Reprinted from MEMOIRS OF THE AMERICAN MATHEMATICAL SOCIETY, Number 6.

PROBABILITY LIMIT THEOREMS ASSUMING ONLY THE FIRST MOMENT I

By

K. L. CHUNG and P. ERDÖS

In this paper we consider sums of mutually independent, identically distributed random variables. An essential feature is that we assume only that the first moment is zero, or that both its positive and negative parts diverge. Part I here deals with lattice distributions. Perhaps the main results are Theorem 3.1 and Theorem 8. We hope to take up other cases later.

1. Let X be a random variable which assumes only integer values

$$P(X = k) = p_k$$

$$p_k \geq 0 \quad \sum p_k = 1^* .$$

A number is said to be a 'possible' value of an integer-valued random variable if its probability is positive. The possible values of X will be denoted by $u_i, i=1,2,\ldots;$ they may be finite or infinite in number. As usual $S_n = \sum_{k=1}^{n} X_k$ where the X_k are mutually independent, each having the same distribution as X.

To avoid minor complications, we shall assume that <u>every integer</u> c <u>is a</u> possible values of S_n if n is sufficiently large: $n \geq n_0(c)$. A set of necessary and sufficient conditions for this is the following:

(1) The u_i are not all of the same sign;

(2) The greatest common divisor of the set of differences $u_i - u_j, i,j=1,2,\ldots$ is equal to 1.

We shall call the following two sets of assumptions (0) and (∞) respectively:

$$(0) \qquad E(|X|) = \sum |k| p_k < \infty, \quad E(X) = \sum k p_k = 0$$

$$(\infty) \qquad \tfrac{1}{2} E(|X|+X) = \sum_{k=0}^{\infty} k p_k = \infty, \ \tfrac{1}{2} E(|X|-X) = -\sum_{k=-\infty}^{0} k p_k = \infty.$$

Thus under (0) or (∞) (1) is always satisfied except in the trivial case $X \equiv 0$, which we exclude. If (2) is not satisfied, there are two possibilities: either all possible

Received by the editors January 5, 1950.

*In an unspecified summation the index runs from $-\infty$ to $+\infty$.

values of S_n are multiples of an integer > 1; or there exists an integer $m > 1$ and a complete residue class mod. m, $r_1, \ldots, r_m$ such that for a fixed j, all possible values of S_{nm+j}, n=1,2,···, belong to the same residue class $r_j \pmod m$. It is not difficult to see how our statements and proofs should be modified for these cases.

In the following the letters a, a' denote arbitrary integers, A,A',B positive constants; ε, ε' arbitrarily small constants.

2. In this section we give some simple theorems on the bounds of $P(S_n=a)$. It is well known that under more restrictive assumptions more precise results can be obtained (see Gnedenko [1], van Kampen and Wintner [2], Esseen [3]).

THEOREM 1. Under no assumptions about moments whatsoever,

$$P(S_n=a) \leq An^{-1/2} \tag{1}$$

where A does not depend on n or a. If $E(X^2) = \infty$, then

$$\lim_{n\to\infty} n^{1/2}P(S_n=a) = 0. \tag{2}$$

Proof. The c.f.* of the d.f. † of X is

$$f(x) = \sum p_k e^{ikx} .$$

The c.f. of S_n if $(f(x))^n$, and we have

$$P(S_n = a) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x))^n e^{-iax}dx.$$

Suppose first that n is even: n=2m. We have

$$\left| \int_{-\pi}^{\pi} (f(x))^{2m}e^{-iax}dx \right| \leq \int_{-\pi}^{\pi} (|f(x)|^2)^m dx.$$

Now $|f(x)|^2$ is the c.f. of a symmetrical d.f., namely that of X + X' where X,X' are mutually independent and X' has the same distribution as −X. Hence we may write

$$|f(x)|^2 = \sum_{k=0}^{\infty} r_k \cos kx, \left(r_k \geq 0, \sum_{k=0}^{\infty} r_k = 1\right)$$

---

*Characteristic function or Fourier-Stieltjes transform.

† Distribution function.

$$= 1 - 2 \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2}.$$

Suppose $r_\ell > 0$. If $x \leq \pi \ell^{-1}$,

$$2 \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2} \geq 2\left(\sum_{k=0}^{\ell} r_k \left(\frac{kx}{\pi}\right)^2\right) \geq \frac{2x^2}{\pi^2} \sum_{k=0}^{\ell} r_k k^2 > Ax^2$$

$$|f(x)|^2 < 1 - A x^2.$$

Hence

$$\int_{-\pi}^{\pi} |f(x)|^{2m} dx < \int_{-\pi\ell^{-1}}^{\pi\ell^{-1}} (1-A x^2)^m dx + \int_{\pi\ell^{-1}<|x| \leq \pi} |f(x)|^{2m} dx.$$

It is known that if $\eta < |x| \leq \pi$, then $|f(x)| < 1 - \varepsilon(\eta)$. Therefore we have

$$\int_{-\pi}^{\pi} |f(x)|^{2m} dx \leq \int_{-\pi\ell^{-1}}^{\pi\ell^{-1}} e^{-A mx^2} dx + O((1-\varepsilon)^m).$$

(1) follows for even n. Noticing that $|f(x)|^n \leq |f(x)|^{n-1}$ we see that the same proof goes through for an odd n.

To prove (2), notice that the assumption $E(X^2) = \infty$ implies that $E((X+X')^2) = \infty$. Hence

$$\sum_{k=0}^{\infty} k^2 r_k = \infty,$$

and the A' in the foregoing can be taken arbitrarily large. q.e.d.

A lower bound for $P(S_n=a)$, under the assumption (0) or (∞), will be given in Theorem 2.2; we shall also show that our estimate is close to the best possible by exhibiting an example in Theorem 2.3. In one special case, however, we can prove a much stronger result, and this is Theorem 2.1.

THEOREM 2.1. If the d.f. of X is symmetrical, and $E(|X|) < \infty$, then

(3) $$\lim_{n \to \infty} n\, P(S_n=a) = \infty.$$

Proof. Since $p_k = p_{-k}$, $f(x)$ is real. Since $f(0) = 1$ and $f(x)$ is continuous, there exists a $\delta > 0$ such that if $|x| < \delta$, $f(x) > 0$. We have

$$P(S_n=a) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x))^n \cos ax\,dx,$$

$$\geq \frac{1}{4\pi} \int_{-\delta}^{\delta} (f(x))^n dx - O((1-\varepsilon)^n),$$

if $\delta < \frac{\pi}{3a}$. As in the proof of Theorem 1, we can write

$$1-f(x) = \sum_{k=0}^{\infty} 2r_k \sin^2 \frac{kx}{2}.$$

Since $\sum_{k=0}^{\infty} kr_k < \infty$,

$$\lim_{x \to 0} \frac{1}{|x|} \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2} = 0.$$

Hence given $\varepsilon > 0$, if $|x| < \delta_0(\varepsilon) < \delta$, $1 - f(x) \leq \varepsilon|x|$. Now

$$\int_{-\delta}^{\delta} (f(x))^n dx \geq \int_{-\delta_0}^{\delta_0} (1-\varepsilon|x|)^n dx = \frac{2-2(1-\delta_0\varepsilon)^{n+1}}{(n+1)\varepsilon}.$$

Since ε is arbitrary (3) follows.

THEOREM 2.2. <u>Under</u> (0) <u>or</u> (∞) <u>we have for every</u> $\varepsilon > 0$

(4) $$P(S_n=a) \geq (1-\varepsilon)^n$$

<u>if</u> $n \geq n(\varepsilon, a)$.

Proof. If the possible values of X are bounded, then $E(X^2) < \infty$. In this case it is well known and also easy to show that

$$\lim_{n \to \infty} n^{1/2} P(S_n=a) = A < \infty.$$

This is a much sharper result than (4). Hence we may assume that there are possible values of arbitrarily large magnitude.

Given $\varepsilon > 0$, there exist arbitrarily large z_1 and z_2 such that

$$\sum_{-z_1}^{z_2} p_k > 1-\varepsilon$$

and if $k > z_2, p_k < \varepsilon$. Now choose h' so large that

$$\left|\sum_{-h'}^{0} kp_k\right| > \sum_{0}^{z_2} kp_k$$

this is possible under (0) or (∞). Also there is a unique h such that

$$\sum_{0}^{h-1} kp_k < \left| \sum_{-h'}^{0} kp_k \right| \leq \sum_{0}^{h} kp_k.$$

Then $h > z_2$ and

$$hp_h \geq \sum_{-h'}^{h} kp_k = C \geq 0.$$

Define $p'_k = p_k$ if $k \neq h$, but

$$p'_h = p_h - Ch^{-1} \geq 0.$$

Then

$$b = \sum_{-h'}^{h} p'_k > 1 - \varepsilon - p_h > 1 - 2\varepsilon$$

$$\sum_{-h'}^{h} kp'_k = 0.$$

Now define a random variable X' as follows:

$$P(X' = k) = p'_k b^{-1} \qquad \text{if} \qquad -h' \leq k \leq h$$

$$= 0 \qquad \text{otherwise.}$$

Let $S'_n = \sum_{k=1}^{n} X'_k$ where the X'_k are mutually independent and each has the same distribution as X'. Since $p'_k \leq p_k$ for all k

$$P(S'_n = a) \geq b^{-n} P(S_n = a \mid -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n).*$$

Hence

$$P(S_n = a) \geq P(S_n = a;\ -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n)$$

$$\geq P(-h' \leq x \leq h)^n P(S_n = a \mid -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n)$$

$$\geq (1-\varepsilon)^n b^n P(S'_n = a) \geq (1-\varepsilon)^n (1-2\varepsilon)^n A n^{-1/2}$$

where A depends on ε by definition of X'. This being true for all ε is equivalent to (4).

The idea of truncation in the preceding proof is due to Shizuo Kakutani. Theorem 2.2 was first proved under (0) by W. H. J. Fuchs using a result in Chung and Fuchs [4], namely

---

*$P(E|F)$ denotes the conditional probability of E under the hypothesis F.

$$\sum_{n=1}^{\infty} P(S_n = a) = \infty.^* \tag{5}$$

A similar proof using (5) was also given by Kakutani. We sketch the latter proof as follows.

From (5) it follows by the Cauchy-Hadamard criterion

$$\overline{\lim_{n\to\infty}} \ (P(S_n=0))^{1/n} = 1.$$

Hence given $\varepsilon > 0$, there exists arbitrarily large m such that

$$P(S_m = 0) \geq (1-\varepsilon)^m.$$

Consequently for all integers $k > 0$,

$$P(S_{km} = 0) \geq (1-\varepsilon)^{km}.$$

We can also choose the aforesaid m so large that

$$\min_{m<v\leq 2m} P(S_v=a) = A' > 0.$$

Now fix m. If $n=(k+1)m+r$, $k > 0$, $\varepsilon < r < m$, we have

$$P(S_n=a) \geq P(S_{m+r} = a)P(S_{km} = 0)$$

$$\geq A'(1-\varepsilon)^{km} \geq A'(1-\varepsilon)^{-m}(1-\varepsilon)^n. \qquad \text{q.e.d.}$$

THEOREM 2.3. <u>We can construct an example satisfying</u> (0) <u>and such that for every given</u> $B > 0$ <u>there exists a sequence</u> $\{n_v\}$ <u>for which</u>

$$P(S_{n_v} \geq 0) = O(n_v^{-B}).$$

Proof. Let $A_v, v=1,2,\ldots$ be a sequence of positive integers increasing to ∞ so fast that for every $\varepsilon > 0$,

$$A_v = O(A_{v+1}^{\varepsilon}).$$

Define

$$X = \begin{cases} -1 \text{ with prob. } \frac{1}{2} \\ \\ A_v \text{ with prob. } 2^{-v}A_v^{-1} \end{cases} \qquad v=2,3,\cdots .$$

Then $E(X) = 0$. If k is sufficiently large

$$P(\max_{1\leq v\leq n} X_v > A_k) \leq n \sum_{v=k+1}^{\infty} \frac{1}{2^v A_v} \leq \frac{n}{A_{k+1}}.$$

---

*However, the assumption (∞) does not imply the truth of (5) (see [4]); thus the following proof does not hold under (∞).

Let

$$X^* = \begin{cases} X \text{ if } X \leq A_k \\ 0 \text{ if } X > A_k \end{cases} .$$

$S^*_n = \sum_{v=1}^{n} X^*_v$ where the X^*_v are mutually independent and each has the distribution of X^*.

We have

$$E(X^*) = -2^{-(k+1)}, \; E(X^{*2}) = O(A_k)$$

$$P(S_n \geq 0) | \max_{1 \leq v \leq n} X_v \leq A_k) = P(S^*_n \geq 0)$$

$$\leq P(|S^*_n - E(S^*_n)| \geq |E(S^*_n)|).$$

Let m be an integer > 0, a routine computation shows that

$$E(|S^*_n - E(S^*_n)|^{2m} \leq K_m (\sum_{v=1}^{n} E(X^* + 2^{-(k+1)})^2)^m$$

$$\leq K'_m A_k^m n^m$$

where K_m, K'_m are two positive constants depending only on m. Hence

$$P(|S^*_n - E(S^*_n)| \geq |E(S^*_n)|) \leq O(A_k^m \, 2^{(k+1)2m} n^{-m}).$$

Now choose

$$n_k \sim A_{k+1}^{\varepsilon}$$

we have, by the property of the sequence A_k,

$$P(S_{n_k} \geq 0) \leq P(\max_{1 \leq v \leq n_k} X_v > A_k) + P(S_{n_k} \geq 0 | \max_{1 \leq v \leq n_k} X_v \leq A_k)$$

$$\leq O(n_k^{-1/\varepsilon + 1} + n_k^{-m+\varepsilon}) = O(n_k^{-B})$$

by choice of ε and m.

Theorem 2.3 should be compared with a result due to Feller [5].

3. The theorems in § 2 were proved by fairly standard analytical methods. We are unable to prove the theorems in this section by similar methods, except in the case where the d.f. of X is symmetrical, i.e., the c.f. is a real-valued function. In this case we have as before

$$|P(S_n=a) - P(S_n=a')| \leq \int_{-\delta}^{\delta} |\cos ax - \cos a'x| \, |f(x)|^n \, dx + O((1-\varepsilon)^n).$$

Choosing δ so small that $\cos ax > 0$, $f(x) > 0$, and $|\cos ax - \cos a'x| < \varepsilon' \cos ax$ for $|x| \leq \delta$, we have

$$|P(S_n=a) - P(S_n=a')| \geq \varepsilon' \int_{-\delta}^{\delta} \cos ax (f(x))^n \, dx + O((1-\varepsilon)^n).$$

On account of Theorem 2.2 it follows that

$$P(S_n=a) - P(S_n=a') = o\,(P(S_n=a))$$

which is equivalent to Theorem 3.1 below. We have not been able to prove the theorem by this method when $f(x)$ is not real-valued. Another relevant remark is the following: if instead of the individual probabilities $P(S_n=a)$ we consider their sums, then it follows from a theorem due to Doeblin [11] on Markov chains that

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} P(S_k=a)}{\sum_{k=1}^{n} P(S_k=a')} = 1.$$

THEOREM 3.1. Under (0) or (∞)

$$\lim_{n \to \infty} \frac{P(S_n=a)}{P(S_n=a')} = 1.$$

Proof. For some $k, u_1-u_0, \cdots, u_k-u_0$ have g.c.d. 1. Thus there exist integers c_i' and c_i such that

$$a'-a = \sum_{i=1}^{k} c_i'(u_i-u_0) = \sum_{i=0}^{k} c_i u_i, \quad \sum_{i=0}^{k} c_i = 0.$$

Let $P(X=u_i) = q_i$. Corresponding to every representation of a in the form

$$a = \sum_{i=0}^{\ell} n_i u_i, \qquad n_i \geq 0, \qquad \sum_{i=0}^{\ell} n_i = n \tag{1}$$

there is a realization of the value a by $X_1 + \cdots + X_n$ with probability

$$\frac{n!}{n_0! \cdots n_\ell!} \prod_{i=0}^{\ell} q_i^{n_i} \tag{2}$$

when n_i of the X's assume the value u_i. The total probability of a is thus

$$\sum \frac{n!}{n_0!\cdots n_\ell!} \prod_{i=0}^{\ell} q_i^{n_i}$$

where the sum runs over all representations (1). Now write this sum as

$$(3) \qquad \sum = \sum_1 + \sum_2$$

where in $\sum_1$ the conditions

$$|n_i - nq_i| < \varepsilon n, \qquad 0 \leq i \leq \ell$$

are satisfied, while $\sum_2$ is the rest.

Consider the event $X = u_i$ with probability q_i; n_i is the number of its occurrences in n mutually independent, identical trials. It is well known that the probability that $|n_i - nq_i| > \varepsilon n$ is

$$O(e^{-\varepsilon'^2 n}).$$

Hence

$$(4) \qquad \sum_2 \leq (\ell+1)O(e^{-\varepsilon'^2 n}) = o(P(S_n = a))$$

by Theorem 2.2, for every $\varepsilon > 0$.

Now consider a representation (1) with $|n_i - nq_i| < \varepsilon n$ for $0 \leq i \leq \ell$. If ε is sufficiently small and n sufficiently large, we have $n_i > n(q_i - \varepsilon) > \varepsilon n > |c_i|$. Corresponding to every representation of a in the form (1) there is a representation of a' in the form

$$(5) \qquad a' = \sum_{i=0}^{k} (n_i + c_i)u_i + \sum_{i=k+1}^{\ell} n_i u_i = \sum_{i=0}^{\ell} n_i' u_i$$

where $|n_i' - nq_i| < 2\varepsilon n$. The ratio of two such corresponding probabilities is

$$= \frac{n_0!\cdots n_k!}{n_0'!\cdots n_k'!} q_0^{n_0'-n_0} \cdots q_k^{n_k'-n_k}.$$

If $m' > m$, $|m - nq| < \varepsilon n$, $|m' - nq| < 2\varepsilon n$, we have

$$\frac{m!}{m'!} q^{m'-m} \frac{(qn)(qn)\cdots(qn)}{m'(m'-1)\cdots(m+1)} n^{m-m'}$$

$$\leq \left(\frac{qn}{(q-2\varepsilon)n}\right)^{m'-m} n^{m-m'} = \left(\frac{1}{1-2\varepsilon}\right)^{m'-m} n^{m-m'}$$

$$\frac{m!}{m'!} q^{m'-m} \geq \left(\frac{1}{1+2\varepsilon}\right)^{m'-m} n^{m-m'}.$$

Since $\sum_{i=0}^{\ell} n_i' = \sum_{i=0}^{\ell} n_i$ it follows that

$$(1-\varepsilon')^C \leq \lambda \leq (1+\varepsilon')^C$$

where $C = \sum_{i=0}^{k} |c_i|$. Since ε' is arbitrarily small we have $\lim_{n\to\infty} \lambda = 1$.

Let us write the corresponding formulas (1) and (2) for a':

$$a' = \sum_{i=0}^{\ell} n_i' u_i, \qquad \sum_{i=0}^{\ell} n_i' = 0$$

(6) $$\sum' = \sum_1' + \sum_2'$$

where in $\sum_1'$ the condition $|n_i' - nq_i| < 2\varepsilon n$ are satisfied find $0 \leq i \leq \ell$. We have just proved that

$$\overline{\lim_{n\to\infty}} \frac{\sum_1}{\sum_1'} \leq 1.$$

Using (4) we conclude that

$$\overline{\lim_{n\to\infty}} \frac{P(S_n=a)}{P(S_n=a')} = 1.$$

Since a and a' are interchangeable we obtain Theorem 3.

THEOREM 3.2. For those values of n for which

(7) $$P(S_n=a) \geq n^{-B}$$

for some fixed $B > 0$, we have for every $\varepsilon > 0$

(8) $$|P(S_n=a) - P(S_n=a')| \leq P(S_n=a)An^{-1/2+\varepsilon}$$

where A may depend on a, a' but not on n.

Proof. In (3) we re-define $\sum_1$ to be the sum of those terms for which

$$|n_i - nq_i| < n^{1/2+\varepsilon}, \qquad 0 \leq i \leq \ell.$$

As before, let (5) correspond to (1), but now we assume n so large that $n^{1/2+\varepsilon} > |C_i|$, so that

$$|n_i'-nq_i| < 2n^{1/2+\varepsilon}.$$

We re-define $\sum_1'$ in (6) to be the sum of those terms for which this is true. By well known estimates on the binomial distribution we have

$$(9) \qquad \sum_2 = O(e^{-An^{\varepsilon}}), \quad \sum_2' = O(e^{-An^{\varepsilon}}).$$

Now consider the difference of two corresponding probabilities (1) and (6)

$$d = \frac{n!}{n_0!\cdots n_\ell!} q_0^{n_0}\cdots q_\ell^{n_\ell}(1-\lambda).$$

If $m=nq+r$, $m'=nq+r'$ where $|r-r'| \leq C$ and $|r| \leq n^{1/2+\varepsilon}$, $|r'| \leq 2n^{1/2+\varepsilon}$, an easy application of Stirling's formula yields

$$\frac{m!}{m'!} q^{m'-m} = n^{r-r'}(1+O(n^{-1/2+3\varepsilon})), \qquad \lambda = 1+O(kn^{-1/2+3\varepsilon}).$$

Since $\sum_{i=0}^{k} (r_i-r_i') = 0$ we have

$$|d| \leq \frac{n!}{n_0!\cdots n_\ell!} q_0^{n_0}\cdots q_\ell^{n_\ell} O(n^{-1/2+3\varepsilon}).$$

Hence

$$|P(S_n=a)-P(S_n=a')| \leq O(\textstyle\sum_1 . n^{-1/2+3\varepsilon}) + \sum_2 + \sum_2'.$$

The first term on the right is $O(P(S_n=a).n^{-1/2+3\varepsilon})$, and the other two terms by (7) and (9) are of smaller order of magnitude. Thus (8) follows.

THEOREM 4. <u>Under</u> (0) <u>or</u> (∞)

$$\lim_{n\to\infty} \frac{P(S_n=a)}{P(S_{n+1}=a')} = 1.$$

Proof. For every representation of a in the form (1), there is a representation of $a+u_o$ in the following form

$$a+u_o=(n_o+1)u_o+\sum_{i=1}^{\ell} n_i u_i.$$

The corresponding probability is

$$\frac{(n+1)!}{(n_o+1)!n_1!\cdots n_\ell!} q_o^{n_o+1} q_1^{n_1}\cdots q_\ell^{n_\ell}.$$

The ratio of this to (2) is $(n_o+1)/(n+1)q_o$. If $|n_o-nq_o|<\varepsilon n$, this ratio is between $1-\varepsilon/q$ and $1+\varepsilon/q_o$ as $n\longrightarrow\infty$. The range at values of n_o such that $|n_o-nq_o|>\varepsilon n$ can be neglected as before. It follows exactly as in the proof of Theorem 3 that

$$\overline{\lim_{n\longrightarrow\infty}} \frac{P(S_n=a)}{P(S_{n+1}=a+u_o)} \leq 1.$$

By virtue of Theorem 3 this gives

$$\overline{\lim_{n\longrightarrow\infty}} \frac{P(S_n=a)}{P(S_{n+1}=a)} \leq 1.$$

Considering $a-u_o$ instead of $a+u_o$ in the above in a similar manner we arrive at

$$\overline{\lim_{n\longrightarrow\infty}} \frac{P(S_n=a)}{P(S_{n-1}=a)} \leq 1.$$

These last two relations combined are equivalent to Theorem 4.

We remark that Theorem 4 can be proved in the same way as sketched above for Theorem 3.1, when $f(x)$ is real-valued. It would also seem that we might be able to deduce Theorem 4 directly from Theorem 3.1, but a trivial argument gives only the following. Since

$$P(S_{n+1}=a)=\sum_{a'=-\infty}^{\infty} P(S_n=a')\,P(X=a-a')$$

$$\geq \sum_{a'=-A}^{A} P(S_n=a')P(X=a-a').$$

It follows easily, using Theorem 3.1, that

$$\underline{\lim_{n\longrightarrow\infty}} \frac{P(S_{n+1}=a)}{P(S_n=a)} \geq 1.$$

But the other half of the result seems difficult.

4. In this section we study the number of a-values in the sequence $S_1, \ldots, S_n$. A very special case has been treated more or less completely by Chung and Hunt [6]. More general cases, in which the existence of certain moments are assumed, have been considered by Feller [7] and Chung [8].* In this paper we are considering a more general situation and precise results are not hoped for at this moment. However, we shall prove the relevant Theorem 8 whose truth would perhaps be considered evident but whose proof, as far as we can make it, is by no means simple. Theorem 7 gives the true bounds within an ε power.

Define

$$Y_k = \begin{cases} 1 \text{ if } S_k = a \\ 0 \text{ if } S_k \neq a \end{cases}$$

$$E(Y_k) = P(S_k = a) = m_k$$

$$T_n = \sum_{k=1}^{n} Y_k$$

$$E(T_n) = M_n = \sum_{k=1}^{n} m_k$$

and similarly Y'_k, m'_k, T'_n, M'_n, for a'.

THEOREM 5. <u>Under</u> (0), <u>for every</u> $\varepsilon > 0$,

$$P(|T_n - T'_n| > M_n^{3/4+\varepsilon} \text{ i.o.**}) = 0.$$

Proof. By Theorem 3.1 and the fact that $M_n \longrightarrow \infty$ as $n \longrightarrow \infty$

$$E(|T_n - T'_n|^2) = E\left(\sum Y_k^2 + \sum Y'^2_k + \sum_{j \neq k} Y_j(Y_k - Y'_k) + \sum_{j \neq k} Y'_j(Y'_k - Y_k)\right) \dagger$$

$$\ll \sum m_j + \sum m'_j + \sum m_j \sum_{k \neq j} |m_{j-k} - m'_{j-k}| + \sum m'_j \sum_{k \neq j} |m'_{j-k} - m_{j-k}| \ddagger$$

$$\ll M_n + M_n \sum |m_k - m'_k|.$$

*The results in [8] are stated for the number of crossings of the values a, but in the case of an integer-valued random variable they can be easily translated into the number of a-values. ** i.o. stands for 'infinitely often' or 'for infinitely many values of the index.'

† Henceforth in an unspecified summation the index runs from 1 to n. ‡ $u_n \ll v_n$ means $u_n = O(v_n)$

According to Theorem 3.2, the m_k in the last sum can be divided into two classes: either $m_k \leq k^{-2}$, the sum over such k being $O(1)$, or the estimate (8) holds. Hence

$$E(|T_n - T'_n|^2) \leq O(M_n) + O(M_n \sum m_k k^{-(1-\varepsilon)/2}).$$

By Hölder's inequality

$$\sum m_k k^{-(1-\varepsilon)/2} \leq (\sum m_k^{2/(1+2\varepsilon)})^{1/2+\varepsilon} (\sum (k^{-(1-\varepsilon)/2})^{2/(1-2\varepsilon)})^{1/2-\varepsilon}$$

$$\leq A(\sum m_k^{2/(1+2\varepsilon)})^{1/2+\varepsilon} \leq AM_n^{1/2+\varepsilon} .$$

By Chebychev inequality

$$(10) \qquad P(|T_n - T'_n| > M_n^{3/4+\varepsilon}) \leq M_n^{-\varepsilon} .$$

Since $m_n \longrightarrow 0$ by Theorem 1, we can choose an increasing sequence n_k such that

$$M_{n_k} \sim k^{(1+\varepsilon)/\varepsilon} .$$

Now suppose that for some $n, n_k < n \leq n_{k+1}$ we have

$$(11) \qquad |T_n - T'_n| > 2M_n^{3/4+\varepsilon} .$$

Let n be the smallest such integer, for which (11) is true, then either Y_n or Y'_n must be 1, hence S_n=a or a'. We call this event E_n. According as S_n=a or a', $T_{n_{k+1}} - T_n$ is the number of 0's or (a-a')'s in the sequence of partial sums of $X_{n+1}, \cdots, X_{n_{k+1}}$. Let the event

$$|T_{n_{k+1}} - T'_{n_{k+1}} - (T_n - T'_n)| \leq M_{n_{k+1}}^{3/4+\varepsilon}$$

be denoted by $E_{n,n_{k+1}}$. By (10), if k is sufficiently large,

$$P(E_{n,n_{k+1}} | E_n) \geq 1 - M_{n_{k+1}}^{-\varepsilon} \geq \frac{1}{2} .$$

Further it is clear that

$$P(E_{n,n_{k+1}} | E'_{n_k} \cdots E'_{n-1} E_n)^* = P(E_{n,n_{k+1}} | E_n) \geq \frac{1}{2}$$

* If E,F are two events, E' denotes the negation of E, EF denotes the conjunction of E and F.

(this follows from the Markov property of the sequence S_n.) Now $E_{n,n_{k+1}}$ and $E'_{n_k}\cdots E'_{n-1}E_n$ together imply

$$|T_{n_{k+1}}-T'_{n_{k+1}}| > M_{n_{k+1}}^{3/4+\varepsilon}.$$

Hence

$$P(|T_{n_{k+1}}-T'_{n_{k+1}}| > M_{n_{k+1}}^{3/4+\varepsilon}) \geq \sum_{n=n_k}^{n_{k+1}} P(E'_{n_k}\cdots E'_{n-1}E_n)P(E_{n,n_{k+1}}|E'_{n_k}\cdots E'_{n-1}E_n)$$

$$\geq \frac{1}{2}\sum_{n=n_k}^{n_{k+1}} P(E'_{n_k}\cdots E'_{n-1}E_n) = \frac{1}{2} P(\max_{n_k\leq n\leq n_{k+1}} |T_n-T'_n|M_n^{-3/4-\varepsilon} > 2).$$

Thus by (10)

$$\sum_k P(\max_{n_k \leq n \leq n_{k+1}} |T_n-T'_n|M_n^{-3/4-\varepsilon} >2) \leq 2 \sum_k M_{n_{k+1}}^{-\varepsilon} < \infty.$$

It follows from the Borel-Cantelli lemma that

$$P(|T_n-T'_n| > 2M_n^{3/4+\varepsilon} \quad \text{i.o.}) = 0.$$

This is equivalent to the statement of Theorem 5.

The next theorem is a new type of limit theorem. The sequence of random variables $Y_1,Y_2,\ldots$ does not obey the usual law of large numbers in the sense that constants A_n do not exist so that with probability 1,

$$\lim_{n\to\infty} \frac{Y_1+\ldots+Y_n}{A_n} = 1.$$

By analogy with the situation for sums of independent random variables with finite first moments, we should expect to take A_n to be the M_n above. That this is not true is shown already in the simplest case of Bernoullian variables $X_1,X_2,\ldots$ where each $X_k = \pm 1$ each with probability 1/2. In this case $m_k \sim A\, k^{-1/2}$. $M_n \sim 2A_n^{1/2}$, but the sum $Y_1+\ldots+Y_n$ oscillates between $A'n^{1/2}(\log n)^{-1-\varepsilon}$ and $A''\, n^{1/2}(\log\log n)^{1/2}$ with probability 1 (see [6]). However we shall show in the next theorem that, in a certain sense, Y_k does behave like its expectation m_k, as follows.

THEOREM 6. Under (0),

$$P\Big(\lim_{n\to\infty} \frac{1}{\log M_n} \sum_{k=1}^{n} \frac{Y_k}{M_k} = 1\Big) = 1.$$

Notice that in this formula if we replace Y_k by m_k, the limit relation holds without the intervention of probability. If we regard $(Y_1+\ldots+Y_n)/M_n$ as a sort of 'arithmetical average,' the quantity

$$\frac{1}{\log M_n} \sum_{k=1}^{n} \frac{Y_k}{M_k}$$

may be called a 'logarithmic average.' Evidently the existence of the mathematical average implies the existence (and equality therewith) of the logarithmic. The first instance of considering such an average in probability is due to P. Levy [9], p.270.

Proof of Theorem 6. We have

$$E\Big(\sum \frac{Y_k}{M_k}\Big) = \sum \frac{m_k}{M_k} = \log M_n + O(1).$$

Next

$$E\Big(\Big(\sum \frac{Y_k}{M_k}\Big)^2\Big) = \sum \frac{m_k}{M_k^2} + 2 \sum_{j<k} \frac{m_j m_{k-j}}{M_j M_k}$$

$$= O(1) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \sum_{k=j+1}^{n} \frac{m_{k-j}}{M_k}.$$

Hence

$$0 \le E\Big(\Big(\sum \frac{Y_k}{M_k}\Big)^2\Big) - E^2\Big(\sum \frac{Y_k}{M_k}\Big)$$

$$\le O(1)+O\Big(\sum \frac{m_k}{M_k^2}\Big)^2 + 2 \sum_{j=1}^{n} \frac{m_j}{M_j}\Big(\sum_{k=j+1}^{n} \frac{m_{k-j}}{M_k} - \sum_{k=j+1}^{n} \frac{m_k}{M_k}\Big)$$

$$\le O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j}\left\{\sum_{k=1}^{n-j} \frac{m_k}{M_{k+j}} - \sum_{k=j+1}^{n} \frac{m_k}{M_k}\right\}$$

$$\leq O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \sum_{k=1}^{j} \frac{m_k}{M_{k+j}}$$

$$\leq O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} = O(\log M_n).$$

By Chebychev inequality

$$P(\left| \sum \frac{Y_k}{M_k} - \log M_n \right| > \varepsilon \log M_n) \leq O(\frac{1}{\log M_n}).$$

Choose an increasing sequence n_k such that

$$M_{n_k} \sim e^{k^2}.$$

By the Borel-Cantelli lemma,

$$P(\lim_{k \to \infty} \frac{1}{\log M_{n_k}} \sum_{i=1}^{n_k} \frac{Y_i}{M_i} = 1) = 1.$$

Now if $n_k \leq n \leq n_{k+1}$,

$$\frac{1}{\log M_{n_{k+1}}} \sum_{k=1}^{n_k} \frac{Y_i}{M_i} \leq \frac{1}{\log n} \sum_{i=1}^{n} \frac{Y_i}{M_i}$$

$$\leq \frac{1}{\log M_{n_k}} \sum_{i=1}^{n_{k+1}} \frac{Y_i}{M_i} .$$

Since $\log M_{n_{k+1}} / \log M_{n_k} \to 1$ as $k \to \infty$ the extreme sides of these inequalities $\to 1$ with probability 1, by what has just been proved. Theorem 6 follows.

THEOREM 7. _Under_ (0), _for every_ $\varepsilon > 0$

$$P(M_n^{1-\varepsilon} < T_n < M_n^{1+\varepsilon} \text{ for all sufficiently large } n) = 1.$$

Proof. This is equivalent to the following two statements:

(1) $$P(T_n > M_n^{1+\varepsilon} \text{ i.o.}) = 0$$

(2) $$P(T_n < M_n^{1-\varepsilon} \text{ i.o.}) = 0.$$

The proof of (1) is similar to that of Theorem 5 and will be omitted. To prove (2), we choose $v=v(n)$ such that $M_v \sim M_n^{1-\varepsilon}$; this is possible because $m_n \longrightarrow 0$ and $M_n \uparrow \infty$. From Theorem 6 we have, with probability 1,

$$\lim_{n\longrightarrow} \frac{1}{\log M_n} \sum_{k=1}^{v} \frac{Y_k}{M_k} = 1 - \varepsilon$$

$$\lim_{n\longrightarrow\infty} \frac{1}{\log M_n} \sum_{k=v+1}^{n} \frac{Y_k}{M_k} = \varepsilon.$$

Upon subtraction it follows that

$$\lim_{n\longrightarrow\infty} \frac{1}{\log M_n} \frac{T_n - T_v}{M_v} \geq \varepsilon$$

or

$$\lim_{n\longrightarrow\infty} \frac{T_n}{M_n^{1-\varepsilon} \log M_n} \geq \varepsilon.$$

This is equivalent to (2).

Remark. Part (2) of Theorem 7 would also have followed from a general theorem of Feller (Theorem 2 in [10]), but for the condition (13) there. To verify this condition (or rather a slightly weaker one) it would be sufficient to show that

$$M_{2n} \leq 2M_n.$$

We are unable to prove or disprove this relation.

THEOREM 8. Under (0)

$$P\left(\lim_{n\longrightarrow\infty} \frac{T_n}{T_n'} = 1\right) = 1.$$

Proof. This is an immediate consequence of Theorems 5 and 7. Actually we have even, for every $\varepsilon > 0$

$$P\left(\left|\frac{T_n - T_n'}{T_n}\right| > M_n^{-(1+\varepsilon)/4} \quad \text{i.o.}\right) = 0.$$

We are indebted to Dr. Miriam Lipschutz for several corrections on the MS.

Cornell University
University of Illinois

BIBLIOGRAPHY

[1] B. V. Gnedenko. (a) On a local limit theorem of the theory of probability. (In Russian) Uspehi Matem. Nauk. (N.S.) 3 (1948), pp. 187-197; (b) On a local theorem for the region of normal attraction of stable laws. (In Russian) Doklady Akad. Nauk. SSSR (N.S.) 66(1949), pp. 325-326.

[2] E. R. van Kampen and A. Wintner. A limit theorem for probability distributions in lattices. Amer. J. Math. 61 (1939), pp. 965-973.

[3] C. G. Esseen. Fourier analysis of distribution functions. Acta Math. 77 (1945), pp. 1-125.

[4] K. L. Chung and W. H. J. Fuchs. On the distribution of values of sums of random variables.

[5] W. Feller. Note on the law of large numbers and 'fair' games. Ann. Math. Stat. 16 (1945), pp. 301-304.

[6] K. L. Chung and G. A. Hunt. On the zeros of $\sum_{1}^{n} \pm 1$. Ann. Math. 50(1949), pp. 385-400.

[7] W. Feller. Fluctuation theory of recurrent events. Trans. Amer. Math. Soc. 67(1949), pp. 98-119.

[8] K. L. Chung. Fluctuations of sums of independent random variables. To appear in Ann. Math.

[9] P. Lévy. Théorie de l'addition des variables aléatoires. Paris, 1937.

[10] W. Feller. A limit theorem for random variables with infinite moments. Amer. J. Math. 68(1946), pp. 257-262.

[11] W. Doeblin. Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables, Bull. Soc. Math. de France 66(1938), pp. 210-220.