[8] W. LeVeque, *The distribution* mod 1 *of trigonometric sequences*, Duke Math. J. 20 (1953), pp. 367–374.
[9] H. Weyl, *Über die Gleichverteilung die Zahlen* mod *Eins*, Math. Annalen 77 (1916), pp. 313–352.

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
Jerusalem, Israel

Received on 27.6.1986 (1659)

ACTA ARITHMETICA
XLIX (1988)

Large deviations of sums of independent random variables

by

HUGH L. MONTGOMERY* (Ann Arbor, Mich.) and
ANDREW M. ODLYZKO (Murray Hill, N.J.)

Dedicated to Pál Erdős on the occasion of his 75-th birthday

1. Statement of results. Our object is to estimate the probability that a sum of independent random variables is large. In this direction we derive a rather precise upper bound, and a corresponding lower bound.

THEOREM 1. *Let $X_1, X_2, \ldots$ be independent random variables such that $P(X_n = 1) = 1/2$, $P(X_n = -1) = 1/2$. Let $\{r_n\}$ be a non-increasing sequence of non-negative real numbers for which*

$$\sigma^2 = \sum_{n=1}^{\infty} r_n^2 < \infty, \tag{1}$$

and put $X = \sum_{n=1}^{\infty} r_n X_n$. If N and V are chosen so that $\sum_{n \leqslant N} r_n \leqslant V/2$, then

$$P(X \geqslant V) \leqslant \exp\Big(-\tfrac{1}{8} V^2 \Big(\sum_{n>N} r_n^2\Big)^{-1}\Big). \tag{2}$$

If $\sum_{n \leqslant N} r_n \geqslant 2V$ then

$$P(X \geqslant V) \geqslant 2^{-22} \exp\Big(-120 V^2 \Big(\sum_{n>N} r_n^2\Big)^{-1}\Big). \tag{3}$$

Also, if $\sum_{n \leqslant N} r_n \geqslant V$ then

$$P(X \geqslant V) \geqslant 2^{-N-1}. \tag{4}$$

---

* Research supported in part by National Science Foundation Grant NSF DMS 85-02804.

The upper bound (2) was proved earlier by Saltzberg [8]. We include the easy proof, since it is short and the inequalities used in this proof are used in the proof of (3) as well.

By Kolmogorov's three series theorem we see that condition (1) ensures that the series defining X converges a.e. It is well known (see Petrov [6], p. 58) that if $V \geqslant 0$ then

(5) $$P(X \geqslant V) \leqslant \exp\left(-V^2/(2\sigma^2)\right).$$

This is sharp for $V \approx \sigma$ if σ is large compared to $\max_n r_n$, but otherwise $P(X \geqslant V)$ is significantly smaller. Since $\sum_{n>N} r_n^2 \to 0$ as $N \to \infty$, we see from (2) that $P(X \geqslant V)$ tends to 0 more quickly as $V \to \infty$ than it would if X were normally distributed. As the X_n are symmetrically and identically distributed, the requirement that the r_n be positive and nonincreasing does not occasion any loss of generality. From Theorem 1 we see that

$$\exp(-c_1 V^4) \leqslant P\left(\sum_n X_n/n^{3/4} \geqslant V\right) \leqslant \exp(-c_2 V^4)$$

for $V \geqslant 1$. In this situation the lower bounds (3) and (4) are comparable, but if $\sum r_n^2$ converges slowly, e.g. $r_n = n^{-1/2}(\log n)^{-1}$, then (3) is superior to (4). On the other hand, if $\sum r_n$ diverges slowly, then (3) is inferior to (4) and (2) can be refined by taking more care in the choice of parameters. For example, the method we use to derive (2) can be used to show that

$$P\left(\sum_n X_n/n \geqslant V\right) \leqslant \exp(c_3 e^V)$$

for $V \geqslant 0$, while (4) gives

$$P\left(\sum_n X_n/n \geqslant V\right) \geqslant \exp(c_4 e^V)$$

for $V \geqslant 0$. S.O.Rice [7] has determined these probabilities for small V. In general, when the r_n decrease in a regular way, an asymptotic expansion of $P(X \geqslant V)$ as $V \to \infty$ can be determined by the saddle point method.

R.Monach [5] has calculated similar probabilities in connection with the distribution of the error term in the prime number theorem and with the distribution of $\arg L(1, \chi)$ for Dirichlet characters $\chi \pmod q$. Chowla and Erdős [1] proved that if $s > 3/4$ is fixed then the numbers

$$L_d(s) = \sum_{n=1}^{\infty} \left(\frac{d}{n}\right) n^{-s}$$

have a limiting distribution in the sense that there is a function $F(s, V)$ such that

$$\lim_{X \to \gamma} \frac{2}{X} \operatorname{card}\{d\colon 0 < d \leqslant X,\ d \equiv 0 \text{ or } 1 \pmod 4,\ d \neq s^2,\ L_d(s) \leqslant V\} = F(s, V).$$

Elliott [2] extended the range of validity to $s > 1/2$, and established corresponding results for complex s for which $\operatorname{Re} s > 1/2$. From his analysis and Theorem 1 it can be shown that there are constants $a_i = a_i(s)$ such that if $1/2 < s < 1$ then

(6) $$\exp\left(-a_1 (\log V)^{1/(1-s)} (\log\log V)^{s/(1-s)}\right) < 1 - F(s, V) < \exp\left(-a_2 \left((\log V)^{1/(1-s)} (\log\log V)^{s/(1-s)}\right)\right)$$

for $V \geqslant 4$. Moreover, these inequalities remain valid with $1-F(s, V)$ replaced by $F(s, 1/V)$.

Concerning other independent random variables, we note that (5) depends only on the fact that $E(e^{\lambda X_n}) \leqslant \exp(\lambda^2/2)$. The sharper bound (2) requires additionally that $|X_n| \leqslant 1$. Our method yields the following more general result, which we state without proof.

THEOREM 2. *For $n = 1, 2, \ldots$ let Y_n be independent real valued random variables such that $E(Y_n) = 0$ and $|Y_n| \leqslant 1$. Suppose there is a constant $c > 0$ such that $E(Y_n^2) \geqslant c$ for all n. Put $Y = \sum r_n Y_n$ where $\sum r_n^2 < \infty$. If $\sum_{|r_n| \geqslant \alpha} |r_n| \leqslant V/2$, then*

$$P(Y \geqslant V) \leqslant \exp\left(-\tfrac{1}{16} V^2 \Big(\sum_{|r_n| < \alpha} r_n^2\Big)^{-1}\right).$$

If $\sum_{|r_n| \geqslant \alpha} |r_n| \geqslant 2V$, then

$$P(Y \geqslant V) \geqslant a_1 \exp\left(-a_2 V^2 \Big(\sum_{|r_n| < \alpha} r_n^2\Big)^{-1}\right).$$

Here $a_1 > 0$ and $a_2 > 0$ depend only on c.

D. Joyner [3] has used Theorem 2 to estimate the asymptotic distribution of $|\zeta(\sigma + it)|$ for given σ, $1/2 < \sigma < 1$. His result is similar to the estimate (6).

We derive our main results from simple inequalities for the characteristic functions of the random variables involved. In § 4 we indicate how these results can instead be derived from inequalities for moments. This latter approach is convenient when investigating complex valued random variables, or random variables which are only approximately independent.

2. Basic lemmas. We employ the following elementary inequalities (see Kahane [4], p. 6).

LEMMA 1: *If $Z \in L^1(\Omega)$ and $Z \geqslant 0$, then*

$$P(Z \geqslant aE(Z)) \leqslant 1/a$$

for all $a > 0$. If $Z \in L^2(\Omega)$ and $0 < a < 1$, then

$$P(Z \geqslant aE(Z)) \geqslant (1-a)^2 E(Z)^2/E(Z^2).$$

Clearly $E(e^{\lambda X_n}) = \cosh\lambda$. Bounds for this quantity are provided by

LEMMA 2. *For any* $u \geqslant 0$,

(7) $$\cosh u \leqslant e^u$$

and

(8) $$\cosh u \leqslant e^{u^2/2}.$$

Moreover,

(9) $$\cosh u > e^{u^2/6}$$

for $0 < u \leqslant 3$, *and*

(10) $$\cosh u > 2e^{u/2}$$

for $u \geqslant 3$.

Proof. Clearly $\cosh u = \frac{1}{2}e^u + \frac{1}{2}e^{-u} \leqslant e^u$ for $u \geqslant 0$, and

$$\cosh u = \sum_{n=0}^{\infty} \frac{u^{2n}}{(2n)!} \leqslant \sum_{n=0}^{\infty} \frac{(u^2/2)^n}{n!} = e^{u^2/2},$$

since $n!\,2^n \leqslant (2n)!$. If $0 < v \leqslant 9$ then $1+v/2 > e^{v/6}$, and hence

$$\cosh u = \sum_{n=0}^{\infty} \frac{u^{2n}}{(2n)!} > 1+u^2/2 > e^{u^2/6}$$

for $0 < u \leqslant 3$. Finally, $\cosh u = \frac{1}{2}e^u + \frac{1}{2}e^{-u} > \frac{1}{2}e^u$, so that

$$\cosh u > \tfrac{1}{2}e^u > 2e^{u/2}$$

for $u \geqslant 3$.

3. Proof of Theorem 1. By the first part of Lemma 1 with $Z = e^{\lambda X}$, $a = e^{\lambda V}/E(e^{\lambda X})$ we see that

(11) $$P(X \geqslant \lambda) \leqslant e^{-\lambda V} E(e^{\lambda X}).$$

Since the X_n are independent,

(12) $$E(e^{\lambda X}) = \prod_{n=1}^{\infty} E(e^{\lambda r_n X_n}) = \prod_{n=1}^{\infty} \cosh(\lambda r_n).$$

Let N be arbitrary. We use (7) for $1 \leqslant n \leqslant N$, and (8) for $n > N$, to see that

(13) $$E(e^{\lambda X}) \leqslant \exp\Big(\lambda \sum_{n \leqslant N} r_n + \tfrac{1}{2}\lambda^2 \sum_{n>N} r_n^2\Big).$$

Hence if $\sum_{n \leqslant N} r_n \leqslant V/2$ then

$$P(X \geqslant V) \leqslant \exp\Big(-\tfrac{1}{2}\lambda V + \tfrac{1}{2}\lambda^2 \sum_{n>N} r_n^2\Big),$$

and (2) follows on taking $\lambda = \frac{1}{2}V\big(\sum_{n>N} r_n^2\big)^{-1}$.

Alternatively, we can derive (2) from (5) by writing

$$X = \sum_{n \leqslant N} r_n X_n + \sum_{n>N} r_n X_n = X' + X'',$$

say. Then $|X'| \leqslant V/2$, and we obtain (2) by using (5) to estimate $P(X'' \geqslant V/2)$.

To derive (3) we use the second part of Lemma 1 with $Z = e^{\lambda X}$ and $a = 1/2$; we choose λ so that

(14) $$\tfrac{1}{2}E(e^{\lambda X}) = e^{\lambda V}.$$

Such a λ must exist, since both sides above are continuous functions of λ, the left-hand side is smaller than the right when $\lambda = 0$, and the reverse is true when $\lambda \geqslant 3/r_N$, since by (12) and (10),

$$\tfrac{1}{2}E(e^{\lambda X}) = \tfrac{1}{2}\prod_{n=1}^{\infty} \cosh(\lambda r_n) > \tfrac{1}{2}\prod_{n \leqslant N} \cosh(\lambda r_n)$$
$$> \exp\Big(\tfrac{1}{2}\lambda \sum_{n \leqslant N} r_n\Big) \geqslant \exp(\lambda V).$$

By Lemma 1 and (14) we see that

$$P(X \geqslant V) = P(e^{\lambda X} \geqslant e^{\lambda V}) \geqslant \tfrac{1}{4}E(e^{\lambda X})^2/E(e^{2\lambda X}),$$

and by (14) again this is

$$= e^{2\lambda V}/E(e^{2\lambda X}).$$

By applying (13) to estimate $E(e^{2\lambda X})$, we obtain

(15) $$P(X \geqslant V)\, e^{-2\lambda V} \geqslant \exp\Big(-2\lambda \sum_{n \leqslant M} r_n - 2\lambda^2 \sum_{n>M} r_n^2\Big)$$

for any M. We take M so that $\lambda r_n \geqslant 3$ for $n \leqslant M$, and $\lambda r_n < 3$ for $n > M$. By (9), (10), (12), and (14) we see that

(16) $$2e^{\lambda V} = E(e^{\lambda X}) > \prod_{n \leqslant M} (2e^{\lambda r_n/2}) \prod_{n>M} e^{\lambda^2 r_n^2/6} > \exp\Big(\tfrac{1}{2}\lambda \sum_{n \leqslant M} r_n + \tfrac{1}{6}\lambda^2 \sum_{n>M} r_n^2\Big).$$

We raise both sides of this inequality to the 12-th power, and multiply the two sides of (15) by these quantities to see that

$$P(X \geqslant V)\, 2^{12} e^{10\lambda V} \geqslant 1.$$

If $\lambda V \leqslant \log 2$ then we have $P(X \geqslant V) \geqslant 2^{-22}$, and we are done. Thus we may suppose that

$$\log 2 \leqslant \lambda V. \tag{17}$$

To complete our proof of (3) it suffices to show that

$$\lambda \leqslant 12V\Big(\sum_{n>N} r_n^2\Big)^{-1}. \tag{18}$$

To this end we first show that $M \leqslant N$. If it were the case that $M > N$ then by (16) we would have

$$2e^{\lambda V} > \prod_{n \leqslant M} (2e^{\lambda r_n/2}) \geqslant 2\exp\big(\tfrac{1}{2}\lambda \sum_{n \leqslant N} r_n\big),$$

which contradicts our hypothesis that $\sum_{n \leqslant N} r_n \geqslant 2V$. Hence $M \leqslant N$, so that by (16),

$$2e^{\lambda V} > \prod_{n>M} e^{\lambda^2 r_n^2/6} \geqslant \exp\big(\tfrac{1}{6}\lambda^2 \sum_{n>N} r_n^2\big),$$

which gives

$$\tfrac{1}{6}\lambda^2 \sum_{n>N} r_n^2 \leqslant \lambda V + \log 2.$$

By (17) this is $\leqslant 2\lambda V$, so we have (18), and the proof of (3) is complete.

To derive (4) we write

$$X = \sum_{n \leqslant N} r_n X_n + X'.$$

The variable X' is independent of the X_n for $n \leqslant N$. Also, $X \geqslant V$ if $X_n = 1$ for $n \leqslant N$ and $X' \geqslant 0$. Hence

$$P(X \geqslant V) \geqslant P(X' \geqslant 0) \prod_{n \leqslant N} P(X_n = 1) = 2^{-N-1}.$$

4. Moment inequalities. To illustrate the use of moments in the present context we now establish the following further result. It is convenient here to suppose that our basic independent random variables are uniformly distributed on the unit circle $|z| = 1$.

LEMMA 3. *Let $Z_1, \ldots, Z_n$ be independent random variables, each one uniformly distributed on the unit circle $|z| = 1$ in the complex plane, suppose that $r_1 \geqslant r_2 \geqslant \ldots \geqslant r_N \geqslant 0$, and put $Z = \sum_{n=1}^{N} r_n Z_n$. Then*

$$E(|Z|^{2k}) \leqslant k!\, \sigma^{2k} \tag{19}$$

for every non-negative integer k, where $\sigma^2 = E(|Z|^2) = \sum_{n=1}^{N} r_n^2$. Moreover, if

$$\sum_{n<k} r_n^2 \leqslant \tfrac{1}{2}\sigma^2 \tag{20}$$

then

$$E(|Z|^{2k}) \geqslant 2^{-k} k!\, \sigma^{2k}. \tag{21}$$

Proof. Since $E(Z_n^m) = 0$ for all integers $m \neq 0$ we see that

$$E(|Z|^{2k}) = \sum \binom{k}{a_1\, a_2 \ldots a_N}^2 r_1^{2a_1} r_2^{2a_2} \ldots r_N^{2a_N}$$

where the sum is over all N-tuples of non-negative integers a_n for which $a_1 + a_2 + \ldots + a_N = k$. Since the multinomial coefficient never exceeds $k!$, the above is

$$\leqslant k! \sum \binom{k}{a_1\, a_2 \ldots a_N} r_1^{2a_1} r_2^{2a_2} \ldots r_N^{2a_N} = k!\, \sigma^{2k}.$$

To obtain the lower bound (21) we restrict our attention to those N-tuples for which $a_n = 0$ or 1 for all n. In this case the multinomial coefficient is exactly $k!$, so that

$$\begin{aligned} E(|Z|^{2k}) &\geqslant k!^2 \sum_{n_1 < n_2 < \ldots < n_k} r_{n_1}^2 r_{n_2}^2 \ldots r_{n_k}^2 \\ &= k! \sum_{n_i \text{ distinct}} r_{n_1}^2 r_{n_2}^2 \ldots r_{n_k}^2 \\ &= k! \sum_{n_1=1}^{N} r_{n_1}^2 \sum_{\substack{n_2=1 \\ n_2 \neq n_1}}^{N} r_{n_2}^2 \ldots \sum_{\substack{n_k=1 \\ n_k \neq n_j (j<k)}}^{N} r_{n_k}^2. \end{aligned} \tag{22}$$

Since the r_n are nonincreasing we see that

$$\sum_{\substack{n_i=1 \\ n_i \neq n_j (j<i)}}^{N} r_{n_i}^2 \geqslant \sum_{n=i}^{N} r_n^2.$$

By hypothesis (20) this last sum is $\geqslant \frac{1}{2}\sigma^2$, and consequently the expression (22) is $\geqslant k!(\frac{1}{2}\sigma^2)^k$.

An upper bound for $P(|Z| \geqslant V)$ can be derived from (19) by using the inequality $P(|Z| \geqslant V)V^{2k} \leqslant E(|Z|^{2k})$ which is a special case of the first inequality in Lemma 1. A lower bound can be derived by appealing to the fact that if $c > 1$ then

$$P(|Z| \geqslant V)(c-1)^2 V^{4k}/4 \leqslant E\big((|Z|^{2k} - V^{2k})(cV^{2k} - |Z|^{2k})\big).$$

References

[1] S. Chowla and P. Erdős, *A theorem on the distribution of the values of L-functions*, J. Indian Math. Soc. (N.S.) 15 (1951), pp. 11–18.

[2] P. D. T. A. Elliott, *On the distribution of the values of quadratic L-series in the half-plane* $\sigma > 1/2$, Invent. math. 21 (1973), pp. 319–338.

[3] David Joyner, *Distribution theorems of L-functions*, Longman Scientific, Harlow 1986.

[4] J.-P. Kahane, *Some random series of functions*, Heath, Lexington 1968.

[5] William Reynolds Monach, *Numerical investigation of several problems in number theory*, Univ. of Michigan Ph.D. Dissertation, Ann Arbor 1980.

[6] V. V. Petrov, *Sums of independent random variables*, Springer-Verlag, Berlin 1975.

[7] S. O. Rice, *Distribution of* $\sum a_n/n$, a_n *randomly equal to* ± 1, Bell System Technical J. 52 (1973), pp. 1097–1103.

[8] B. S. Saltzberg, *Intersymbol interference error bounds with applications to ideal bandlimited signaling*, IEEE Trans. Information Theory IT-14 (1968), pp. 563–568.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
Ann Arbor, MI 48109
AT&T BELL LABORATORIES
Murray Hill, NJ 07974

Received on 23.7.1986
and in revised form on 25.10.1986 (1663)