CHANGES OF SIGN OF SUMS OF RANDOM VARIABLES

P. ERDŐS AND G. A. HUNT

1. Introduction. Let $x_1, x_2, \cdots$ be independent random variables all having the same continuous symmetric distribution, and let

$$s_k = x_1 + \cdots + x_k.$$

Our purpose is to prove statements concerning the changes of sign in the sequence of partial sums $s_1, s_2, \cdots$ which do not depend on the particular distribution the x_k may have.

The first theorem estimates the expectation of N_n, the number of changes of sign in the finite sequence $s_1, \cdots, s_{n+1}$. Here and later we write $\phi(k)$ for

$$\frac{2([k/2]+1)}{k+1} \binom{k}{[k/2]} 2^{-k} \approx (2\pi k)^{-1/2}.$$

THEOREM 1.

$$\sum_{k=1}^{n} \frac{1}{2(k+1)} \leq \mathrm{E}\{N_n\} \leq \frac{1}{2} \sum_{k=1}^{n} \phi(k).$$

It is known (see [1]) that, with probability one,

$$\limsup_{n \to \infty} \frac{N_n}{(n \log \log n)^{1/2}} = 1 \tag{1}$$

when the x_k are the Rademacher functions. We conjecture, but have not been able to prove, that (1) remains true, provided the equality sign be changed to $\leq$, for all sequences of identically distributed independent symmetric random variables. We have had more success with lower limits:

THEOREM 2. *With probability one,*

Received November 13, 1952. The preparation of this paper was sponsored (in part) by the Office of Naval Research, USN.

Pacific J. Math. 3 (1953), 673-687

$$\liminf_{n\to\infty} \frac{N_n}{\log n} \geq \frac{1}{2}.$$

By considering certain subsequences of the partial sums we obtain an exact limit theorem which is still independent of the distribution of the x_k: Let α be a positive number and a the first integer such that $(1+\alpha)^a \geq 2$; let $1', 2', \cdots$ be any sequence of natural numbers satisfying $(k+1)' \geq (1+\alpha)k'$; and let N_n' be the number of changes of sign in the sequence $s_1', \cdots, s_{n+1}'$, where s_k' stands for $s_{k'}$.

THEOREM 3. $\mathrm{E}\{N_n'\} \geq [n/a]/8$, *and, with probability one,*

$$\lim_{n\to\infty} \frac{N_n'}{\mathrm{E}\{N_n'\}} = 1.$$

For $k' = 2^k$, it is easy to see that $\mathrm{E}\{N_n'\} = n/4$; so with probability one the number of changes of sign in the first n terms of the sequence $s_1, s_2, \cdots, s_{2^k}, \cdots$ is asymptotic to $n/4$.

The basis of our proofs is the combinational Lemma 2 of the next section. When translated into the language of probability, this gives an immediate proof of Theorem 1. We prove Theorem 3 in §3 and then use it to prove Theorem 2. A sequence of random variables for which $N_n/\log n \longrightarrow 1/2$ is exhibited in §4; thus the statement of Theorem 2 is in a way the best possible. Finally we sketch the proof of the following theorem, which was discovered by Paul Lévy [2] when the x_k are the Rademacher functions.

THEOREM 4. *With probability one,*

$$\sum_{k=1}^{n} \frac{\operatorname{sgn} s_k}{k} = o(\log n).$$

Our results are stated only for random variables with continuous distributions. Lemma 3, slightly altered to take into account cases of equality, remains true however for discontinuous distributions; the altered version is strong enough to prove the last three theorems as they stand and the first theorem with the extreme members slightly changed. The symmetry of the x_k is of course essential in all our arguments.

2. Combinatorial lemmas. Let $a_1, \cdots, a_n$ be positive numbers which are free in the sense that no two of the sums $\pm a_1 \pm \cdots \pm a_n$ have the same value.

These sums, arranged in decreasing order, we denote by $S_1, \cdots, S_{2^n}$; q_i is the excess of plus signs over minus signs in S_i; and $Q_i = q_1 + \cdots + q_i$. It is clear that $Q_{2^n} = 0$ and that $Q_i = Q_{2^n-i}$ for $1 \leq i < 2^n$.

LEMMA 1. *For* $1 \leq i \leq 2^{n-1}$,

$$0 \leq Q_i - i \leq ([n/2]+1)\binom{n}{[n/2]} - 2^{n-1}.$$

The proof of the first inequality, which is evident for $n = 1$, goes by induction. Suppose $n > 1$ and $i \leq 2^{n-1}$. Define S_j' and Q_j' for $1 \leq j \leq 2^{n-1}$ just as S_j and Q_j were defined above, but using only $a_1, \cdots, a_{n-1}$. Let k and l be the greatest integers such $S_k' - a_n \geq S_i$ and $S_l' + a_n \geq S_i$. It may happen that no such k exists; then $i = l$ and the proof is relatively easy. Otherwise $k \leq l$, $k \leq 2^{n-2}$, and $i = k + l$. If $l \leq 2^{n-2}$ then

$$Q_i = Q_k' - k + Q_l' + l = (Q_k' - k) + (Q_l' - l) + 2l \geq i.$$

If $2^{n-2} < l < 2^{n-1}$ then

$$Q_i = Q_k' - k + Q_l' + l = Q_k' - k + Q'_{2^{n-1}-l} + l$$

$$= (Q_k' - k) + (Q_{2^{n-1}-l} - 2^{n-1} + l) + 2^{n-1} - l + l \geq 2^{n-1} \geq i.$$

Finally, if $l = 2^{n-1}$ then, recalling $Q'_{2^{n-1}} = 0$, we get

$$Q_i = Q_k' - k + Q'_{2^{n-1}} + 2^{n-1} \geq 2^{n-1} \geq i.$$

In order to prove the second inequality we note that for each i the maximum of Q_i is attained if the a_i are given such values that $S_j > S_k$ implies $q_j \geq q_k$, —this happens if the a_j are nearly equal. Assume this situation. Then if n is odd q_i is positive for $i \leq i_0 = 2^{n-1}$ and $Q_i - i$ is maximum for $i = i_0$. We have

$$Q_{i_0} - i_0 = \sum_{k=0}^{[n/2]} (n-2k)\binom{n}{k} - 2^{n-1} = ([n/2]+1)\binom{n}{[n/2]} - 2^{n-1}.$$

A similar computation for n even gives

$$2^{n-1} - \binom{n}{n/2}$$

for the index i_0 of the maximum and the same expression for $Q_{i_0} - i_0$. This completes the proof.

If $c_1, \cdots, c_{n+1}$ are real numbers let $m(c_1, \cdots, c_{n+1})$ be the number of indices j for which

$$|c_j| > |\sum_{i \neq j} c_i|.$$

We now consider $n+1$ positive numbers $a_1, \cdots, a_{n+1}$ which are 'free' in the sense explained above, and define

$$M = M(a_1, \cdots, a_{n+1}) = \sum m(\pm a_1, \cdots, \pm a_{n+1}),$$

the summation being taken over all combinations of plus signs and minus signs.

LEMMA 2.

$$2^{n+1} \leq M \leq 4([n/2]+1)\binom{n}{[n/2]}.$$

It is clear that $M = 2^{n+1}$ if

$$a_{n+1} > a_1 + \cdots + a_n,$$

and we reduce the other cases to this one by computing the change in M as a_{n+1} is increased to $a_1 + \cdots + a_n + 1$. Using the notation of Lemma 1, we suppose that $S_{i+1} < a_{n+1} < S_i$, where i of course is not greater than 2^{n-1}, and that a'_{n+1} is a number slightly greater than S_i. We now compare $M(a_1, \cdots, a_n, a_{n+1})$ with $M(a_1, \cdots, a_n, a'_{n+1})$. The inequality $a_{n+1} < S_i$ becomes $a'_{n+1} > S_i$ if a_{n+1} is replaced by a'_{n+1}, and we see that there is a contribution $+4$ to M coming from the terms $\pm a'_{n+1}$ in the four sums $\pm S_i \pm a'_{n+1}$. In like manner, each $+a_j$ occurring in S_i contributes -4 to M, and each $-a_j$ in S_i contributes $+4$ if j is less than $n+1$. So

$$M(a_1, \cdots, a_n, a_{n+1}) - M(a_1, \cdots, a_n, a'_{n+1}) = 4(q_i - 1),$$

where q_i has the meaning explained at the beginning of this section. Thus increasing a_{n+1} to $a_1 + \cdots + a_n + 1$ decreases M by

$$4(Q_i - i) = 4 \sum_{j \leq i} (q_j - 1),$$

and Lemma 2 follows from Lemma 1.

There is another more direct way of establishing the first inequality of Lemma 2. Since the inequality is trivial for $n = 1$, we proceed by induction. Considering the numbers $(a_1 + a_2)$, $a_3, \cdots, a_{n+1}$ we assume that there are at least 2^{n-2} inequalities of the form

$$a_j > U \qquad (j > 2) \tag{2}$$

or

$$(a_1 + a_2) > V, \tag{3}$$

where the right members are positive, and U is a sum over $(a_1 + a_2)$, $a_3, \cdots$, $a_{j-1}, a_{j+1}, \cdots, a_{n+1}$ with appropriate signs, and V is a sum over $a_3, \cdots, a_{n+1}$. From (2) we obtain an inequality (2′) by dropping the parentheses from $(a_1 + a_2)$ in U; from (3) we obtain an inequality (3′): $a_1 > a_2 - V$ or $a_1 > V - a_2$ according as a_2 is greater or less than V (we assume without loss of generality that $a_1 > a_2$). We consider also the numbers $(a_1 - a_2)$, $a_3, \cdots, a_{n+1}$ and inequalities

$$a_j > \overline{U} \qquad (j > 2) \tag{4}$$

$$(a_1 - a_2) > \overline{V}, \tag{5}$$

of which we assume there are at least 2^{n-2}. From (4) we derive an inequality (4′) by dropping the parentheses from $(a_1 - a_2)$ in $\overline{U}$, and from (5) we derive an inequality (5′): $a_1 > a_2 + \overline{V}$. It is easy to see that no two of the primed inequalities are the same. Hence there must be at least $2 \cdot 2^{n-2} = 2^{n-1}$ inequalities

$$a_i > \sum_{j \neq i} \pm a_j \qquad (1 \leq i \leq n + 1)$$

in which the right member is positive. Taking into account the four possibilities of attributing signs to the members of each inequality we get the first statement of the lemma.

We now translate our result into terms of probability.

LEMMA 3.

$$\frac{1}{n+1} \le \Pr\{|x_{n+1}| > |x_1 + \cdots + x_n|\} \le \phi(n).$$

Here of course the random variables satisfy the conditions imposed at the beginning of § 1, and $\phi(n)$ is the function defined there. Since the joint distribution of the x_i is unchanged by permuting the x_i or by multiplying an x_i by -1, we have

$$\begin{aligned}
\Pr\left\{|x_{n+1}| > \left|\sum_1^n x_i\right|\right\} &= \frac{1}{n+1}\sum_{i=1}^{n+1} \Pr\left\{|x_i| > \left|\sum_{j\neq i} x_j\right|\right\} \\
&= \frac{1}{n+1}\,\mathrm{E}\{m(x_1, \cdots, x_{n+1})\} \\
&= \frac{1}{n+1}\,\mathrm{E}\left\{\frac{1}{2^{n+1}}\sum_{+,-} m(\pm|x_1|, \cdots, \pm|x_{n+1}|)\right\} \\
&= \frac{1}{(n+1)2^{n+1}}\,\mathrm{E}\{M(|x_1|, \cdots, |x_{n+1}|)\},
\end{aligned}$$

where m and M are the functions defined above. Since $|x_1|, \cdots, |x_{n+1}|$ are 'free' with probability one (because the distribution of the x_i is continuous), Lemma 3 follows at once from Lemma 2.

Our later proofs could be made somewhat simpler than they stand if we could use the inequality

$$\frac{m}{m+n} \le P_{m,n} \equiv \Pr\left\{\left|\sum_1^n x_i\right| < \left|\sum_{n+1}^{n+m} x_i\right|\right\} \le \phi([n/m])$$

for $m \le n$. This generalization of Lemma 3 we have been unable to prove; and indeed a corresponding generalization of Lemma 2 is false. However, we shall use

$$(6) \qquad P_{m,n} \le 6\phi([n/m]) < 3[n/m]^{-1/2},$$

and establish it in the following manner:

Let $a = [n/m]$, and write

$$u = x_1 + \cdots + x_{am},$$

$$v = x_{am+1} + \cdots + x_n,$$

$$w = x_{n+1} + \cdots + x_{n+m},$$

$$z = y_{n+1} + \cdots + y_{am+m},$$

where the y_k have the same distribution as the x_j, and the x_j and y_k taken together form an independent set of random variables. Let E be the set on which the four inequalities

$$|w| < |u \pm v \pm z|$$

hold; by Lemma 3 the probability of any one of these inequalities is at least $1 - \phi(a+1)$; hence E has probability at least $1 - 4\phi(a+1)$. Similarly the probability of the set F on which the two inequalities $|v \pm z| < |u|$ hold in at least $1 - 2\phi(a)$. Now clearly $|u+v| > |w|$ on EF and also

$$\Pr\{EF\} \geq 1 - 2\phi(a) - 4\phi(a+1) \geq 1 - 6\phi(a).$$

3. Proofs of Theorems 1, 2, 3. It is easy to see that the probability of s_k and s_{k+1} differing in sign is one-half the probability of s_{k+1} being larger in absolute value than s_k. Thus

$$\mathrm{E}\{N_n\} = \sum_1^n \Pr\{s_k\, s_{k+1} < 0\} = \frac{1}{2}\sum_1^n \Pr\{|x_{k+1}| > |s_k|\},$$

and Lemma 3 implies Theorem 1.

Let us turn to Theorem 3. Clearly the probability of s_k' and

$$s_{2k'} = \sum_1^{2k'} x_j$$

differing in sign is 1/4. Also, $s_{k+a} - s_{2k'}$ is independent of both s_k' and $s_{2k'}$, for

$$(k+a)' \geq (1+\alpha)^a k' \geq 2k'.$$

Thus $s_{k+a}' - s_{2k'}$ has an even chance of taking on the same sign as $s_{2k'}$; so we must have

$$\Pr\{s_k' s_{k+a}' < 0\} \geq \frac{1}{2}\Pr\{s_k' s_{2k'} < 0\} = 1/8.$$

Now, if $s_k' s_{k+a}' < 0$ then must be at least one change of sign in the sequence $s_k', s_{k+1}', \cdots, s_{k+a}'$. Hence, if p_k is the probability of s_k' and s_{k+1}' differing in sign, we have

$$p_k + \cdots + p_{k+a-1} \geq \frac{1}{8},$$

and consequently

$$\mathrm{E}\{N_n'\} = \sum_1^n p_k \geq \frac{1}{8}[n/a]. \tag{7}$$

This proves the first half of the theorem.

As a preliminary to proving the second half of the theorem we show that the variance of N_n' is $O(n)$ by estimating the probabilities

$$p_{i,j} = \Pr\{s_i' s_{i+1}' < 0 \ \& \ s_j' s_{j+1}' < 0\}.$$

Suppose that $i < j$; set

$$u = s_i', \quad v = s_{i+1}' - s_i', \quad w = s_j' - s_{i+1}', \quad z = s_{j+1}' - s_j';$$

and define the events

$$\begin{aligned}
A &: uv < 0, \\
B &: |u| < |v|, \\
C &: (u+v+w)z < 0, \\
D &: |u+v+w| < |z|, \\
D' &: |w| < |z|, \\
E &: |z-w| > |u+v|.
\end{aligned}$$

Then

$$p_i = \Pr\{AB\},\ p_j = \Pr\{CD\}, \text{ and } p_{i,j} = \Pr\{ABCD\}.$$

One sees immediately that A, B, C, D' are independent, and that $ED = ED'$. Writing $\widetilde{E}$ for the complement of E, we have

$$ABCD = \widetilde{E}ABCD + EABCD' \subset \widetilde{E} + ABCD',$$

and

$$D' \subset \widetilde{E} + D.$$

Hence

$$\begin{aligned}\Pr\{ABCD\} &\leq \Pr\{\widetilde{E}\} + \Pr\{ABC\}\Pr\{D'\}\\ &\leq \Pr\{\widetilde{E}\} + \Pr\{ABC\}(\Pr\{\widetilde{E}\} + \Pr\{D\})\\ &\leq \Pr\{AB\}\Pr\{C\}\Pr\{D\} + 2\Pr\{\widetilde{E}\} = p_i\, p_j + 2\Pr\{\widetilde{E}\}.\end{aligned}$$

Note now that $z - w$ is the sum of $(j+1)' - (i+1)'$ of the x's, and $u + v$ is the sum of $(i+1)'$, of the x's, and that moreover

$$(j+1)' - (i+1)' \geq [(1+\alpha)^{j-i} - 1]\ (i+1)'.$$

We may thus apply the inequality (6) following Lemma 3 to obtain

$$\Pr\{\widetilde{E}\} < 3\,[(1+\alpha)^{j-i} - 2]^{-1/2}$$

provided $j - i \geq a$. This yields an upper bound for $p_{i,j}$; a similar argument yields a corresponding lower bound. We have finally

$$p_{i,j} = p_i\, p_j + O\{|1+\alpha|^{-|i-j|/2}\}$$

for all i and j. This estimate shows that

$$\begin{aligned}(8)\qquad \mathrm{E}\{N_n'^{\,2}\} &= \sum_{1\leq i,j\leq r} p_{ij}\\ &= \sum p_i\, p_j + \sum O\{(1+\alpha)^{-|i-j|/2}\} = \mathrm{E}\{N_n'\}^2 + O(n).\end{aligned}$$

Let us denote $\mathrm{E}\{N_k'\}$ by b_k. It follows from (7), (8), and Tchebycheff's inequality that

$$\Pr\left\{\left|\frac{N'_k}{b_k} - 1\right| > \epsilon\right\} < \frac{c}{\epsilon^2 k}$$

for an appropriate constant c and for all positive ϵ. Thus

$$\Pr\left\{\left|\frac{N'_{k^2}}{b_{k^2}} - 1\right| > \epsilon\right\}$$

is the kth term of a convergent series, so that according to the lemma of Borel and Cantelli

$$\frac{N'_{k^2}}{b_{k^2}} \longrightarrow 1$$

with probability one. Note also that

$$\frac{b_{k^2}}{b_{(k+1)^2}} \longrightarrow 1.$$

Now for every natural number n we have

$$\frac{N'_{k^2}}{b_{(k+1)^2}} \leq \frac{N'_n}{b_n} \leq \frac{N'_{(k+1)^2}}{b_{k^2}},$$

with k so chosen that $k^2 \leq n < (k+1)^2$. Since the extreme members tend to one as n increases, the proof of the second half of Theorem 3 is complete.

Theorem 2 is obtained from Theorem 3 in the following way. Let r be a large integer and let $1', 2', \cdots$ be the sequence

$$\begin{array}{l} r, (r+1), \\ r^2, r(r+1), (r+1), (r+1)^2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ r^l, r^{l-1}(r+1), \cdots, (r+1)^l, \\ r^m, r^{m-1}(r+1), \cdots, (r+1)^m, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot, \end{array}$$

where m is defined by

$$r^{m+1} \geq (r+1)^{l+1} > r^m .$$

Let us call j 'favorable' if $(j+1)' = (1+1/r)j'$. Then it is easy to see that:

a) $(1+1/r)j' \leq (j+1)' \leq (1+r)j'$ for all j;

b) there are $k + o(k)$ favorable j less than k (as $k \longrightarrow \infty$);

c) $\log k' = k \log (1+1/r) + o(k)$.

Now, if j is favorable then

$$j' = r\{(j+1)' - j'\}$$

and we may apply Lemma 3 to s_j' and $s_{j+1}' - s_j$. Thus

$$\Pr\{s_j' s_{j+1}' < 0\} = \frac{1}{2} \Pr\{|s_{j+1}' - s_j'| > |s_j'|\} \geq \frac{1}{2(1+r)} .$$

Hence

$$E\{N_k'\} = \sum_{j=1}^{k} \Pr\{s_j' s_{j+1}' < 0\}$$

$$\geq \sum_{j \text{ favorable}} \Pr\{s_j' s_{j+1}' < 0\} \geq \frac{k}{2(r+1)} + o(k).$$

Note that for every natural number n

$$\frac{N_n}{\log n} \geq \frac{N_k'}{\log (k+1)'} ,$$

where k is chosen so that $k' \leq n < (k+1)'$. Consequently

$$\liminf_{n \to \infty} \frac{2N_n}{\log n} \geq \liminf_{k \to \infty} \frac{2N_k'}{\log (k+1)'} = \liminf \frac{2N_k'}{(k+1)\log (1+1/r)}$$

$$\geq \liminf \frac{N_k'}{E\{N_k'\}(r+1)\log(1+1/r)} = \frac{1}{(r+1)\log(1+1/r)}.$$

Letting $r \longrightarrow \infty$ we have Theorem 2.

4. An example. Our construction of a sequence $x_1, x_2, \cdots$ for which $N_n/\log n \longrightarrow 1/2$ with probability one depends on the following observations. For given k define the random index $i = i(k)$ by the condition

$$|x_i| = \max_{1 \leq j \leq k+1} |x_j|,$$

and let A_k be the event $|x_i| > \sum |x_j|$, where the summation is over $j \neq i$, $1 \leq j \leq k+1$. Let f_k be the characteristic function of the event '$s_k\ s_{k+1} < 0$,' and g_k is the characteristic function of the event '$i(k) = k+1$ and further $(x_1 + \cdots + x_k)x_{k+1} < 0$'. It is clear that $g_1, g_2, \cdots$ are independent random variables, that

$$2\,\mathrm{Pr}\{g_k = 1\} = \frac{1}{(k+1)},$$

and that the strong law of large numbers applies to the sequence $g_1, g_2, \cdots$ also $f_k = g_k$ on A_k; if moreover $\sum \mathrm{Pr}\{\widetilde{A}_k\} < \infty$ (here $\widetilde{A}_k$ is the complement of A_k) then, with probability one, $f_k = g_k$ for all but a finite number of indices. In this case we have, with probability one,

$$N_n = \sum_{k=1}^{n} f_k = \sum_{k=1}^{n} g_k + O(1) = \sum_{k=1}^{n} \frac{1}{2(k+1)} + o(\log n),$$

the last step being the strong law of large numbers applied to $g_1, g_2, \cdots$. Thus, in order to produce the example, we have only to choose the x_j so that, say,

$$\mathrm{Pr}\{\widetilde{A}_k\} = O(k^{-2}).$$

To do this we take $x_j = \pm \exp(\exp 1/u_j)$, where $u_1, u_2, \cdots$ is a sequence of independent random variables each of which is uniformly distributed on the interval $(0, 1)$ and the $\pm$ stands for multiplication by the jth Rademacher function. For a given k let y and z be the least and the next to least of $u_1, \cdots, u_{k+1}$. The joint density function of y and z is

$$(k+1)\,k\,(1-z)^{k-1} \qquad (0 < y < z < 1).$$

Consequently the event

$$D_k : \frac{1}{y} > \frac{1}{z} + \frac{1}{k^2}$$

has probability

$$k(k+1) \int_0^{k^2/(k^2+1)} dy \int_{k^2 y/(k^2-y)}^{1} (1-z)^{k-1} dz = 1 + O(k^{-2}),$$

and the event $E_k : 1/z > 3 \log k$ also has probability $1 + O(k^{-2})$. It is easy to verify that the event A_k defined above contains $D_k E_k$; thus

$$\Pr\{\widetilde{A}_k\} = O(k^{-2}),$$

and our example is completed.

5. Proof of Theorem 4. We prove Theorem 4 in the form

$$T_n \equiv \sum_{\substack{1 \le k \le n \\ s_k > 0}} \frac{1}{k} = \frac{1}{2} \log n + o(\log n)$$

by much the same method as we proved Theorem 2. First,

$$\mathrm{E}\{T_n\} = \frac{1}{2} \sum_{1}^{n} \frac{1}{k} = \frac{1}{2} \log n + o(1).$$

Next, the inequality following Lemma 3 yields

$$\Pr\{|s_l - s_k| < |s_k|\} \le 3\left[\frac{k}{l-k}\right]^{1/2} \qquad (l \ge 2k),$$

so that

$$\Pr\{|s_l - s_k| < |s_k|\} = O\left(\frac{k}{l}\right)^{1/2}$$

for $l > k$. Consequently

$$\Pr\{s_k > 0 \ \& \ s_l > 0\} = \frac{1}{4} + O\left(\frac{k}{l}\right)^{1/2} \qquad (l > k).$$

This implies that

$$E\{T_n^2\} = \sum_{1\le k,\, l\le n} \frac{1}{kl} \Pr\{s_k > 0 \;\&\; s_l > 0\}$$

$$= \sum_{1\le k\le n} \frac{1}{k^2} \Pr\{s_k > 0\} + 2 \sum_{1\le k< l\le n} \frac{1}{kl} \Pr\{s_k > 0 \;\&\; s_l > 0\}$$

$$= \frac{1}{2} \sum_{1}^{n} \frac{1}{k^2} + 2 \sum_{1\le k\le l\le n} \frac{1}{kl} \left\{ \frac{1}{4} + O\left(\frac{k}{l}\right)^{1/2} \right\}$$

$$= \frac{1}{4} (\log n)^2 + O(\log n).$$

Thus the variance of T_n is of the order of $\log n$. Setting $n(k) = 2^{k^2}$, we have, according to Tchebycheff's inequality,

$$\Pr\left\{ \left| \frac{T_{n(k)}}{\log n(k)} - 1 \right| > \epsilon \right\} \le \frac{c}{\epsilon^2 k^2}$$

for an appropriate constant c and all positive ϵ. Since the right member is the kth term of a convergent series, the lemma of Borel and Cantelli implies that

$$\frac{T_{n(k)}}{\log n(k)} \longrightarrow 1$$

with probability one. Note also that

$$\frac{\log n(k+1)}{\log n(k)} \longrightarrow 1.$$

Now, for any n,

$$\frac{T_{n(k)}}{\log n(k+1)} \le \frac{T_n}{\log n} \le \frac{T_{n(k+1)}}{\log n(k)},$$

where k is so chosen that $n(k) \le n \le n(k+1)$. Here the extreme members almost certainly tend to one as n increases. This proves Theorem 4.

References

1. Kai Lai Chung and G. A. Hunt, *The zeros of* $\sum \pm 1$, Ann. of Math. 50 (1949), 385-400.

2. P. Lévy, *Théorie de l'addition des variables aléatoires*, Paris, 1937.

University College of London

Cornell University

National Bureau of Standards, Los Angeles