

$$\left|\sum_1^m V_i\right| < q_m \tag{51}$$

for a.a. m, wp 1. If (51) holds for two values $m = M$ and $m = k > M$, we have

$$\left|\sum_{M+1}^k V_i\right| \leqq \left|\sum_1^k V_i\right| + \left|\sum_1^M V_i\right| < q_k + q_M. \tag{52}$$

If f is any function from N into N with $f(n) > n$ for all n, we conclude that, for a.a. M in N, wp 1,

$$\max_{M < k \leqq f(M)} \left|\sum_{M+1}^k V_i\right| < q_M + q_{f(M)} < 2\, q_{f(M)}. \tag{53}$$

In particular, selecting only the values $M=n_r$, $r\geqq\rho$, with f chosen so that $f(n_r)=n_r^*$, and writing

$$k_r=\sum_{j=\rho}^{r}(n_r^*-n_r) \tag{54}$$

for $r\geqq\rho$, we obtain

$$\max_{1\leqq i\leqq\delta n_r}|U_{n_r+i}-U_{n_r}-i|<2q_{k_r} \qquad \text{for a.a. } r\geqq\rho, \text{ wp } 1. \tag{55}$$

Since $k_r\sim n_r\,\delta\gamma/(\gamma-1)\sim m_r\,\delta\gamma^2/(\gamma-1)^2$, we have

$$2q_{k_r}\sim[2\delta^{\frac{1}{2}}\,\gamma/\varepsilon(\gamma-1)]\,\varepsilon\,q_{m_r}. \tag{56}$$

Hence, if $\delta<\varepsilon^2(\gamma-1)^2/4\gamma^2$, we have $2q_{k_r}<\varepsilon\,q_{m_r}$ for all large r, and then (28c) follows from (55) and the definition (27).

We turn finally to (28b). While this condition can be verified in several ways, an expeditious proof relies on the Hartman-Wintner truncation scheme [10] and the validity of the analogue of (28b) in Kolmogorov's lower class proof [16], as we shall now show.

Firstly, given $0<\varepsilon<1$, let $\varepsilon'>0$ be such that

$$1-2\varepsilon'>(1-\varepsilon)(1+\varepsilon)^{\frac{1}{2}} \quad\text{and}\quad 1+3\varepsilon'<(1+\varepsilon)^{\frac{1}{2}}. \tag{57}$$

If T_i is truncated and centered (by constants depending only on i), say to Z_i, exactly as in [10], so that $|Z_n|=o(n/q_n)$, then it is shown in [10] that

$$\sum_{1}^{n}(T_i-Z_i)=o(q_n) \qquad \text{wp } 1. \tag{58}$$

Moreover, the developments of Kolmogorov's lower class proof [16] (also to be found in Loève [17], pp. 260–262) show, after minor arithmetic to verify the negligible difference (in the appropriate sense) between the first two moments of $\sum_{1}^{n}Z_i$ and U_n, that if

$$H_r=\sum_{i=n_{r-1}+1}^{n_r}Z_i-m_r, \tag{59}$$

then for each small positive ε' there is a small positive ε'' such that, for all large r,

$$P\{H_r>(1-\varepsilon')[2\beta\,m_r\log\log m_r]^{\frac{1}{2}}\}>r^{\varepsilon''-1}. \tag{60}$$

On the other hand, the analogue of the first inequality of (53) for the Z_i (and their LIL) rather than the V_i yields

$$|H_r|<(1+\varepsilon')(2\beta)^{\frac{1}{2}}\{(n_r\log\log n_r)^{\frac{1}{2}}+(n_{r-1}\log\log n_{r-1})^{\frac{1}{2}}\} \quad \text{for a.a. } r, \text{ wp } 1. \tag{61}$$

If γ is so large that $(1+\varepsilon')(\gamma^{\frac{1}{2}}+1)/(\gamma-1)^{\frac{1}{2}}<(1+2\varepsilon')$, the Borel-Cantelli lemma for the independent H_r of (61), together with (60), yields

$$\sum_r P\{1-\varepsilon'<H_r[2\beta\,m_r\log\log m_r]^{-\frac{1}{2}}<1+2\varepsilon'\}=+\infty. \tag{62}$$

By (58) and the analogue of (53) for the $T_i - Z_i$ instead of the V_i,

$$\left|\sum_{i=n_{r-1}+1}^{n_r} (T_i - Z_i)\right| < \varepsilon' [2\beta m_r \log\log m_r]^{\frac{1}{2}} \qquad \text{for a.a. } r, \text{ wp } 1. \tag{63}$$

The sequence (in r) of sums on the left side of (63) being independent, the Borel-Cantelli lemma and (63), with (62), yield

$$\sum_r P\{1-2\varepsilon' < [2\beta m_r \log\log m_r]^{-\frac{1}{2}} (U_{n_r} - U_{n_{r-1}} - m_r) < 1 + 3\varepsilon'\} = +\infty. \tag{64}$$

From (57), the definition of (27), and the Borel-Cantelli lemma applied to the independent events of (64), we obtain (28 b). This completes the proof of Theorem 3.

References

1. Bahadur, R.: A note on quantiles in large samples. Ann. math. Statistics **37**, 577–580 (1966).
2. Breiman, L.: On the tail behavior of sums of independent random variables. Z. Wahrscheinlichkeitstheorie verw. Geb. **9**, 20–25 (1967).
3. – Probability. Reading, Mass.: Addison-Wesley 1968.
4. Brillinger, D. R.: An asymptotic representation of the sample d.f. Bull. Amer. math. Soc. **75**, 545–547 (1969).
5. Dubins, L.: On a theorem of Skorokhod. Ann. math. Statistics **39**, 2094–2097 (1968).
6. Feller, W.: The general form of the so-called law of the iterated logarithm. Trans. Amer. math. Soc. **54**, 373–402 (1953).
7. – An introduction to probability theory and its applications, vol. II. New York: John Wiley 1966.
8. – The asymptotic distribution of the range of sums of independent random variables. Ann. math. Statistics **22**, 427–432 (1951).
9. Hall, W. J.: On the Skorokhod embedding theorem. (To be published.)
10. Hartman, P., Wintner, A.: On the law of the iterated logarithm. Amer. J. Math. **63**, 169–176 (1941).
11. Ito, K., McKean, H. P., Jr.: Diffusion processes and their sample paths. New York: Academic Press 1965.
12. Kiefer, J.: On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. Pacific J. Math. **11**, 649–660 (1961).
13. – On Bahadur's representation of sample quantiles. Ann. math. Statistics **38**, 1323–1342 (1967).
14. – On the range of a general random walk. (To be published.)
15. – Old and new methods for studying order statistics and sample quantiles. Proc. 1st Intl. Conference on Nonparametric Inference.
16. Kolmogorov, A.: Das Gesetz des iterierten Logarithmus. Math. Ann. **101**, 126–135 (1929).
17. Loève, M.: Probability theory, 2nd edn. New York: Van Nostrand 1960.
18. Skorokhod, A.: Studies in the theory of Random processes. Reading, Mass.: Addison-Wesley 1965.
19. Strassen, V.: An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie verw. Geb. **3**, 211–226 (1964).
20. – A converse to the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie verw. Geb. **4**, 265–268 (1966).
21. – Almost sure behavior of sums of independent random variables and martingales. Proc. Fifth Berkeley Sympos. math. Statist. Probab. Vol. II (Part I), 315–343 (1967).

Professor Dr. J. Kiefer
Department of Mathematics
Cornell University
Ithaca, N.Y. 14850, USA

(Received August 23, 1968)