ON FLUCTUATIONS OF SUMS OF RANDOM VARIABLES¹

LOUIS J. COTE

Introduction. These results are generalizations of a theorem of Chung [1] on the limiting distribution of the number of crossings of a value, c, by the successive partial sums of a sequence of independent random variables, and of theorems by Chung and Erdös [3] on the lower limits of such sums. Two of these theorems concern the case of independent, equidistributed random variables whose distributions have an absolutely continuous component (Case A). The others are for binomial variates (Case B). We extend the results of Case A to sums of independent random variables whose distributions need not be the same, and those of Case B to sums of independent and equidistributed random variables of the lattice type.

Let $\{X_n\}$ be a sequence of independent random variables whose c.d.f.'s, $\{F_n(x)\}$, need not be the same. We use the usual notations:

$$EX_{i}^{k} = \alpha_{ki}, \qquad E |X_{i}|^{k} = \beta_{ki}, \qquad E(X_{i} - \alpha_{1i})^{2} = \sigma_{i}^{2},$$

$$\sum_{i=1}^{n} X_{i} = S_{n}, \qquad \sum_{i=1}^{n} \sigma_{i}^{2} = S_{n}^{2},$$

$$P\{S_{n} \leq x\} = \overline{F}_{n}(x) = F_{1} * F_{2} * \cdots * F_{n}(x).$$

The crossings of a line, Case A. For a sequence, $\{c_n\}$, we say that S_n crosses $\{c_n\}$ from above at k if $S_k > -c_k$, $S_{k+1} < -c_k$. The number of such crossings for $1 \le k \le n$ is a random variable T_n . We have the following theorem on its limiting distribution.

THEOREM 1. Let $\{X_n\}$ be a sequence of random variables whose means are zero and which satisfy the conditions:

(1) For the resolution $F_n(x) = g_n G_n(x) + (1 - g_n) G'_n(x)$, where $G_n(x)$ is absolutely continuous and $G'_n(x)$ has no absolutely continuous component.

(a)
$$\frac{dG_n(x)}{dx} \text{ is of bounded total variation, } v_n, \text{ in } (-\infty, \infty),$$

(b)
$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^n \frac{g_k}{1+v_k^2} = \infty;$$

Received by the editors June 1, 1954.

¹ These results are part of a dissertation submitted in partial fulfillment of requirements for a Ph.D. at Columbia University. The research was supported by the O. N. R.

(2)
$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=k+1}^{k+n} \beta_{ri} = B_r \quad uniformly in \ k \ (r=1, 2),$$

$$\frac{1}{n} \sum_{i=k+1}^{k+n} \beta_{4i} < M \qquad uniformly in k,$$

and let $\{c_n\}$ be a sequence of real numbers satisfying

(4)
$$\lim_{n\to\infty}\frac{c_{k+n}-c_k}{n^{1/2}}=0 \qquad uniformly in k.$$

Then we have

(1)
$$\lim_{n\to\infty} P\left\{T_n \leq x \frac{B_1}{2B_2^{1/2}} n^{1/2}\right\} = \left(\frac{2}{\pi}\right)^{1/2} \int_0^x e^{-t^2/2} dt.$$

When the variables $\{X_n\}$ are equidistributed with a c.d.f. having an absolutely continuous component whose derivative is of bounded total variation in $(-\infty, \infty)$, the conditions $(1), \cdots, (3)$ are automatically fulfilled. Then, taking $c_n = c$, we have the case treated by Chung's Theorem 1 [1]. Our generalization is two-fold; the variables need not be equidistributed and we are not confined to crossings of a horizontal line. Condition (4) admits the line $c_n = n^{1/2-\epsilon}$ (though not $c_n = o(n^{1/2})$); on the other hand it follows from the law of the iterated logarithm that the conclusion (1) is no longer valid for $c_n = n^{1/2+\epsilon}$.

Chung's theorem assumes only the existence of β_3 and that $F_n(x)$ is not of lattice type. This additional strength is a consequence of the availability of a stronger expansion theorem in the case of equidistributed random variables (Esseen [6]) than in the case of non-equidistributed random variables (Cramér [5]). Since the proof follows that of [1] we present only an outline.

First we show

(2)
$$E(T_n) = \sum_{k=1}^n \int_{-\infty}^{0} [\overline{F}_k(-c_k - x) - \overline{F}_k(-c_k)] dF_{k+1}(x)$$
$$= n^{1/2} \frac{B_1}{(2\pi B_2)^{1/2}} + o(n^{1/2})$$

using Cramér's theorem to estimate \overline{F}_k . We proceed to calculate $E(T_n^m)$ which can be expressed as a sum of multiple integrals as follows: by the multinomial theorem

(3)
$$E(T_n^m) = \sum_{l=1}^m \sum_{m_l} \frac{m!}{m'_1 \cdots m_l!} \sum_{n_l} E(Y_{n_1}^{m_1} \cdots Y_{n_l}^{m_l}),$$

where

$$\sum_{m_i}' \text{ is a sum over all } m_i > 0 \text{ such that } \sum_{i=1}^l m_i = m,$$

$$\sum_{n_i}^{\prime\prime}$$
 is a sum over all n_i such that $0 < n_1 < \cdots < n_i \le n$,

and

$$Y_k = \begin{cases} 1 & \text{if } S_n \text{ crosses } \{c_n\} \text{ from above at } k, \\ 0 & \text{otherwise.} \end{cases}$$

Let $S_{n_k} - S_{n_{k-1}+1} = R_{n_k}$ have the c.d.f. $G_{n_k}(x)$. We can express the last sum on the right of (3) as

(4)
$$\sum_{n_i} \prod_{k=1}^{l} \int_{-\infty}^{0-} \int_{-C_k}^{-C_k - y_k} dG_{n_k}(x_k) dF_{n_k + 1}(y_k)$$

where \prod denotes iterated integration, and

$$C_k = c_{n_k} + \sum_{i=1}^{k-1} (x_i + y_i).$$

The integral in (4) is a generalization of the one in (2); it can be evaluated up to terms of $o(n^{l/2})$ using induction on l. It is for this evaluation that the uniformity conditions of assumptions (1) to (4) are needed.

The results of these calculations show that the moments of $T_n/(n^{1/2}B_1/2B_2^{1/2})$ approach those of the semi-normal distribution. With Pólya's continuity theorem of the moment problem [7], this gives the final result.

The lower limit of sums of random variables, Case A.

THEOREM 2. Let $\{X_n\}$ be a sequence of independent random variables with zero means which satisfy the condition (1) of Theorem 1 and

(i)
$$\inf_{n} \beta_{2n} > 0,$$

(ii)
$$\sup_{n} \beta_{5n} < \infty,$$

(iii)
$$\lim_{k\to\infty}\inf\frac{s_{n+k}^2-s_n^2}{s_n^2}\geq 1 \qquad uniformly \ in \ n.$$

Then given a monotone nonincreasing sequence of positive numbers $\{a_n\}$,

(5)
$$P\left\{ \liminf_{n \to \infty} \frac{|S_n|}{a_n} = 0 \right\} = \left\{ \begin{cases} 1, \\ 0, \end{cases} \right\}$$

according as

(6)
$$\sum_{n} \frac{a_{n}}{n^{1/2}} \begin{cases} converges, \\ diverges. \end{cases}$$

For the sake of brevity we have stated the conditions in a form somewhat stronger than necessary, for instance the divergence part requires only $\sup_n \beta_{4n} < \infty$. When the X_n are equidistributed with a c.d.f. having an absolutely continuous component, conditions (i), (ii), and (iii) follow from the existence of the fifth moment and the result is the same as Theorem 3 of Chung and Erdös [3].

Again since the proof is similar in method to that of [3] we present an outline. Using Cramér's expansion theorem we first prove the following lemma.

LEMMA. Let $\{X_n\}$ be a sequence of independent random variables with zero means, satisfying conditions (1) of Theorem 1, (i) of Theorem 2, and (ii'), $\sup_n \beta_{pn} < \infty$ (p an integer ≥ 3). Let a(n, k) < b(n, k) be sequences of numbers $(k=1, 2, \cdots; n=k, k+1, \cdots)$ satisfying

$$\lim_{n-k\to\infty} a(n, k) = \lim_{n-k\to\infty} b(n, k) = x \quad uniformly \ in \ k;$$

then

(7)
$$P\left\{a(n, k) \leq \frac{S_n - S_k}{(s_n^2 - s_k^2)^{1/2}} \leq b(n, k)\right\} = \phi(x) \left[b(n, k) - a(n, k)\right] + o\left[b(n, k) - a(n, k)\right] + O\left[(n - k)^{-(p-2)/2}\right]$$

where $\phi(x)$ is the normal density function, and the order terms denote limits as $n-k\to\infty$ which are uniform in k.

The convergence part of Theorem 2 follows directly from the lemma with p=5 without using the uniformity of the order terms in (7). The divergence part is considerably more delicate and is proved by an elaboration of the methods of [3]; it requires condition (iii) and the uniformity of the lemma, but only p=4 in (ii').

The lower limit of sums of random variables, Case B. This section

² Dr. Chung asked me to remark that in the statement of Lemma 6 of [3], $z_n = O(n^{1/2})$ should be replaced by an assumption such as (4) of Theorem 1 above. It should be remarked that (4) is not implied by $c_n = O(g(n))$ if $g(n) \to \infty$ however slowly. The proof in [3] requires only $z_n = O(1)$ which implies (4).

is devoted to proving some lower limit theorems for sums of independent, equidistributed lattice type random variables. The corollary to Theorem 3b and Theorem 5 generalize Theorems 1 and 2 of [3] which are stated for binomial random variables.

THEOREM 3. Let $\{X_n\}$ be a sequence of independent, equidistributed random variables whose possible values are integers and $EX_n = x$. Let $\{a_n\}$ be a nonincreasing, positive sequence and define $\{n_i\}$ to be the (increasing) sequence of integers for which $|I(nx) - nx| < a_n \epsilon$ ($\epsilon > 0$ and I(nx) =the nearest integer to nx). Then

(a) If, for some $\delta > 0$, $E|X_n|^{2+\delta} < \infty$, and $\sum_i n_i^{-1/2}$ diverges,

(8)
$$P\left\{ \liminf_{n\to\infty} \frac{\left|S_n - nx\right|}{a_n} = 0 \right\} = 1.$$

(b) Assuming only the first moment to be finite, when $\sum_{i} n_{i}^{-1/2}$ converges,

(9)
$$P\left\{ \liminf_{n\to\infty} \frac{|S_n - nx|}{a_n} = 0 \right\} = 0.$$

The techniques of the proof require that the sums S_n have the same lattice for each n. This can be accomplished for general lattice type variables by making a linear transformation to an integral lattice. Having done so, however, we can no longer assume the mean to be zero but we can transform so that $0 \le x < 1$. We consider the event

$$\left\{ \liminf_{n \to \infty} \frac{\left| S_n - nx \right|}{a_n} = 0 \right\} = \lim_{\epsilon \to 0} \left\{ \left| S_n - nx \right| < a_n \epsilon \text{ i.o.} \right\}$$
$$= \lim_{\epsilon \to 0} \lim_{n \to \infty} \bigcup_{n \to 0} \left\{ \left| S_n - nx \right| < a_n \epsilon \right\}.$$

Let $E_n = \{ |S_n - nx| < a_n \epsilon \}$. S_n is always an integer whereas nx may not be. We may take ϵ so that $a_n \epsilon < 1$ for all n, then for those values of n for which nx is not within $a_n \epsilon$ of an integer, the event E_n is impossible. Therefore for the sequence $\{n_i\}$ defined in the theorem,

$$E_n = \begin{cases} 0 & \text{for } n \in \{n_i\}, \\ G_i = \{S_{n_i} = I(n_i x)\} & \text{for } n_i \in \{n_i\}. \end{cases}$$

Obviously

$${E_n \text{ i.o.}} = {G_i \text{ i.o.}}$$

and we need consider only the G_i .

We shall use Lemmas 1 and 2 from [3] which link $\{a_n\}$ with $\{n_i\}$. We state them somewhat differently but omit the proofs which are unchanged.

LEMMA 1. If $a_n \downarrow 0$, then for integers i, k > 0,

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

LEMMA 2. If $\sum_{n} a_n/n^{1/2}$ diverges, then $\sum_{i} 1/n_i^{1/2}$ diverges.

For any integers a, b let $F_k = \{S_{nk} = I[(n_k + a)x] + b\}$. We express these events as a disjoint union

$$F_{i} = \bigcup_{k=h}^{i} \left[F_{i} \cap \left(F_{k} \cap \bigcap_{j=h}^{i} F_{j}' \right) \right] \qquad (i \ge h).$$

Thus

(10)
$$\sum_{i=h}^{n} P(F_{i}) = \sum_{i=h}^{n} \sum_{k=h}^{i} P\left[F_{i} \cap \left(F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}^{\prime}\right)\right] \\ = \sum_{k=h}^{n} P\left(F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}^{\prime}\right) \sum_{i=k}^{n} P\left(F_{i} \mid F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}^{\prime}\right).$$

Now

(11)
$$P\left(F_{i} \mid F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}'\right) = P(F_{i} \mid F_{k})$$

$$= P\left\{S_{n,i-n_{k}} = I\left[(n_{i} + a)x\right] - I\left[(n_{k} + a)x\right]\right\}.$$

Lemma 1, p. 44 of Esseen [6] may be used to extend Theorem 5, p. 63 of that paper to the case where $E|X|^{2+\delta} < \infty$, giving for a discontinuity point, ξ , of S_n

(12)
$$P\left\{S_n = \xi\right\} = \frac{1}{\sigma n^{1/2}} \phi\left(\frac{\xi - nx}{\sigma n^{1/2}}\right) + O\left(\frac{1}{n^{(1+\delta)/2}}\right).$$

Now $I[(n_i+a)x]-I[(n_k+a)x]=I[(n_i-n_k)x]+\theta$ where $\theta=0$ or ± 1 . By Lemma 2, $n_i-n_k \ge n_p$, where p=[(i-k)/2]. From (12) and (11) we have

$$\lim_{i-k\to\infty}\frac{P(F_p)n_p^{1/2}}{P(F_i|F_k)(n_i-n_k)^{1/2}}=1.$$

Thus for any $\eta > 0$ we can find $N(\eta)$ such that for $i - k \ge N(\eta)$

$$P(F_i | F_k) \le (1 + \eta) P(F_p) \quad (p = [(i - k)/2]).$$

From (10) we have

$$\sum_{i=h}^{n} P(F_{i}) \leq \sum_{k=h}^{n} P\left(F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}'\right) \cdot \left[\sum_{i=k}^{N-1} P(F_{i} \mid F_{k}) + 2(1+\eta) \sum_{i=\lfloor N/2 \rfloor}^{\lfloor n/2 \rfloor} P(F_{i})\right].$$

Whence

$$\sum_{k=h}^{n} P\left(F_{k} \cap \bigcap_{j=h}^{k-1} F_{j}^{\prime}\right) = P\left(\bigcup_{k=h}^{n} F_{k}\right) \ge \frac{\sum_{i=h}^{n} P(F_{i})}{N + 2(1+\eta) \sum_{[N/2]}^{[n/2]} P(F_{i})}.$$

Since by (12), $\sum_i P(F_i)$ diverges, we can find a $K(\zeta)$ for $0 < \zeta < 1/2$ such that $k \ge K(\zeta)$ implies

$$P\left(\bigcup_{j=h}^{k} F_{j}\right) = P\left(\bigcup_{j=h}^{k} \left\{S_{n_{j}} = I\left[(n_{j} + a)x + b\right\}\right) \ge \zeta.$$

This statement implies the statement: for $\zeta > 0$, there is a $K(\zeta)$ such that for h > 0 and $k \ge K(\zeta)$,

(13)
$$P\left(\bigcup_{i=h}^{k} \left\{ S_i = I[(j+a)x] + b \right\} \right) \ge \zeta^{1/2}.$$

Now we choose a sequence of integers, $k_1 = h < k_2 < k_3 < \cdots$, as follows: Suppose k_r is already chosen. Let

(14)
$$A_i = \bigcup_{\substack{j=k_{i-1}\\j=k_{i-1}}}^{k_i-1} \left\{ S_j = v(j) = I(jx) \right\} \qquad (i = 1, 2, \dots, r+1)$$

where k_{r+1} is to be chosen. We have

$$P\left(A_{r+1} \middle| \bigcap_{i=1}^{r} A_{i}^{\prime}\right)$$

$$= \frac{1}{P\left(\bigcap_{i=1}^{r} A_{i}^{\prime}\right)} \int_{\sigma \neq v(k_{r}-1)} P(A_{r+1} \middle| S_{k_{r}-1} = \sigma) dP(\sigma)$$

$$\geq \underset{|\sigma| \leq c_{r}}{\text{Min}} P(A_{r+1} \middle| S_{k_{r}-1} = \sigma) P\left\{ \middle| S_{k_{r}-1} \middle| \leq c_{r} \middle| \bigcap_{i=1}^{r} A_{i}^{\prime} \right\}.$$

We choose the c_r in advance so $P\{|S_{k_r-1}| \le c_r | \bigcap_{i=1}^r A_i'\} \ge \zeta^{1/2}$. Since the

minimum of the first term is taken over a finite set of integers, σ , it is achieved for some σ_r . Furthermore we have, for v(j) defined in (14),

$$\begin{split} P(A_{r+1} \mid S_{k_r-1} &= \sigma_r) \\ &= P\left(\bigcup_{j=k_r}^{k_{r+1}-1} \left\{ S_j = v(j) \right\} \mid S_{k_r-1} = \sigma_r \right) \\ &= P(\bigcup \left\{ S_j - S_{k_r-1} = v(j) - \sigma_r \right\}) \\ &= P(\bigcup \left\{ S_{j-k_r+1} = I[(j-k_r+1+k_r-1)x] - \sigma_r \right\}). \end{split}$$

We choose k_r by (13) so that this is greater than $\zeta^{1/2}$.

We have now a sequence $\{k_i\}$ and an associated sequence of events $\{A_i\}$ which have the property

$$P\left(A_r \middle| \bigcup_{i=1}^{r-1} A_i^{\prime}\right) \ge \zeta \qquad \text{for every } r = 1, 2, \cdots.$$

Thus

$$1 - P(A_i \text{ i.o.}) = \lim_{h \to \infty} \lim_{n \to \infty} P\left(\bigcap_{i=h}^{n} A_i'\right)$$

$$= \lim_{h \to \infty} \lim_{n \to \infty} P(A_h') P(A_{h'+1} | A_h') \cdot \cdot \cdot \cdot P\left(A_n' | \bigcap_{i=h}^{n-1} A_i'\right)$$

$$\leq \lim_{h \to \infty} \lim_{n \to \infty} (1 - \zeta)^{n-h}$$

$$= 0.$$

This proves Theorem 3a.

For Theorem 3b, we have from Theorem 1 of Chung and Erdös [2]

$$P(G_i) = P\{S_{n_i} = I(n_i x)\} < A n_i^{-1/2}.$$

Therefore $\sum_{i} P(G_i)$ converges, and by the Borel-Cantelli Lemma, Theorem 3b is proved

We now use these results to achieve generalizations of the results of [3].

COROLLARY TO THEOREM 3a. If $\sum_{n} a_n/n^{1/2}$ diverges and $E|X|^{2+\delta} < \infty$, the statement (8) is true.

The proof is an application of Lemma 2.

THEOREM 4. Assuming only the first moment to be finite, if x is rational the statement (8) is true regardless of the choice of $\{a_n\}$.

Let x = p/q and transform to Y = qX - p, setting $T_n = \sum_{i=1}^n Y_i$.

By Theorem 4, Chung and Fuchs [4],

$$P\{T_n = 0 \text{ i.o.}\} = 1.$$

THEOREM 5. When $\sum_{n} a_n/n^{1/2}$ converges, then for almost all (Lebesgue) real numbers, x, the statement (9) is true.

This proof was communicated to me by Dr. P. Erdös. Choose ϵ so that $\epsilon a_n < 1$ for all n. Then let

$$\{n, x\} = \{n, x: |nx - I(nx)| < \epsilon a_n\},$$

and let $\chi(n, x)$ be the set characteristic function of $\{n, x\}$. We have for any n

$$\int_0^1 \chi(n, x) dx \le 2\epsilon a_n$$

and for a given x

$$\sum_{n=1}^{\infty} \frac{\chi(n, x)}{n^{1/2}} = \sum_{i=1}^{\infty} n_i^{-1/2},$$

where $\{n_i\}$ is the sequence defined in the statement of Theorem 3. It is, of course, a function of x. From the above we have

$$\int_{0}^{1} \sum_{n=1}^{N} \frac{\chi(n, x)}{n^{1/2}} dx \le \sum_{n=1}^{N} \frac{2a_n \epsilon}{n^{1/2}} \le A \qquad \text{for all } N.$$

Since A does not depend on N, the set of x for which $\sum_{n} \chi(nx)/n^{1/2}$ converges has measure one. With Theorem 1b the proof is completed.

The corollary and Theorem 5 give us almost the same results as hold in the continuous case, except that when (6) converges there is an exceptional set on [0, 1) of measure zero which contains all the rational numbers and such that if EX is in this set the result (8) is true. It is not known whether the exceptional set contains any irrationals, but a simple argument shows that for any irrational number, one can choose the sequence $\{a_n\}$ such that the result (9) is true. Thus, unlike the rationals, the irrationals of the exceptional set depend on the choice of the normalizing sequence. To sketch this argument, for x irrational let $n_1=1$, $n_{i+1}=\min_{n>n_i}$ $\{n: |nx-I(nx)| < |n_ix-I(n_ix)| \}$. By a theorem of Weyl [8, p. 378], |nx-I(nx)| is dense on the unit interval and thus $\{n_i\}$ is an infinite set. Now for $\epsilon > 0$ choose a_n so that for $n_{i-1} \le n < n_i$, $a_n \epsilon < |n_ix-I(n_ix)|$. Since $a_n \epsilon < |nx-I(nx)|$ for all n, the events E_n are all impossible.

Acknowledgment. I am indebted to Professor K. L. Chung for suggesting the problems and for advising me during the research.

BIBLIOGRAPHY

- 1. K. L. Chung, Fluctuation of sums of independent random variables, Ann. of Math. vol. 51 (1950) pp. 697-706.
- 2. K. L. Chung and P. Erdös, Probability limit theorems assuming only the first moment. I, Memoirs of the American Mathematical Society, no. 6, pp. 13-19.
- 3. ——, On the lower limit of sums of independent random variables, Ann. of Math. vol. 48 (1947) pp. 1003-1013.
- 4. K. L. Chung and W. H. J. Fuchs, On the distribution of values of sums of random variables, Memoirs of the American Mathematical Society, no. 6, pp. 1-9.
- 5. H. Cramér, Random variables and probability distributions, Cambridge Tracts in Mathematics, no. 36, 1937.
- 6. C. G. Esseen, Fourier analysis of distribution functions, Acta Math. vol. 77 (1945) pp. 5-145.
- 7. G. Pólya, Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentproblem, Math. Zeit. vol. 8 (1920) pp. 171-181.
- 8. G. H. Hardy and E. M. Wright, An introduction to the theory of numbers, Oxford, 1938.

Columbia University

ON PROJECTIVE GEOMETRY OVER FULL MATRIX RINGS

RIMHAK REE

1. Introduction. In this note we show that projective geometry over a ring R and that over the full matrix ring R_n are essentially the same, and extend the fundamental theorem of projective geometry [1, p. 44] to the case of Φ_n -modules, where Φ is a division ring. (By a projective geometry over R we mean a lattice of all R-submodules of an R-module.) As a special case of these results we have the following: If $n \ge 3$, any lattice isomorphism of the lattice of all left ideals of Φ_n and that of Ψ_m where Φ and Ψ are division rings, is induced by an isomorphism of Φ_n and Ψ_m . We obtain also an extension of the basis theorem for vector spaces to Φ_n -modules.

Other extensions of the fundamental theorem of projective geometry have been made by Baer, for the case of R-modules, where R is a "primary ring" in his sense [2, p. 304], and the ring of rational integers [3].

2. Main theorems. In the following, by a ring we always mean an associative ring with unit element. Let R be a ring with unit