ANNALS OF MATHEMATICS Vol. 48, No. 4, October, 1947

## ON THE LOWER LIMIT OF SUMS OF INDEPENDENT RANDOM VARIABLES

# By Kai-Lai Chung and Paul Erdös (Received October 30, 1946)

**1.** Let  $X_1, X_2, \dots, X_n \dots$  be independent random variables and let  $S_n = \sum_{r=1}^n X_r$ . In the so-called law of the iterated logarithm, completely solved by Feller recently, the upper limit of  $S_n$  as  $n \to \infty$  is considered and its true order of magnitude is found with probability one. A counterpart to that problem is to consider the lower limit of  $S_n$  as  $n \to \infty$  and to make a statement about its order of magnitude with probability one.

THEOREM 1. Let  $X_1, \dots, X_n, \dots$  be independent random variables with the common distribution:  $\Pr(X_n = 1) = p$ ,  $\Pr(X_n = 0) = 1 - p = q$ . Let  $\psi(n) \downarrow \infty$  and

(1.1) 
$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty,$$

Then we have

(1.2) 
$$\Pr\left(\lim_{n \to \infty} n^{1/2} \psi(n) \mid S_n - np \mid = 0\right) = 1.$$

Theorem 1 is a best possible theorem. In fact we shall prove the following THEOREM 2. Let  $X_n$  be as in Theorem 1 but let p be a quadratic irrational. Let  $\phi(n) \uparrow \infty$  and

(1.3) 
$$\sum_{n=1}^{\infty} \frac{1}{n\phi(n)} < \infty$$

Then we have

(1.4) 
$$\Pr\left(\lim_{n \to \infty} n^{1/2} \phi(n) \mid S_n - np \mid = 0\right) = 0.$$

By making use of results on uniform distribution mod 1 we can prove (1.4) for almost all p, however the proof is omitted here.

In order to extend the theorem to more general sequences of random variables, we need a theorem about the limiting distribution of  $S_n$  with an estimate of the accuracy of approximation. Cramér's asymptotic expansion is suitable for this purpose. The conditions on F(x) in the following Theorem 3 are those under which the desired expansion holds.

THEOREM 3. Let  $X_1, \dots, X_n, \dots$  be independent random variables having the same distribution function F(x). Suppose that the absolutely continuous part of F(x) does not vanish identically and that its first moment is zero, the second is one, and the absolute fifth is finite. Let  $\psi(n)$  be as in Theorem 1, then

(1.5) 
$$\Pr\left(\lim_{n \to \infty} n^{1/2} \psi(n) \mid S_n \mid = 0\right) = 1,$$

CHUNG AND ERDÖS

On the other hand, let  $\phi(n)$  be as in Theorem 2: then

(1.6) 
$$\Pr\left(\lim_{n\to\infty} n^{1/2} \phi(n) \mid S_n \mid = 0\right) = 0.$$

It seems clear that the result can be extended to other cases, however we shall at present content ourselves with this statement.

2. For x > 0 let I(x) denote the integer nearest to x if x is not equal to [x] $+\frac{1}{2}$ ; in the latter case, let I(x) = [x]; let |x| = x - I(x). We have then for any x > 0, y > 0, the inequality

 $|\{x - y\}| \le |\{x\} - \{y\}|,$ 

We are now going to state and prove some lemmas. The first two lemmas are number-theoretic in nature; the third one supplies the main probability argument; and the fourth one is a form of zero-or-one law.

LEMMA 1. Let p > 0 be a real number. Let  $\psi(n) \uparrow \infty$ . Arrange all the positive integers n for which we have.

 $|\{np\}| < cn^{-1/2} \psi(n)^{-1}$ (2.1)

in an increasing sequence  $n_i$ ,  $i = 1, 2, \dots$ . Then for any pair of positive integers i and k we have

$$n_{i+2k} \ge n_i + n_k$$

PROOF. Suppose the contrary:

 $n_{i+n+} < n_i + n_i$ 

CASE (i):  $k \leq i$ . Consider the 2k + 1 numbers

ni , ni+1 , ..., ni+2k

and the corresponding

 $\{n_ip\}, \{n_{i+1}p\}, \cdots, \{n_{i+2k}p\}.$ (2.2)

There are at least k + 1 numbers among (2.2) which are of the same sign: without loss of generality we may assume that they are non-negative. Let the corresponding  $n_i$  be

 $n_{i_1} < n_{i_2} < \cdots < n_{i_{k+1}}$ 

Then we have

 $0 \leq \{n_{i_j} p\} < c n_{i_j}^{-1/2} \psi(n_j)^{-1} \leq c n_k^{-1/2} \psi(n_k)^{-1}, \quad j = 1, \cdots, k+1;$ 

since  $i_i \geq i \geq k$ ; and

$$|\{n_{i_{k+1}}p - n_{i_{j}}p\}| < cn_{k}^{-1/2}\psi(n_{k})^{-1}, \quad j = 1, \cdots, k;$$
  
$$0 < n_{i_{k+1}} - n_{i_{j}} \le n_{i+2k} - n_{i} < n_{k}.$$

#### SUMS OF INDEPENDENT VARIABLES

Thus there would be k different positive integers  $n_{i_{k+1}} - n_{i_j}$ ,  $j = 1, \dots, k$ all  $< n_k$ , for which

$$\{np\} \mid < cn_k^{-1/2} \psi(n_k)^{-1}.$$

This is a contradiction to the definition of  $n_k$ .

CASE (ii) k > i. Consider the i + k + 1 numbers

$$n_k$$
,  $n_{k+1}$ , ...,  $n_{i+2k}$ 

and the corresponding

$$\{n_kp\}, \{n_{k+1}p\}, \cdots, \{n_{i+2k}p\}.$$

Since i + k + 1 > 2i + 1, there are at least i + 1 of the numbers above which are of the same sign, say non-negative. Let the corresponding  $n_i$  be

$$n_{k_1} < n_{k_2} < \cdots < n_{k_{i+1}}$$

By an argument similar to that in Case (i) we should have *i* numbers  $n_{k_i+1} - n_{k_i}, j = 1, \dots, i$ , all  $< n_i$  for which

$$|\{np\}| < cn_i^{-1/2} \psi(n_i)^{-1}.$$

This leads to a contradiction as before.

LEMMA 2. Let  $n_i$  be defined as in Lemma 1. Then if

(2.3) 
$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty,$$

we have

(2.4) 
$$\sum_{i=1}^{\infty} n_i^{-1/2} =$$

PROOF. Consider the points

$$hcn^{-1/2}\psi(n)^{-1}$$
  $h = \pm 1, \cdots, \pm [2^{-1}c^{-1}n^{1/2}\psi(n)],$ 

00

They divide the interval  $(-\frac{1}{2}, \frac{1}{2})$  into at most  $[c^{-1}n^{1/2}\psi(n)] + 2$  parts. Hence at least one subinterval contains

$$l \ge \frac{n}{[c^{-1}n^{1/2}\psi(n)] + 2}$$

members of the *n* numbers  $\{mp\}, m = 1, 2, \dots, n$ . Let the corresponding  $n_i$  be

 $n_1 < n_2 < \cdots < n_l.$ 

Then

$$0 < |\{n_{i}p - n_{i}p\}| < cn^{-1/2} \psi(n)^{-1} < c(n_{i} - n_{i})^{-1/2} \psi(n_{i} - n_{i})^{-1},$$
  
$$i = 1, \cdots, l - l$$

1005

1.

Hence if g(n) denote the number of numbers among 1,  $\cdots$ , n for which

 $|\{np\}| < cn^{-1/2}\psi(n)^{-1},$ 

we have, for n sufficiently large

$$q(n) > 2^{-1} c n^{1/2} \psi(n)^{-1}$$

Now

$$\sum_{\frac{2^{k-1} \leq n \leq 2^k}{2^k}} n_i^{-1/2} \ge \frac{g(2^k) - g(2^{k-1})}{\sqrt{2^k}}$$

Hence

$$\begin{split} \sum_{k=1}^{\infty} \sum_{2^{k-1} < \tau_i \le 2^k} n_i^{-1/2} &\geq \sum_{k=1}^{\infty} \frac{g(2^k) - g(2^{k-1})}{\sqrt{2^k}} \\ &= -\frac{g(1)}{\sqrt{2}} + \sum_{k=1}^{\infty} g(2^k) \left(\frac{1}{\sqrt{2^k}} - \frac{1}{\sqrt{2^{k+1}}}\right) \\ &\geq -\frac{g(1)}{\sqrt{2}} + \sum_{k=1}^{\infty} \frac{c}{2} \frac{\sqrt{2^k}}{\psi(2^k)} \left(\frac{1}{\sqrt{2^k}} - \frac{1}{\sqrt{2^{k+1}}}\right) \\ &\geq -1 + \frac{c}{2} \left(1 - \frac{1}{\sqrt{2}}\right) \sum_{k=1}^{\infty} \frac{1}{\psi(2^k)} \,. \end{split}$$

It is well-known<sup>1</sup> that if (2.3) holds then

$$\sum_{k=1}^{\infty} \frac{1}{\psi(2^k)} = \infty.$$

Thus (2.4) is proved.

LEMMA 3. Let  $n_i$ ,  $i = 1, 2, \cdots$  be a monotone increasing sequence such that for any pair of positive integers i and k we have

 $(2.5) n_{i+2k} \ge n_i + n_k$ 

and

(2.6) 
$$\sum_{i=1}^{\infty} n_i^{-1/2} = \infty$$

Then if  $\alpha$  and  $\beta$  are two integers, we have for any integer h > 0,

(2.7) 
$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } i \ge h) \ge \frac{1}{2}.$$

<sup>1</sup>See e. g. Theory and Application of Infinite Series, London-Glasgow, Blackie and Son, 1928, p. 120.

**PROOF.** Denoting the joint probability of  $E_1, E_2, \cdots$  by  $\Pr(E_1; E_2; \cdots)$ we have

$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta) = \Pr(S_{n_k} = I(pn_k + p\alpha) + \beta;$$

$$S_{n_i - n_k} = I(pn_i + p\alpha) - I(pn_k + p\alpha))$$

$$+ \Pr(S_{n_k} \neq I(pn_k + p\alpha) + \beta; S_{n_{k+1}} = I(pn_{k+1} + p\alpha) + \beta;$$

$$S_{n_i - n_{k+1}} = I(pn_i + p\alpha) - I(pn_{k+1} + p\alpha))$$

$$+ \cdots$$

$$+ \Pr(S_{n_k} \neq I(pn_k + p\alpha) + \beta; \cdots; S_{n_{i-1}} \neq I(pn_{i-1} + p\alpha) + \beta;$$

$$S_{n_i} = I(pn_i + p\alpha) + \beta.$$

Writing

 $p_i = \Pr(S_{n_i} = I(pn_i + p\alpha) + \beta),$  $w_k = \Pr(S_{n_i} \neq I(pn_i + p\alpha) + \beta \text{ for } h \leq j < k; S_{n_k} = I(pn_k + p\alpha) + \beta).$  $p_{k,i} = \Pr(S_{n_k - n_k} = I(pn_i + p\alpha) - I(pn_k + p\alpha)), \quad p_{k,k} = 1;$ 

and using the assumption of independence, we have

$$p_i = \sum_{k=h}^i w_k p_{k,i}.$$

Summing from h to m we get

(2.8) 
$$\sum_{i=k}^{m} p_i = \sum_{i=k}^{m} \sum_{k=1}^{i} w_k p_{k,i} \leq \sum_{k=1}^{m} w_k \sum_{i=k}^{m} p_{k,i}.$$

Now for any positive x and y, I(x) - I(y) = I(x - y) or  $I(x - y) \pm 1$ ; and it is well-known that for the random variables we have, given any  $\epsilon > 0$ , if  $n > n_0(\epsilon)$ , and  $\theta = \pm 1$ ,

$$\Pr(S_n = I(np) + \theta) \le (1 + \epsilon)\Pr(S_n = I(np))$$

hence we have, if  $i - k \ge m_1(\epsilon)$ ,

(2.9) 
$$p_{k,i} \leq (1 + \epsilon/4) \Pr(S_{n_i - n_k} = I(pn_i - pn_k)).$$

From (2.5) if i > k, we have

$$(2.10) n_i \ge n_k + n_{\{(i-k)/2\}}.$$

Also it is well-known that as  $i \to \infty$ ,

$$(2.11) p_i \sim \frac{1}{\sqrt{2\pi pqn_i}}.$$

Hence from (2.9), (2.10) and (2.11) we have if  $i - k \ge m_2(\epsilon)$  where  $m_2$  is a positive constant,

$$p_{k,i} \leq (1 + \epsilon/2) \Pr(S_{n_{\{(i-k)/2\}}} = I(pn_{\{(i-k)/2\}})).$$

Since  $\alpha$  and  $\beta$  are fixed, to any  $\epsilon > 0$  there exists an integer  $m_0 = m_0(\epsilon) > m_2$ such that if  $i - k \ge m_0(\epsilon)$ ,

(2.12)  $\Pr(S_{n_1(i-k)/21} = I(pn_{1(i-k)/21})) \leq (1 + \epsilon)p_{1(i-k)/21}.$ 

Thus for  $i - k \ge m_0(\epsilon)$ ,

 $p_{k,i} \leq (1 + \epsilon) p_{[(i-k)/2]}.$ 

Using (2.12) in (2.9), we obtain

$$\sum_{i=h}^{m} p_i \leq \sum_{k=1}^{m} w_i \left( \sum_{i=h}^{k+m_0-1} p_{j,i} + (1+\epsilon) \sum_{i=k+m_0}^{m} p_{\lfloor (i-k)/2 \rfloor} \right)$$
$$\leq \sum_{k=1}^{m} w_i \left( m_0 + 2(1+\epsilon) \sum_{i=m_0}^{\lfloor m/2 \rfloor} p_i \right).$$

Therefore

$$\sum_{i=1}^{m} w_i \ge \frac{\sum_{i=1}^{m} p_i}{m_0 + 2(1+\epsilon) \sum_{i=m_0}^{[m/2]} p_i}$$

Since by (2.11) and (2.6) the series  $\sum_{i=1}^{\infty} p_i$  is divergent, we get, letting  $n \to \infty$ ,

$$\sum_{i=1}^{\infty} w_i \ge \frac{1}{2(1+\epsilon)}$$

Since  $\epsilon$  is arbitrary and the left-hand side does not depend on  $\epsilon$  this proves (2.7).

LEMMA 4. If for any integers  $\alpha$ ,  $\beta$  and k > 0, there exists a number  $\eta > 0$  not depending on  $\alpha$ ,  $\beta$  and an integer  $l = l(k, \eta)$  such that, n; being any sequence  $\uparrow \infty$ ,

(2.13) 
$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } k \leq i \leq l) \geq \eta;$$

then

(2.14) 
$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ infinitely often}) = 1.$$

**PROOF.** Take a sequence  $k_1, k_2, \cdots$  and the corresponding  $l_1, l_2, \cdots$  such that

 $k_1 < l_1 < k_2 < l_2 < \cdots$ 

Consider the event

 $\mathbf{E}_r: \qquad S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } k_r \leq i \leq l_r,$ 

and let the probability that  $E_r$  occurs under the hypothesis that none of  $E_1$ ,  $\cdots$ ,  $E_{r-1}$  occurs, be denoted by  $\Pr(E_r | E'_1 \cdots E'_{r-1})$ . Then the latter is a probability mean of the conditional probabilities of  $E_r$  under the various hypotheses:

 $H: \qquad S_{n_i} = \sigma_{n_i}, \qquad k_t \leq i \leq l_t, \qquad 1 \leq t \leq r-1;$ 

#### SUMS OF INDEPENDENT VARIABLES

where the  $\sigma_{n_i}$ 's are such that for all i,  $\sigma_{n_i} \neq I(pn_i + p\alpha) + \beta$  but are otherwise arbitrary. Now under H, E, will occur if the following event F occurs:

F:  $S_{n_i-n_{l_{r-1}}} = I(pn_i + p\alpha) + \beta - \sigma_{n_{l_{r-1}}}$  at least once for  $k_r \leq i \leq l_r$ . Hence

$$\Pr(E_r \mid E'_1 \cdots E'_{r-1}) \ge \min \Pr(E_r \mid H) \ge \Pr(F \mid H) = \Pr(F).$$

Writing the equality in F as

$$S_{n_i - n_{l_{r-1}}} = I(p(n_i - n_{l_{r-1}}) + p(n_{l_{r-1}} + \alpha)) + \beta - \sigma_{n_{l_{r-1}}}$$
  
=  $I(p(n_i - n_{l_{r-1}}) + p\alpha') + \beta'$ 

and consider the random variables  $X_{n_{l_{r+1}}}$ ,  $X_{n_{l_{r+1}+1}}$ ,  $\cdots$  as  $X'_1$ ,  $X'_2$ ,  $\cdots$  we see from (2.13) that

$$\Pr(E_r \mid E'_1 \cdots E'_{r-1}) \geq \Pr(F) \geq \eta.$$

Therefore the probability that none of the events  $E_r$ ,  $r = 1, \dots, s$  occurs is  $\Pr(E'_1 \dots E'_s) = \Pr(E'_1)\Pr(E'_2 | E'_1) \dots \Pr(E'_s | E'_1 \dots E'_{s-1}) \leq (1 - \eta)^s$ . Hence

$$\Pr(S_{n_i} \neq I(pn_i + p\alpha) + \beta \text{ for all } l_r \leq i \leq k_r, r = 1, 2, \dots) = 0$$

Since  $l_1$  can be taken arbitrarily large, (2.14) is proved.

**REMARK.** Lemma 3 and 4 imply an interesting improvement of the wellknown fact that  $Pr(S_n - np = \text{infinitely often}) = 1$  for a rational p. Let  $n_i$  be any monotone increasing sequence such that (2.6) holds; in addition if for a ce tain integer m > 0 and any pair of integers i and k we have

$$(2.15) n_{i+mk} \ge n_i + n_k$$

then

 $\Pr(S_{n_i} - n_i p = 0 \text{ for infinitely many } i) = 1.$ 

That the condition (2.6) alone is not sufficient can be shown by a counterexample. On the other hand, it is trivial that (2.6) is a necessary condition. The condition (2.15) can be replaced e.g. by the following condition:

$$n_{i+1} - n_i \ge A n_i^{1/2}$$
,  $A > 0$ .

The proof is different and will be omitted here.

PROOF OF THEOREM 1. Let the sequence  $n_i$  be defined as in Lemma 1. Then by Lemma 1 and 2 this sequence satisfies the conditions (2.5) and (2.6) in Lemma 3. Hence by Lemma 3 the condition (2.13) in Lemma 4 is satisfied with any  $\eta < \frac{1}{2}$ . Thus by Lemma 4 we have (2.14). Taking  $\alpha = \beta = 0$  therein we obtain

$$Pr(S_{n_i} - n_i p = \{n_i p\} \text{ infinitely often}) = 1.$$

Hence by the definition (2.2)

 $\Pr(|S_n - np| < cn^{-1/2} \psi(n)^{-1} \text{ infinitely often}) = 1.$ 

Since c is arbitrarily small (1.2) is proved.

REMARK. It is clear that (2.14) yields more than Theorem 1 since  $\alpha$  and  $\beta$  are arbitrary. It is easily seen that we may even make  $\alpha$  and  $\beta$  vary with  $n_i$  in a certain way, but we shall omit these considerations here.

**PROOF OF THEOREM 2.** Arrange all the positive integers n for which we have

$$|\{np\}| \leq An^{-1/2}\phi(n)^{-1}, \qquad A > 0.$$

in an ascending sequence  $n_i$ ,  $i = 1, 2, \cdots$ . Since

$$|\{n_i p\}| \leq A n_i^{-1/2} \phi(n_i)^{-1}$$

we have

$$(2.16) \qquad |\{n_{i+1}p - n_ip\}| \leq 2A n_i^{-1/2} \phi(n_i)^{-1}.$$

On the other hand, since p is a quadratic irrational, it is well-known<sup>2</sup> that there exists a number M > 0 such that

(2.17) 
$$|\{n_{i+1}p - n_ip\}| > \frac{M}{n_{i+1} - n_i}.$$

From (2.16) and (2.17) we get with  $A_1 = M/2A$ ,

 $(2.18) n_{i+1} - n_i > A_1 n_i^{1/2} \phi(n_i)$ 

Without loss of generality we may assume that  $\phi(n_i) \leq n_i^{1/2}$ . For we may replace  $\phi(n)$  by  $\phi_1(n)$  defined as follows:

 $\phi_1(n) = \begin{cases} \phi(n) & \text{if } \phi(n) \leq n^{1/2}; \\ n^{1/2} & \text{if } \phi(n) > n^{1/2}. \end{cases}$ 

After this replacement (1.3) remains convergent, while if (1.4) holds for  $\phi_1(n)$ , it holds a fortiori for  $\phi(n)$ .

Now if  $\phi(n_i) \leq n_i^{1/2}$ , and the constant  $A_2$  is such that  $2A_2 + A_2^2 < A_1$ , we have from (2.18)

$$n_{i+1}^{1/2} > n_i^{1/2} + A_2 \phi(n_i).$$

Hence by iterating,

$$a_{i+1}^{1/2} > A_2 \sum_{k=1}^{i} \phi(n_k) > A_2 \sum_{k=\lfloor i/2 \rfloor}^{i} \phi(n_k) > A_2 \frac{i}{2} \phi\left(\left[\frac{i}{2}\right]\right).$$

Therefore by (1.3)

(2.19) 
$$\sum_{i=1}^{\infty} n_i^{-1/2} < \infty.$$

<sup>2</sup>See e. g. HARDY AND WRIGHT, Introduction to the Theory of Numbers, Oxford 1938, p. 157.

Define

$$p_i = \Pr(S_{n_i} = I(pn_i)).$$

As in (2.11) we have

$$p_i \sim \frac{1}{\sqrt{2\pi pqn_i}}.$$

Hence from (2.18)

$$\sum_{i=1}^{\infty} \, p_i \, < \, \infty \, .$$

By the classical Borel-Cantelli lemma it follows that

$$Pr(S_{n_i} = I(pn_i) \text{ infinitely often}) = 0$$

By the definition of  $n_i$  this is equivalent to (1.4).

**3.** LEMMA 5. Let  $X_1, \dots, X_n, \dots$  be independent random variables having the same distribution function F(x) which satisfies the conditions in Theorem 3. Then if  $x_1 < x_2$  and  $x_1 = o(1), x_2 = o(1)$  as  $n \to \infty$ , we have

(3.1) 
$$\Pr(x_1 \le n^{-1/2} S_n \le x_2) = (2\pi)^{-1/2} (x_2 - x_1) + o(x_2 - x_1) + 0(n^{-3/2})$$

**PROOF.** By Cramér's asymptotic expansion<sup>3</sup> we have, if we denote the r<sup>th</sup> moment of F(x) by  $\alpha_r$ ,

$$\Pr\left(\frac{S_n}{\sqrt{n}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} \, dy - \frac{\alpha_3}{6\sqrt{2\pi}\sqrt{n}} (x^2 - 1) \, e^{-x^2/2} \\ + \frac{\alpha_4 - 3\alpha_2^2}{24\sqrt{2\pi}n} \, (-x^3 + 3x) e^{-x^2/2} + \frac{\alpha_3^2}{72\sqrt{2\pi}n} \, (-x^5 + 10x^3 - 15x) e^{-x^2/2} + R(x)$$

where

 $|R(x)| \leq Qn^{-3/2},$ 

and Q is a constant depending only on F(x),

It follows, using elementary estimates, that

$$\Pr\left(x_{1} \leq \frac{S_{n}}{\sqrt{n}} \leq x_{2}\right) = \frac{1}{\sqrt{2\pi}} \int_{x_{1}}^{x_{2}} e^{-y^{2}/2} dy + 0 \left((x_{2} - x_{1})\left(\frac{|x_{1}| + |x_{2}|}{\sqrt{n}} + \frac{1}{n}\right)\right) + 0 \left(\frac{1}{\sqrt{n^{3}}}\right)$$

Since  $x_1 = o(1)$ ,  $x_2 = o(1)$  this reduces immediately to (3.1).

<sup>\*</sup>CRAMÉR, Random Variables and Probability Distributions, Cambridge 1937, Ch. 7. For a simplified proof see P. L. Hsu, The Approximate Distribution of the Mean and Variance of a Sample of Independent Variables, Ann. Math. Statistics, 16 (1945), pp. 1-29.

#### CHUNG AND ERDÖS

LEMMA 6. Let  $z_n$  be any real number such that  $z_n = O(n^{1/2})$ , c any positive number, and h any positive integer. Let  $\psi(n) \uparrow \infty$  and

(3.2) 
$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty,$$

Then if the random variables  $X_n$  satisfy the conditions of Theorem 3, we have

(3.3) 
$$\Pr(|S_n - z_n| \le cn^{-1/2} \psi(n)^{-1} \text{ at least once for } n \ge h) = 1$$

PROOF. Write

$$P_{n} = \Pr\left(|S_{n} - z_{n}| \leq cn^{-1/2} \psi(n)^{-1}\right);$$
  

$$W_{k} = \Pr\left(|S_{i} - z_{j}| > cj^{-1/2} \psi(j)^{-1} \text{ for } h \leq j < k; |S_{k} - z_{k}| \leq ck^{-1/2} \psi(k)^{-1}\right)$$
  

$$P_{k,n} = \Pr\left(|S_{n} - z_{n}| \leq cn^{-1/2} \psi(n)^{-1}| |S_{j} - z_{j}| > cj^{-1/2} \psi(j)^{-1} \text{ for } h \leq j < k; |S_{k} - z_{k}| \leq ck^{-1/2} \psi(k)^{-1}\right)$$

Then by a similar argument as in Lemma 3, we have

(3.4) 
$$\sum_{n=k}^{m} P_n \leq \sum_{k=k}^{m} W_k \sum_{n=k}^{m} P_{k,n}$$

Our next step is to show that to any  $\epsilon > 0$  there exists a constant  $A(\epsilon)$  such that for n - k > A, we have

$$(3.5) P_{k,n} \leq (1 + \epsilon)P_{n-k}.$$

To prove this we divide the x-interval  $|x - z_k| \leq ck^{-1/2}\psi(k)^{-1}$  into disjoint subintervals  $I_j$ ; of lengths  $\leq \epsilon' cn^{-1/2}\psi(n)^{-1}$  where  $\epsilon' > 0$  is arbitrary. If we write

$$P_{k,n}^{(i)} = \Pr\left( |S_n - z_n| \le c n^{-1/2} \psi(n)^{-1} |S_k - z_k \subset I_j \right)$$

we have

$$P_{k,n}^{(i)} \leq \Pr(S_n - S_k \subset I'_i)$$

where  $I'_i$  is an interval of lengths  $\leq (2 + \epsilon')cn^{-1/2}\psi(n)^{-1} \leq (2 + \epsilon')c(n - k)^{-1/2}$  $\psi(n - k)^{-1}$  lying within the interval  $|x - z_n + z_k| \leq cn^{-1/2}\psi(n)^{-1} + ck^{-1/2}\psi(k)^{-1}$ . From Lemma 5 it is seen that if  $n - k \geq A_1(\epsilon')$ ,

$$P_{\mathbf{k},\mathbf{m}}^{(j)} \leq \frac{2(1+\epsilon')c}{\sqrt{2\pi} \ (n-k)\psi(n-k)} \ ;$$

since  $P_{k,n}$  is a probability mean of  $P_{k,n}^{(j)}$ , we have

(3.6) 
$$P_{k,n} \leq \max_{j} P_{k,n}^{(j)} \leq \frac{2(1+\epsilon')c}{\sqrt{2\pi} (n-k)\psi(n-k)}$$

On the other hand, we have again from Lemma 5, if  $n - k \ge A_2(\epsilon')$ ,

(3.7) 
$$P_{n-k} \ge \frac{2(1-\epsilon')}{\sqrt{2\pi} (n-k)\psi(n-k)}.$$

From (3.6) and (3.7) follows (3.5).

Using (3.5) in (3.4) we get

$$\sum_{n=h}^{m} P_n \leq \sum_{k=h}^{m} W_k \left( \sum_{n=k}^{k+A-1} P_{k,n} + (1+\epsilon) \sum_{n=k+A}^{m} P_{n-k} \right)$$
$$\leq \sum_{k=h}^{m} W_k (A + (1+\epsilon) \sum_{n=A}^{m} P_n)$$
$$\sum_{k=h}^{m} W_k \geq \frac{\sum_{n=h}^{m} P_n}{A + (1+\epsilon) \sum_{n=A}^{m} P_n}$$

Now  $\sum_{n=k}^{\infty} P_n = \infty$  by (3.7) and (3.1). It follows from (3.8) by letting  $n \to \infty$  that

$$\sum_{k=h}^{\infty} W_k \ge \frac{1}{1+\epsilon}.$$

Since  $\epsilon$  is arbitrary and the left-hand side does not depend on  $\epsilon$  we have

$$(3.9) \qquad \qquad \sum_{k=k}^{\infty} W_k \ge 1.$$

Thus (3.3) follows.

(3.8)

**PROOF OF THEOREM 3.** Taking z = 0 in (3.9) and denoting by  $E_{h}$  the event  $|S_{n}| \leq c n^{-1/2} \psi(n)^{-1}$ ,

we can write (3.9) as follows:

$$\Pr\left(\sum_{n=k}^{\infty} E_n\right) = 1,$$

where the sign  $\sum$  denotes disjunction of events. Now the event which consists in the realization of an infinite number of the  $E_n$ 's can be written as

$$\prod_{h=1}^{\infty} \left( \sum_{n=h}^{\infty} E_n \right)$$

where the sign II denotes conjunction of events. Hence

$$\Pr\left(\prod_{h=1}^{\infty} \left(\sum_{n=h}^{\infty} E_n\right)\right) = \lim_{h \to \infty} \left(\sum_{n=h}^{\infty} E_n\right) = 1.$$

Thus (1.5) is proved. The proof of (1.6) follows immediately from Lemma 5 and Borel-Cantelli lemma.

PRINCETON UNIVERSITY