GENERALIZATION OF A THEOREM OF CHUNG AND FELLER

MIRIAM LIPSCHUTZ

1. Introduction. Chung and Feller [1] have obtained the following result on fluctuations in coin tossing. Let X_k be a sequence of independent random variables, each assuming the values ± 1 with probability 1/2. Let $S_n = X_1 + \cdots + X_n$ and let N_n denote the number of those terms in the sequence $S_1, S_2, \cdots, S_n$ which either are positive, or which are zero but follow a positive term. With the familiar notation for conditional probabilities the Chung-Feller theorem states that for each integer $r \leqq n$

$$\text{(1)} \qquad \mathrm{P}\{N_{2n} \leqq 2r \mid S_{2n} = 0\} = \frac{r+1}{n+1}\cdot$$

In other words, under the hypothesis $S_{2n}=0$ the variable N_{2n} becomes uniformly distributed. This is in contrast to the unconditional distribution for N_{2n} which is given by the arc sine law.

The aim of the present paper is to generalize this result to arbitrary lattice variables.[1] In this case the result of the form (1) will have only an asymptotic character.

We shall prove the following.

THEOREM. *Let X_k be a sequence of independent random variables with a common distribution $F(x)$ such that*: (A) *The r.v. have mean zero, variance* 1, *and finite fourth moment*; (B) *The r.v. have a lattice distribution such that the minimum distance between jumps is one unit.* (*The r.v. X_k thus assume integer values only.*) *Then*

$$\text{(2)} \qquad \mathrm{P}(N_n \leqq \alpha n \mid S_n = 0) = \frac{[\alpha n]+1}{n+1} + g(n), \qquad 0 \leqq \alpha \leqq 1,$$

where

$$g(n) = O\left(\frac{\log n}{n^{1/30}}\right) \textit{ if the r.v. have third moment zero,}$$

and

$$g(n) = O\left(\frac{\log n}{n^{1/72}}\right) \textit{ if the r.v. have third moment differing from zero.}$$

Received by the editors April 7, 1951 and, in revised form, September 23, 1951.

[1] I wish to thank Professor K. L. Chung for suggesting this problem to me.

2. Method of proof. The proof is carried out in three steps.

The first step adapts the Erdös-Kac invariance method [2; 3] to conditional probabilities of the form $P(N_n \leqq \alpha n | S_n = 0)$. This leads to (10).

The next step involves the application of a result of Esseen [4] on the multi-dimensional central limit theorem for lattice distributions to the upper and lower bounds in this inequality (10). In order to do so we use Euler's formula to extend Esseen's theorem to regions of summation appropriate to our case. This gives (18) which may be of independent interest.

Having thus obtained bounds, with errors which tend to zero in the limit, for $P(N_n \leqq \alpha n, S_n = 0)$, we note that these bounds are independent of the distribution of the individual X_k. We complete the proof by substituting for these the known limiting value in the special case of Chung-Feller.

We add that, unfortunately, the order of magnitude of the error term in the final approximation to the uniform distribution is still very large.

3. Generalization of the invariance principle of Erdös-Kac. Let

$$\phi(S_r) = \begin{cases} 1 & \text{if } S_r > 0, \quad S_n = 0, \quad r < n,^2 \\ 0 & \text{otherwise,} \end{cases}$$

and note that

$$P(N_n < \alpha n, S_n = 0) = P\left(\frac{1}{n} \sum_{r=1}^{n} \phi(S_r) < \alpha\right).$$

Let now $k = n^{1/5}$, $\epsilon = k^{-1/3}$, $\delta = k^{-1/6}$,

$$n_i' = \left[\frac{in}{k}\right], \qquad n_i = \left[\frac{(i - 1/2)n}{k}\right] \qquad \text{for } i = 1, 2, \cdots, k,$$

and put

$$D_n = \frac{1}{n}\left(\sum_{r=1}^{n} \phi(S_r) \; - \sum_{i=1}^{k} (n_i' - n_{i-1}')\phi(S_{n_i})\right).$$

We have

$$E(|D_n|) \leqq \frac{1}{n} \sum_{i=1}^{k} \sum_{r=n_{i-1}'+1}^{n_i'} E(|\phi(S_{n_i}) - \phi(S_r)|)$$

[2] Note the convention made in the introduction concerning the positiveness of the sums S_r.

and we now wish to estimate $E(|\phi(S_r)-\phi(S_{n_i})|)$ for $n'_{i-1} \leqq r \leqq n'_i$. Notice that

$$E(|\phi(S_r) - \phi(S_{n_i})|) = P(S_{n_i} > 0, S_r < 0, S_n = 0) + P(S_{n_i} < 0, S_r > 0, S_n = 0) \tag{3}$$

and that for $\epsilon > 0$

$$\begin{aligned} P(S_{n_i} > 0, S_r < 0, S_n = 0) &= P(\epsilon n_i^{1/2} > S_{n_i} > 0, S_r < 0, S_n = 0) \\ &\quad + P(S_{n_i} > \epsilon n_i^{1/2}, S_r < 0, S_n = 0) \\ &\leqq P(\epsilon n_i^{1/2} > S_{n_i} > 0, S_n = 0) \\ &\quad + P(S_{n_i} - S_r > \epsilon n_i^{1/2}, S_n = 0). \end{aligned} \tag{4}$$

We have, for $n'_{i-1} < r \leqq n_i$,

$$\begin{aligned} &P(S_{n_i} - S_r > \epsilon n_i^{1/2}, S_n = 0) \\ &\qquad = \sum_{y > \epsilon n_i^{1/2}} P(S_{n_i} - S_r = y, S_n = 0) \\ &\qquad = \sum_{y > \epsilon n_i^{1/2}} P(S_{n_i} - S_r = y, S_n - (S_{n_i} - S_r) = -y) \\ &\qquad = \sum_{y > \epsilon n_i^{1/2}} P(S_{n_i} - S_r = y) P(S_{n-(n_i-r)} = -v), \end{aligned} \tag{5}$$

whereas for $n_i < r \leqq n'_i$,

$$\begin{aligned} &P(S_{n_i} - S_r > \epsilon n_i^{1/2}, S_n = 0) \\ &\qquad = \sum_{y > \epsilon n_i^{1/2}} P(S_r - S_{n_i} = -y) P(S_{n-(r-n_i)} = y). \end{aligned} \tag{6}$$

Now, since in both cases considered we have $0 \leqq |n_i - r| \leqq n/2k$ and $n \geqq n - (|n_i - r|) \geqq n - n/2k$, it follows by the central limit theorem that

$$P(S_{n-|n_i-r|} = y) < \frac{c}{(n - n/2k)^{1/2}} \sim \frac{c}{n^{1/2}}. \tag{7}$$

By Chebycheff's inequality we then obtain

$$(5) < \frac{n_i - r}{\epsilon^2 n_i} \frac{c}{n^{1/2}},$$

$$(6) < \frac{r - n_i}{\epsilon^2 n_i} \frac{c}{n^{1/2}}.$$

We consider next

$$
\begin{aligned}
\mathrm{P}(\epsilon n_i^{1/2} > S_{n_i} > 0, S_n = 0) &= \sum_{0<y<\epsilon n_i^{1/2}} \mathrm{P}(S_{n_i} = y, S_n = 0) \\
&= \sum_{0<y<\epsilon n_i^{1/2}} \mathrm{P}(S_{n_i} = y, S_n - S_{n_i} = -y) \\
&= \sum_{0<y<\epsilon n_i^{1/2}} \mathrm{P}(S_{n_i} = y)\mathrm{P}(S_{n-n_i} = -y) \\
&< \frac{c}{(n-n_i)^{1/2}} \mathrm{P}(0 < S_{n_i} < \epsilon n_i^{1/2}).
\end{aligned}
$$

We operate in exactly the same way with the second term of (3) and obtain

$$
\begin{aligned}
\mathrm{E}(|D_n|) \leqq \frac{1}{n} \sum_{i=1}^{k} \Bigg\{ & \frac{(n_i - n'_{i-1})(n_i - n'_{i-1} - 1) + (n'_i - n_i)(n'_i - n_i - 1)}{\epsilon^2 n_i} \frac{c}{n^{1/2}} \\
& + (n'_i - n'_{i-1})\mathrm{P}(-\epsilon n_i^{1/2} < S_{n_i} < \epsilon n_i^{1/2}) \frac{c}{(n-n_i)^{1/2}} \Bigg\}
\end{aligned} \tag{8}
$$

Now for large n, $n_i - n'_{i-1} = n/2k + \theta'_i$ and $n'_i - n_i = n/2k + \theta''_i$, where $|\theta'_i|$ and $|\theta''_i|$ are less than 1. Thus the first term in (8) is equal to

$$
\begin{aligned}
\frac{2}{n\epsilon^2} \sum_{i=1}^{k} \frac{[(n/2k) + \theta'_i][(n/2k) + \theta''_i]}{in/k} &= \mathrm{I} + \mathrm{II} \\
&= \frac{1}{k\epsilon^2} \sum_{i=1}^{k} \frac{1}{i} + o(\mathrm{I}) < \frac{1 + \log k}{k\epsilon^2}.
\end{aligned}
$$

Since

$$
\mathrm{P}(-\epsilon n_i^{1/2} < S_{n_i} < \epsilon n_i^{1/2}) = \frac{1}{(2\pi)^{1/2}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{t^2}{2}\right) dt + \frac{\theta Q}{(n_i)^{1/2}},
$$

where[3] $|\theta| < 1$ and Q is a function of the distribution of X only, the second term in (8) equals

$$
\begin{aligned}
\frac{1}{n} \sum_{i=1}^{k} (n'_i - n'_{i-1}) & \left(\epsilon + \frac{\theta Q}{(n_i)^{1/2}}\right) \frac{1}{(n-n_i)^{1/2}} \\
&= \frac{\epsilon}{k} \sum_{i=1}^{k} \frac{1}{(n-n_i)^{1/2}} + \mathrm{II} = \mathrm{I} + \mathrm{II} = \frac{c\epsilon}{(n)^{1/2}} + o(\mathrm{I}).
\end{aligned}
$$

Finally we obtain

[3] The value of these constants need not be the same each time these symbols appear.

$$\mathrm{E}(|D_n|) < \frac{1+\log k}{k\epsilon^2}\frac{c}{n^{1/2}} + \frac{c\epsilon}{n^{1/2}} = R(n, \epsilon, k). \tag{9}$$

From here on we proceed as in Kac-Erdös to conclude that

$$\begin{aligned} \mathrm{P}\left\{\frac{1}{n}\sum_{i=1}^{k}(n_i' - n_{i-1}')\phi(S_{n_i}) < \alpha - \delta\right\} &- \frac{R(n, \epsilon, k)}{\delta} \\ &\leqq \mathrm{P}(N_n < \alpha n) \\ &\leqq \mathrm{P}\left\{\frac{1}{n}\sum_{i=1}^{k}(n_i' - n_{i-1}')\phi(S_{n_i}) < \alpha + \delta\right\} + \frac{R(n, \epsilon, k)}{\delta}. \end{aligned} \tag{10}$$

4. **Application of the theorem of Esseen.** We now proceed to evaluate

$$\mathrm{P}\left(\frac{1}{n}\sum_{i=1}^{k}(n_i' - n_{i-1}')\phi(S_{n_i}) < \alpha\right).$$

Write

$$V_i^{(p)} = \sum_{j\equiv p \bmod (n_1),\, j\leqq n_i} X_j, \qquad i = 1, \cdots, k,$$
$$V_{k+1}^{(p)} = V_k^{(p)} + X_{n_k+p},$$

and consider the $(k+1)$-dimensional vector

$$V^{(p)} = (V_1^{(p)}, V_2^{(p)}, \cdots, V_k^{(p)}, V_{k+1}^{(p)}).$$

We have

$$\sum_{p=1}^{n_1} V^{(p)} = (\tilde{S}_{n_1}, \cdots, \tilde{S}_{n_k}, \tilde{S}_n) \sim (S_{n_1}, S_{n_2}, \cdots, S_{n_k}, S_n).$$

(The $\tilde{}$ on the sums are necessary since $\sum_{p=1}^{n_1} V_i^{(p)}$ may not be exactly equal to S_{n_i}. For this sum equals

$$X_1 + X_2 + \cdots + X_{n_1+n_1} + X_{n_1'+1} + \cdots + X_{n_1'+n_1} + X_{n_2+1} + \cdots + X_{n_{i-1}'+n_1-1} + X_{n_{i-1}'+n_1}$$

but $n_1' - 2 < n_1 + n_1 \leqq n_1'$, $n_2 - 2 < n_1' + n_1 \leqq n_2, \cdots, n_i - 2 < n_{i-1}' + n_1 \leqq n_i$. Thus $\tilde{S}_{n_i} = S_{n_i} - T_i$ where $T_i \leqq 2S_i$ and $i \leqq k$. Since we are however within the domain of the central limit theorem, the sum T_i is negligible for large n.)

The $V^{(p)}$ are identically distributed r.v. in R_{k+1} with mass concentrated in the lattice points $\sum_{i=1}^{k+1} \nu_i \boldsymbol{a}_i$. The $\boldsymbol{a}_i$ are $(k+1)$-dimensional vectors with components a_{ij} equal to δ_{ij} and the $\nu_i = 0, \pm 1, \pm 2, \cdots$.

The correlation matrix $\Delta/2^{k-1}=(\mu_{ij})/2^{k-1}$ and its inverse $\Delta^{-1}=(\Delta_{ij})/|\Delta|$ are given by

$$\Delta = \begin{pmatrix} 1, 1, 1, & \cdots, & 1 \\ 1, 3, 3, & \cdots, & 3 \\ 1, 3, 5, 5, & \cdots, & 5 \\ 1, 3, 5, 7, & \cdots, & 7 \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ 1, 3, 5, 7, & \cdots, 2k-1, & 2k-1 \\ 1, 3, 5, 7, & \cdots, 2k-1, & 2k \end{pmatrix},$$

$$\Delta^{-1} = \begin{pmatrix} 3/2, & -1/2, & 0, & 0, & & \cdots, 0 \\ -1/2, & 1, & -1/2, & 0, & 0, & \cdots, 0 \\ 0, & -1/2, & 1, & -1/2, & 0, & 0, \cdots, 0 \\ 0, & 0, & -1/2, & 1, & -1/2, & 0, \cdots, 0 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ 0, & 0, \cdots, & & & & -1/2,\ 1,\ -1/2,\ 0 \\ 0, & 0, \cdots, & & & & 0,\ -1/2,\ 3/2,\ -1 \\ 0, & 0, \cdots, & & & & 0,\ 0,\ -1,\ 1 \end{pmatrix}.$$

The determinant $|\Delta|$ of Δ equals 2^{k-1}.[4]

We write for all $p=1, \cdots, n_1$,

$$\alpha_j^{\nu} = \int (V_j^{(p)})^{\nu} dP, \tag{11}$$

and note that

$$\alpha_j^3 = (2j-1)^3\gamma \quad \text{and} \quad \sum_{j=1}^{k} \alpha_j^3 = Ck^4 \quad \text{if} \quad \gamma = E(X^3) \neq 0. \tag{12}$$

We introduce the following transformation of variables for all $p=1, \cdots, n_1$,

$$Y^{(p)} = (Y_1^{(p)}, \cdots, Y_{k+1}^{(p)})$$

with $Y_i^{(p)} = \sum_{j=1}^{k+1} c_{ij} V_j^{(p)}$, and the c_{ij} so determined that

[4] This is easily seen by noting that in the expansion of $|\Delta|$ by the first row, only the first and second terms differ from zero since all the other determinants in the expansion have at least one column which is a multiple of the first column. The same is again true for the new determinant obtained, etc.

$$E(Y_i^{(p)})^2 = 1 \quad \text{and} \quad E(Y_i^{(p)}Y_j^{(p)}) = 0 \quad \text{for} \quad i \neq j.$$

It then follows that $|c_{ij}| = 1/|\Delta|^{1/2}$ and

$$\sum_{i=1}^{k} (Y_i)^2 = \sum_{i,j=1}^{k} (\Delta_{ij}/\Delta) V_i V_j \tag{13}$$

where we have written Y_i for $Y_i^{(p)}$ since the $Y_i^{(p)}$ are also identically distributed. The mass of the $Y^{(p)}$ will be concentrated in the points $\sum_{i=1}^{k+1} \nu_i \mathbf{c}_i$ where the $\mathbf{c}_i$ are $(k+1)$-dimensional vectors with components c_{ij} and where $\nu_i = 0, \pm 1, \pm 2, \cdots$. Then

$$\begin{aligned} &\mathrm{P}(V_1^{n_1} = \eta_1, V_2^{n_1} = \eta_2, \cdots, V_k^{n_1} = \eta_k, V_{k+1}^{n_1} = 0) \\ &\mathrm{P}\left(Y_1^{n_1} = \sum_{i=1}^{k} c_{1i}\eta_i,\ Y_2^{n_1} = \sum_{i=1}^{k} c_{2i}\eta_i, \cdots, Y_{k+1}^{n_1} = \sum_{i=1}^{k} c_{k+1,i}\eta_i\right), \end{aligned} \tag{14}$$

where

$$\sum_{p=1}^{n_1} Y_j^{(p)} = Y_j^{n_1}, \qquad \sum_{p=1}^{n_1} V_j^{(p)} = V_j^{n_1}.$$

We now quote the following theorem of Esseen [**4**]:

THEOREM. *Let $Y^{(1)}, Y^{(2)}, \cdots, Y^{(n)}, \cdots$ be a sequence of independent, identically distributed r.v. in R_k with the lattice distribution defined above for the $Y^{(p)}$. Let an arbitrary r.v. of the sequence have the properties*: (1) *the mean values are equal to zero*, (2) *the dispersions are equal to one*, (3) *the mixed moments of first order are zero*, (4) *the fourth moments are finite. Then the probability distribution*[5] *$P_n(E)$ of $(Y^{(1)}+Y^{(2)}+\cdots+Y^{(n)})/n^{1/2}$ is also a lattice distribution and the probability mass $q_n(\epsilon_1, \epsilon_2, \cdots, \epsilon_k)$ at a discontinuity point $(\epsilon_1, \epsilon_2, \cdots, \epsilon_k)$ of $P_n(E)$ is expressed by*[5]

$$\begin{aligned} q_n(\epsilon_1\epsilon_2 \cdots \epsilon_k) = \left(\frac{1}{2\pi n}\right)^{k/2} |c_{ij}| \Bigg\{ &\exp\left(-\left(\frac{\epsilon_1^2 + \cdots + \epsilon_k^2}{2}\right)\right) \\ &- \frac{1}{6n^{1/2}}\left(\alpha_1 \frac{\partial}{\partial\epsilon_1} + \alpha_2 \frac{\partial}{\partial\epsilon_2} + \cdots + \alpha_k \frac{\partial}{\partial\epsilon_k}\right)^3 \\ &\cdot \exp\left(-\left(\frac{\epsilon_1^2 + \cdots + \epsilon_k^2}{2}\right)\right)\Bigg\} + O\left(\frac{1}{n^{(k+2)/2}}\right). \end{aligned}$$

We also recall the following facts derived from the Euler summation formula; let $f(x)$ be continuous and have continuous first and second derivatives and $f(\infty) = f'(\infty) = f''(\infty) = f'''(\infty) = 0$. Then

[5] Note the definition (11).

$$
\begin{aligned}
\sum_{t_i>0} f(t_i) &= \int_{1/2}^{\infty} f(t)dt - \int_{1/2}^{\infty} f'(t)P_1(t)dt \\
&= \int_{1/2}^{\infty} f(t)dt - P_2\left(\frac{1}{2}\right)f'\left(\frac{1}{2}\right) + \int_{1/2}^{\infty} P_2(t)f''(t)dt \qquad (15)\\
&= \int_{1/2}^{\infty} f(t)dt - P_2\left(\frac{1}{2}\right)f'\left(\frac{1}{2}\right) - \int_{1/2}^{\infty} P_3(t)f'''(t)dt
\end{aligned}
$$

where

$$
\begin{aligned}
P_1(t) &= [t] - t + 1/2, \\
P_2(t) &= \sum_{\nu=1}^{\infty} \frac{\cos 2\pi\nu t}{2(\nu\pi)^2}, \\
P_3(t) &= \sum_{\nu=1}^{\infty} \frac{\sin 2\pi\nu t}{2(\nu\pi)^2},
\end{aligned}
$$

and

$$
P_2(1/2) < P_2(0) = \frac{1}{12}, \qquad P_3(1/2) = 0.
$$

We apply the theorem of Esseen to the expression in (14). We define $\eta_0 = 0$. From (13) and the fact that $\eta_{k+1} = 0$ we have that (14) equals

$$
\begin{aligned}
(14) = \left(\frac{k}{2\pi n}\right)^{(k+1)/2} \frac{1}{2^{(k-1)/2}} \Bigg\{ &\exp\left[-\frac{k}{4n}\sum_{i=0}^{k}(\eta_i - \eta_{i+1})^2\right] \\
&- \frac{n}{6k}\sum_{p=1}^{k} \alpha_p^3 \frac{\partial^3}{\partial \eta_p^3} \exp\left[-\frac{k}{4n}\sum_{i=0}^{k}(\eta_i - \eta_{i+1})^2\right]\Bigg\} \qquad (16)\\
&+ O\left(\frac{1}{n^{(k+2)/2}}\right).
\end{aligned}
$$

Let now

$$
\begin{aligned}
f(\eta_1, \cdots, \eta_k) = \exp&\left[-\frac{k}{4n}\sum_{i=0}^{k}(\eta_i - \eta_{i+1})^2\right] \\
&- \frac{n}{6k}\sum_{p=1}^{k} \alpha_p^3 \frac{\partial^3}{\partial \eta_p^3} \exp\left[-\frac{k}{4n}\sum_{i=0}^{k}(\eta_i - \eta_{i+1})^2\right].
\end{aligned}
$$

Note that

$$\frac{\partial}{\partial \eta_p} f(\eta_1, \cdots, \eta_k)$$
$$= -\frac{1}{2}\frac{k}{n}(2\eta_p - \eta_{p-1} - \eta_{p+1}) f(\eta_1, \cdots, \eta_k) + o\left(\frac{k}{n}\right),$$
$$\frac{\partial^2}{\partial \eta_p^2} f(\eta_1, \cdots, \eta_k)$$
$$= \left\{-\frac{k}{n} + \frac{k^2}{4n^2}(2\eta_p - \eta_{p-1} - \eta_{p+1})^2\right\} f(\eta_1, \cdots, \eta_k) + o\left(\frac{k}{n}\right)^2,$$
$$\frac{\partial^3}{\partial \eta_p^3} f(\eta_1, \cdots, \eta_k)$$
$$= \left\{\frac{3}{2}\left(\frac{k}{n}\right)^2 (2\eta_p - \eta_{p-1} - \eta_{p+1}) + o\left(\frac{k}{n}\right)^2\right\} f(\eta_1, \cdots, \eta_k).$$

Let y^α be a k-dimensional vector with $y_i = 0$ or 1 for any i and such that $\sum_{i=1}^k (y_i^\alpha/k) < \alpha$. Denote by F_α the set of all such vectors and by E_0 and E_1 the set of all integers less than zero and greater than zero, respectively. We have for any vector $y' \in F_\alpha$

$$P(S_{n_1} \in E_{y_1'}, S_{n_2} \in E_{y_2'}, \cdots, S_{n_k} \in E_{y_k'}, S_n = 0)$$
$$= \sum_{\eta_1 \in E_{y_1'}, \cdots, \eta_k \in E_{y_k'}} \left(\frac{k}{2\pi n}\right)^{(k+1)/2} \frac{1}{2^{(k-1)/2}} f(\eta_1, \cdots, \eta_k)$$
$$+ O\left(\frac{1}{n^{(k+2)/2}}\right)$$
$$= \left(\frac{k}{2\pi n}\right)^{(k+1)/2} \frac{1}{2^{(k-1)/2}} \sum_{\eta_1 \in E_{y_1'}} \sum_{\eta_2 \in E_{y_2'}} \cdots \sum_{\eta_k \in E_{y_k'}} f(\eta_1, \cdots, \eta_k)$$
$$+ O\left(\frac{1}{n^{(k+2)/2}}\right).$$

Using formula (15) above and remembering that the regions $E_{y_i'}$ are either $(1, \infty)$ or $(-\infty, -1)$, we obtain

$$\sum_{\eta_1 \in E_{y_1'}} f(\eta_1, \cdots, \eta_k) = \int_{E_{y_1'}}^h f(\eta_1, \cdots, \eta_k) d\eta_1 + h_1,$$

where

$$| h_1 | < (k/24n)(1 - \eta_2) f(1/2, \eta_2, \cdots, \eta_k).$$

Similarly

$$\sum_{\eta_2 \in E_{y_2'}} \sum_{\eta_1 \in E_{y_1'}} f(\eta_1, \cdots, \eta_k) = \int_{E_{y_1'}} \int_{E_{y_2'}} f(\eta_1, \cdots, \eta_k) d\eta_1 d\eta_k$$
$$+ \int_{E_{y_2'}} h_1 d\eta_2 + \int_{E_{y_1'}} h_2 d\eta_1 + h_{12},$$

where

$$|h_2| < (k/24n)(1 - \eta_1 - \eta_3) f(\eta_1, 1/2, \eta_3, \cdots, \eta_k)$$

and

$$|h_{12}| < (k/12n) f(1/2, 1/2, \eta_3, \cdots, \eta_k).$$

We proceed in this way and obtain after integration of the error terms

$$\begin{aligned} &\mathrm{P}(S_{n_1} \in E_{y_1'}, \cdots, S_{n_k} \in E_{y_k'}, S_n = 0) \\ (17) \quad &= \left(\frac{k}{2\pi n}\right)^{(k+1)/2} \frac{1}{2^{(k-1)/2}} \int_{E_{y_1'}} \cdots \int_{E_{y_k'}} f(\eta_1, \cdots, \eta_k) d\eta_1 \cdots d\eta_k \\ &\quad + \frac{k^2}{n} \theta Q + O\left(\frac{1}{n^{(k+2)/2}}\right) \end{aligned}$$

and

$$\begin{aligned} &\mathrm{P}\left(\frac{1}{n} \sum_{i=1}^{k} (n_i' - n_{i-1}') \phi(S_{n_i}) < \alpha\right) \\ &= \sum_{(y_1', y_2', \cdots, y_k') \in F_\alpha} \mathrm{P}(S_{n_1} \in E_{y_1'}, \cdots, S_{n_k} \in E_{y_k'}, S_n = 0) \\ &= \left(\frac{k}{n}\right)^{1/2} \left\{ \int_{F_\alpha} \psi \left[\left(\frac{2n}{k}\right)^{1/2} \eta_1, \left(\frac{2n}{k}\right)^{1/2} \eta_2, \cdots, \right. \right. \\ (18) \quad &\qquad \left. \left(\frac{2n}{k}\right)^{1/2} \eta_k \right] d(\eta_1 \cdots \eta_k) \\ &\quad \left. + \left(\frac{k}{n}\right)^{1/2} \theta Q \sum_{p=1}^{k} \alpha_p^3 \right\} + \frac{k^2}{n} \theta Q \\ &= \left(\frac{1}{n}\right)^{1/2} \left\{ \Phi_{n,k}(F_\alpha) + g(n, k) \right\}. \end{aligned}$$

But for large $n = 2n'$

$$\Phi_{n,k}(F_\alpha) + g(n, k) \cong \Phi_{n',k}(F_\alpha) + g(n', k).$$

5. **The limiting distribution.** We now refer back to the introduc-

tion of our paper, where we stated that in the case $P(X=1) = P(X=-1)=1/2$, we have (1). From this it follows that for $n=2n'$ and for any α, $0\leqq\alpha\leqq 1$,

$$P(N_n < n\alpha, S_n = 0) = P(S_n = 0)\frac{\lfloor n'\alpha\rfloor + 1}{n'+1}\cdot$$

We utilize the result (10) to obtain

$$\begin{aligned}\frac{[(\alpha-\delta)n']+1}{n'+1}\frac{c}{(2n')^{1/2}} - \frac{R(2n',\epsilon,k)}{\delta} - \frac{g(n',k)}{n^{1/2}} &\leqq \frac{1}{(2n')^{1/2}}\Phi_{n',k}(F_\alpha)\\ &\leqq \frac{[(\alpha+\delta)n']+1}{n'+1}\frac{c}{(2n')^{1/2}} + \frac{R(2n',\epsilon,k)}{\delta} + \frac{g(n',k)}{n^{1/2}}.\end{aligned}$$

From (11) and (18) it follows that

$$g(n,k) = \frac{\theta Qk}{n^{1/2}}\sum_{p=1}^{k}\alpha_p^3 + \frac{k^2}{n^{1/2}}\theta Q = \frac{\theta Qk}{n^{1/2}}Ck^4 + \frac{k^2}{n^{1/2}}\theta Q.$$

We recall that $k=n^{1/5}$, $\epsilon=k^{-1/3}$, $\delta=k^{-1/6}$; thus we have

$$\begin{aligned}\frac{R(n,\epsilon,k)}{\delta} &= \left(\frac{\log k}{k^{1/6}} + \frac{1}{k^{1/6}}\right)\frac{c}{n^{1/2}}\\ &= \left(\frac{\log n^{1/5}}{n^{1/30}} + \frac{1}{n^{1/30}}\right)\frac{c}{n^{1/2}}\cdot\end{aligned}$$

If the third moments γ equal zero, this gives

$$\frac{g(n,k)}{n^{1/2}} + \frac{R(n,\epsilon,k)}{\delta} = O\left(\frac{\log n^{1/5}}{n^{1/30}}\right)\frac{c}{n^{1/2}}\cdot$$

Whereas, if the third moments differ from zero, we stipulate instead that $k=n^{1/12}$, $\epsilon=k^{-1/3}$, $\delta=k^{-1/6}$ and obtain

$$\frac{g(n,k)}{n^{1/2}} + \frac{R(n,\epsilon,k)}{\delta} = O\left(\frac{\log n^{1/12}}{n^{1/72}}\right)\frac{c}{n^{1/2}}\cdot$$

Thus

$$\Phi_{n',k}(F_\alpha) = \frac{[\alpha n']+1}{n'+1} + O\left(\frac{\log n^{1/5}}{n^{1/30}}\right) \text{ in the first case,}$$

and

$$\Phi_{n',k}(F_\alpha) = \frac{[\alpha n'] + 1}{n' + 1} + O\left(\frac{\log n^{1/12}}{n^{1/72}}\right) \text{ in the second case.}$$

We use (10) once more and conclude that for any sequence of r.v. satisfying the conditions of the theorem

$$P(N_n \leqq \alpha n, S_n = 0) = \frac{[\alpha n] + 1}{n + 1} + O\left(\frac{\log n}{n^{1/30}}\right)$$

or

$$P(N_n \leqq \alpha n, S_n = 0) = \frac{[\alpha n] + 1}{n + 1} + O\left(\frac{\log n}{n^{1/72}}\right).$$

Q.E.D.

References

1. K. L. Chung and W. Feller, *On fluctuations in coin-tossing*, Proc. Nat. Acad. Sci. U.S.A. vol. 35 (1949) p. 605.

2. P. Erdös and M. Kac, *On the number of positive sums of independent random variables*, Bull. Amer. Math. Soc. vol. 53 (1947) p. 1011.

3. K. L. Chung, *An estimate concerning the Kolmogoroff limit distribution*, Trans. Amer. Math. Soc. vol. 67 (1949) p. 36.

4. G. G. Esseen, *Fourier analysis of distribution functions*, Acta Math. vol. 77 (1945) p. 1.

University of Maryland