AN ESTIMATE CONCERNING THE KOLMOGOROFF LIMIT DISTRIBUTION[1]

BY

KAI-LAI CHUNG

1. We consider a sequence of independent random variables having the common distribution function $F(x)$ which is assumed to be continuous. Let $nF_n(x)$ denote the number of random variables among the first n of the sequence whose values do not exceed x. Write

$$(1.1) \qquad d_n = \sup_{-\infty<x<\infty} \left| n(F_n(x) - F(x)) \right|.$$

Kolmogoroff [1][2] proved that the probability

$$(1.2) \qquad P(d_n \leqq \lambda n^{1/2}),$$

where λ is a positive constant, tends as $n \to \infty$ uniformly in λ to the limiting distribution

$$(1.3) \qquad \Phi(\lambda) = \sum_{-\infty}^{\infty} (-1)^j e^{-2j^2\lambda^2}.$$

Smirnoff [2] extended this result and recently Feller [3] has given new proofs of these theorems[3].

In this paper we shall obtain an estimate of the difference between (1.2) and (1.3) as a function of n, valid not only for λ equal to a constant but also for λ equal to a function $\lambda(n)$ of n which does not grow too fast. Since this estimate of the "remainder" will be of the order of magnitude of a negative power of n, it is futile to consider $\lambda(n)$ which is beyond the order of magnitude of lg n. In fact, a glance at (1.3) will show that already for $\lambda(n) = \lg n$ we have $\Phi(\lambda)$ differ from 1 by a term of an order of magnitude smaller than that of any negative power of n, thus smaller than our estimate of the remainder. For a similar reason it is also futile to consider $\lambda(n)$ whose order is less than $(\lg n)^{-1}$. Keeping these facts in mind we state our result as follows:

THEOREM 1. *If $A_0 > 0$ is an arbitrary constant and*

$$(1.4) \qquad (A_0 \lg n)^{-1} \leqq \lambda(n) \leqq A_0 \lg n,$$

Presented to the Society, September 10, 1948; received by the editors June 14, 1948.

(1) Research in connection with an ONR project.

(2) Numbers in brackets refer to the references cited at the end of the paper.

(3) More recently, Doob and Kac, independently, have simplified and amplified the matter (oral communication). However, none of these authors considered the error term, or in other words, a fixed large number n without the passage to infinity.

we have

$$(1.5)\quad \left| P(d_n \leqq \lambda(n)n^{1/2}) - \Phi(\lambda(n)) \right| \leqq An^{-1/10}\left\{1 + (\lg n)^{1/2}(\lambda^{-1}(n) + \lambda^{-2}(n))\right\}$$

where A is a constant depending only on A_0.

Henceforth we shall think of A_0 as fixed, for example, 100; then A will simply be a "universal" constant. The form of the estimate can be varied to a certain extent, but it is believed that the present form is about the best obtainable without essential improvement of the method. The rather clumsy situation of having several terms in an estimate is unavoidable if we want to include both ends of the range of $\lambda(n)$. Clearly at a small sacrifice we may replace the right side of (1.5) by $An^{-1/10}(\lg n)^{5/2}$.

We wish to point out that we shall really prove a more general theorem about the so-called "lattice distributions" which is embodied in formula (5.9) and the remark following it.

Our estimate can undoubtedly be improved upon but the limitations of our method are such that it is improbable that we can obtain the best possible result. However, Theorem 1 is amply sufficient for proving the following "strong" theorem.

THEOREM 2(4). *Let $\lambda(n) \uparrow \infty$. Then ("i.o." standing for "infinitely often")*

$$P(d_n > \lambda(n)n^{1/2} \text{ i.o.}) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$\sum \frac{\lambda^2(n)}{n} e^{-2\lambda^2(n)} \begin{Bmatrix} < \\ = \end{Bmatrix} \infty.$$

In particular, for any integer $p \geqq 3$

$$P(d_n > (2^{-1}n)^{1/2}(\lg_2 n + 2\lg_3 n + \lg_4 n + \cdots + \lg_p n + (1+\delta)\lg_{p+1} n)^{1/2} \text{ i.o.}) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$\delta \begin{Bmatrix} > \\ \leqq \end{Bmatrix} 0.$$

The second part of Theorem 2 follows of course from the first part by taking the appropriate sequence and using the Abel-Dini theorem.

The method of proving Theorem 2 by means of Theorem 1 follows the

(4) The idea of considering $\lambda(n)n^{1/2}$ as belonging to the upper or lower class is due to P. Lévy [11].

same pattern as developed by Feller in [4]. To avoid too much repetition even of an excellent thing we refrain from entering into a complete proof of Theorem 2. We shall, however, give a short proof for the particular result corresponding to the original form of the law of the iterated logarithm, due to Khintchine and Kolmogoroff (see [5]). Our reason for doing this is mainly didactic, in order to show how easily a result on the asymptotic distribution *with a suitable remainder* can be used to derive the corresponding strong result.

THEOREM 2*. *We have*

$$P\left(\limsup_{n\to\infty} \frac{d_n}{(2^{-1}n \lg_2 n)^{1/2}} = 1\right) = 1.$$

2. To prove Theorem 1 we shall follow Kolmogoroff's steps up to the last stage where he reduced the problem to one of addition of independent random variables of a special kind. More precisely he proved the following statement (with a different normalization).

Let $\{Y_j\}$, $j=1, \cdots, n$, be independent random variables having the common distribution given below:

$$P(Y_j = i - 1) = \frac{1}{i!} e^{-1}, \qquad i = 0, 1, 2, \cdots, \tag{2.1}$$

and let

$$T_n = \sum_{j=1}^{n} Y_j, \qquad T_n^* = \max_{1 \leqq j \leqq n} |T_j|.$$

Putting

$$P_n = P(T_{n-1}^* \leqq \lambda n^{1/2}, T_n = 0) \tag{2.2}$$

we have the equality

$$P(d_n \leqq \lambda n^{1/2}) = \frac{n! e^n}{n^n} P_n. \tag{2.3}$$

Thus the problem of finding $\Phi_n(\lambda)$ is reduced to that of finding P_n. The asymptotic value of P_n as $n\to\infty$ was obtained by Kolmogoroff by means of a general limit theorem of his employing partial differential equations. We shall use instead a method based in the combinatorial ideas of Erdös-Kac [6] on the one hand and the analytical tools of Esseen [7] on the other hand. It is similar to the one developed in my paper [8] but is adapted for discrete probabilities. Roughly speaking, we shall first show that the asymptotic value of (2.2) is independent of the nature of the distribution of the Y's so long as it satisfies certain general conditions. To be precise, let $\{X_j\}$,

$j=1, \cdots, n$, be independent random variables having the common distribution function $G(x)$ which has the following properties.

(i) $G(x)$ is a step function having all its (positive) jumps at integer points including 0 and such that the minimum distance between the abscissa of two jumps is 1.

(ii) The first moment of $G(y)$ is 0, the second is 1, and the fourth absolute moment is finite.

The following result is due to Esseen (p. 63 in [7]).

LEMMA 1. *Let* $\{X_j\}, j=1, \cdots, n$, *be independent random variables having the common distribution* $G(x)$ *which satisfies the condition* (i) *and* (ii). *Let*

$$S_n = \sum_{j=1}^{n} X_j$$

and let ξ *be such that* $\xi n^{1/2}$ *is a possible value of* S_n (*that is, the abscissa of a jump of the distribution function of* S_n). *Then*

$$(2.4)\quad \begin{aligned} P(S_n = \xi n^{1/2}) = \frac{1}{(2\pi n)^{1/2}}\Bigg\{\phi(\xi) - \frac{\alpha_3}{6n^{1/2}}\phi^{(3)}(\xi) \\ + \frac{1}{n}\left(\frac{\alpha_4 - 3}{24}\phi^{(4)}(\xi) + \frac{\alpha_3^2}{72}\phi^{(6)}(\xi)\right)\Bigg\} + o\left(\frac{1}{n}\right), \end{aligned}$$

where

$$\phi(\xi) = e^{-\xi^2/2}, \qquad \phi^{(i)}(\xi) = \frac{d^i}{d\xi^i}\phi(\xi)$$

and the remainder term $o(1/n)$ *is uniform with respect to* ξ.

In particular

$$P(S_n = \xi n^{1/2}) = \frac{1}{n^{1/2}}\phi(\xi) + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n^{1/2}}\right)$$

where the O-terms are uniform with respect to ξ.

3. We think of n as a very large number and define

$$k = k(n) = [n^{1/5}].$$

Conformably with the statement in Theorem 1, we take $\lambda = \lambda(n)$ to be a function of n not exceeding lg n. We put also

$$(3.1)\qquad \epsilon = \epsilon(n) = \frac{2(\lg n)^{1/2}}{\lambda(n) n^{1/10}}.$$

Further we put

$$n_i = \left[\frac{i}{k} n\right], \qquad i = 1, \cdots, k.$$

With the notation

$$S_n^* = \max_{1 \leqq j \leqq n} |S_j|$$

we write

$$\begin{aligned}
&P(S_{n-1}^* > \lambda n^{1/2}, S_n = 0) \\
&\quad = \sum_{r=1}^{n_{k-1}-1} P(S_{r-1}^* \leqq \lambda n^{1/2}, |S_r| > \lambda n^{1/2}, S_n = 0, |S_{n_{i+1}} - S_r| < \epsilon\lambda n^{1/2}) \\
&\quad + \sum_{r=1}^{n_{k-1}-1} P(S_{r-1}^* \leqq \lambda n^{1/2}, |S_r| > \lambda n^{1/2}, S_n = 0, |S_{n_{i+1}} - S_r| \geqq \epsilon\lambda n^{1/2}) \\
&\quad + \sum_{r=n_{k-1}}^{n} P(S_{r-1}^* \leqq \lambda n^{1/2}, |S_r| > \lambda n^{1/2}, S_n = 0) \\
&\quad = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3 \text{ say,}
\end{aligned}$$

where the n_{i+1} corresponding to each r is defined by $n_i \leqq r < n_{i+1}$.

A moment's reflection shows that

$$\sum\nolimits_1 \leqq P\left(\max_{1 \leqq i \leqq k-1} |S_{n_i}| > (1 - \epsilon)\lambda n^{1/2}, S_n = 0\right). \tag{3.2}$$

Next we may write

$$\begin{aligned}
\sum\nolimits_2 = \sum_{r=1}^{n_{k-1}-1} P(S_{r-1}^* \leqq \lambda n^{1/2}, |S_r| > \lambda n^{1/2}) \sum_y P(|S_{n_{i+1}} - S_r| \\
\geqq \epsilon\lambda n^{1/2}, S_{n_{i+1}} = y) P(S_n - S_{n_{i+1}} = -y)
\end{aligned} \tag{3.3}$$

where y runs through certain (integral) values depending on the value of S_r.

We shall use Q to denote a changeable positive constant depending only on $G(x)$, and A a changeable positive universal constant. From (2.4) we infer that

$$P(S_n - S_{n_{i+1}} = -y) \leqq Q(n - n_{i+1})^{-1/2} \leqq Qk^{1/2}n^{-1/2}.$$

It follows from (3.1) that

$$\sum\nolimits_2 \leqq Qk^{1/2}n^{-1/2} \max_r P(|S_{n_{i+1}} - S_r| \geqq \epsilon\lambda n^{1/2}). \tag{3.4}$$

To estimate the last-written probability we use Tchebecheff's inequality for small values of $n - n_{i+1}$ and the central limit theorem with a remainder (due to A. C. Berry [9] and Esseen [7]) for large values of $n-n_{i+1}$. More precisely, let

$$g = (n\epsilon^2\lambda^2)^{2/3}.$$

Then if $n_{i+1}-r\leqq g$, we have

$$(3.5)\qquad P(|S_{n_{i+1}} - S_r| \geqq \epsilon\lambda n^{1/2}) \leqq \frac{n_{i+1}-r}{n\epsilon^2\lambda^2} \leqq \frac{g}{n\epsilon^2\lambda^2} \leqq \left(\frac{1}{n\epsilon^2\lambda^2}\right)^{1/3}.$$

If $n_{i+1}-r=h>g$, we have

$$P(|S_{n_{i+1}} - S_r| \geqq \epsilon\lambda n^{1/2}) \leqq \left(\frac{2}{\pi}\right)^{1/2}\int_v^\infty e^{-u^2/2}du + Qg^{-1/2}$$

where

$$v = \frac{\epsilon\lambda n^{1/2}}{(n_{i+1}-r)^{1/2}} \geqq \epsilon\lambda n^{1/2}\left(\frac{n}{k}+1\right)^{-1/2}.$$

By the choice of ϵ and k, we have $v\rightarrow\infty$ as $n\rightarrow\infty$. It follows from a well known inequality that

$$(3.6)\qquad P(|S_{n_{i+1}} - S_r| \geqq \epsilon\lambda n^{1/2}) \leqq \frac{A}{\epsilon\lambda k^{1/2}}\exp\left(-\frac{k\epsilon^2\lambda^2}{3}\right) + Q\left(\frac{1}{n\epsilon^2\lambda^2}\right)^{1/3}.$$

In view of (3.5) we conclude that (3.6) holds in general. Hence it follows from (3.4) that

$$(3.7)\qquad \sum\nolimits_2 \leqq Qk^{1/2}n^{-1/2}\left\{\frac{1}{\lambda\epsilon k^{1/2}}\exp\left(-\frac{k\epsilon^2\lambda^2}{3}\right)+\left(\frac{1}{n\epsilon^2\lambda^2}\right)^{1/3}\right\} \leqq Qn^{-2/3}$$

if we substitute the values of k, ϵ, and λ.

Finally, we have

$$(3.8)\qquad \begin{aligned} \sum\nolimits_3 &= \sum_{r=n_{k-1}}^{n}\sum_{|y|>\lambda n^{1/2}} P(S^*_{r-1}\leqq \lambda n^{1/2}, S_r = y). \\ P(S_n - S_r = -y) &\leqq \max_{|y|>\lambda n^{1/2}} P(S_n - S_r = -y). \end{aligned}$$

Similarly to the above, let $h=(n\lambda^2)^{2/5}$. If $n-r\leqq h$, then

$$P(S_n - S_r = -y) \leqq \frac{n-r}{y^2} \leqq \frac{h}{n\lambda^2} \leqq (n\lambda^2)^{-3/5}.$$

If $n-r>h$ we apply Lemma 1 and obtain

$$P(S_n - S_r = -y) = Q(\xi) + o(h^{-3/2})$$

where $Q(\xi)$ stands for the expression involving $\phi(\xi)$ and its derivatives on the right-hand side of (2.4), with $\xi = y(n-r)^{-1/2}$.

Since $|y| > \lambda n^{1/2}$

$$\frac{y}{(n-r)^{1/2}} > \frac{\lambda n^{1/2}}{(n/k)^{1/2}} = k^{1/2}\lambda \geqq \frac{n^{1/10}}{A_0 \lg n},$$

we see that $Q(\xi)$ is negligible in order of magnitude and thus we may write

$$P(S_n - S_r = -y) \leqq Qh^{-3/2} \leqq Q(n\lambda^2)^{-3/5}.$$

(Notice that here Q actually depends on the constant A_0 in (1.4), as $Q(\xi)$ does, but we have agree to fix A_0.) Altogether we conclude from (3.8) that

$$\sum\nolimits_3 \leqq Q(n\lambda^2)^{-3/5}. \tag{3.9}$$

Continuing (3.2), (3.7), and (3.9) we obtain, recalling the order of magnitude of λ,

$$P(S_{n-1}^* > \lambda n^{1/2}, S_n = 0) \leqq P\left(\max_{1 \leqq i \leqq k-1} |S_{n_i}| > (1-\epsilon)\lambda n^{1/2}, S_n = 0\right) + Q(n\lambda^2)^{-3/5}.$$

Equivalently we may write

$$P(S_{n-1}^* \leqq \lambda n^{1/2}, S_n = 0) \geqq P\left(\max_{1 \leqq i \leqq k-1} |S_{n_i}| \leqq (1-\epsilon)\lambda n^{1/2}, S_n = 0\right) - Q(n\lambda^2)^{-3/5}. \tag{3.10}$$

4. Our next step is to approximate

$$P(|S_{n_1}| \leqq x_1, \cdots, |S_{n_{k-1}}| \leqq x_{k-1}, S_n = z)$$

by the corresponding probability associated with certain "discrete normal distributions," the meaning of which will be clear in a moment. We state the following lemma.

LEMMA 2. *Let S_{n_i}, $i = 1, \cdots, k$, be as in the preceding paragraphs and let $x_1, \cdots, x_{k-1}$ be arbitrary positive numbers, z a possible value of S_{n_k}. We write*

$$p_\xi = \left(\frac{k}{2\pi n}\right)^{1/2} \phi\left(\left(\frac{n}{k}\right)^{1/2} \xi\right)$$

where ϕ is defined in Lemma 1. Then we have

$$P(|S_{n_1}| \leqq x_1, \cdots, |S_{n_{k-1}}| \leqq x_{k-1}, S_{n_k} = z) = \sum_{|t_{k-1}| \leqq x_{k-1}} \cdots \sum_{|t_1| \leqq x_1} p_{t_1} p_{t_2 - t_1} \cdots p_{z - t_{k-1}} + \theta Q \frac{k^2}{n} \tag{4.1}$$

where $|\theta| \leqq 1$.

Proof. We shall use θ as a changeable constant such that $|\theta| \leqq 1$. From Lemma 1 we have

$$P(S_{n_1} = z) = p_z + \theta Q\, k/n$$

since $n_1 = k/n + \theta$. Assuming that

$$P(|S_{n_1}| \leqq x_1, \cdots, |S_{n_{i-1}}| \leqq x_{i-1}, S_{n_i} = z)$$
$$= \sum_{|t_{i-1}| \leqq x_{i-1}} \cdots \sum_{|t_1|_1 \leqq x} p_{t_1} p_{t_2 - t_1} \cdots p_{z - t_{i-1}} + \theta Q \frac{ik}{n}$$

we shall show that the same as is true when we replace i by $i+1$. In fact

$$P(|S_{n_1}| \leqq x_1, \cdots, |S_{n_i}| \leqq x_i, S_{n_{i+1}} = z)$$
$$= \sum_{|t_i| \leqq x_i} P(|S_{n_1}| \leqq x_1, \cdots, |S_{n_{i-1}}| \leqq x_{i-1}, S_{n_i} = t_i) p_{z-t_i} + \theta Q \frac{k}{n}$$
$$= \sum_{|t_i| \leqq x_i} \cdots \sum_{|t_1| \leqq x_1} p_{t_1} p_{t_2 - t_1} \cdots p_{z - t_i} + \theta Q \frac{(i+1)k}{n}.$$

Thus Lemma 2 is proved by induction on i.

Taking $x_1 = \cdots = x_{k-1} = (1-\epsilon)\lambda n^{1/2}$, $z = 0$ in (4.1) we have

$$P\left(\max_{1 \leqq i \leqq k-1} |S_{n_i}| \leqq (1-\epsilon)\lambda n^{1/2}, S_n = 0\right)$$
$$= \sum_{|t_i| \leqq (1-\epsilon)\lambda n^{1/2}, 1 \leqq i \leqq k-1} \cdots \sum p_{t_1} p_{t_2 - t_1} \cdots p_{-t_{k-1}} + \theta Q \frac{k^2}{n}$$
$$= \Psi(1-\epsilon)\lambda) + \theta Q n^{-3/5}$$

where Ψ is defined by the multiple sum.

Combining this with (3.10) we obtain

$$P(S^*_{n-1} \leqq \lambda n^{1/2}, S_n = 0) \geqq \Psi((1-\epsilon)\lambda) - Q n^{-3/5}(1 + \lambda^{-6/5}). \tag{4.2}$$

On the other hand it is obvious that

$$P(S^*_{n-1} \leqq \lambda n^{1/2}, S_n = 0) \leqq P\left(\max_{1 \leqq i \leqq k-1} |S_{n_i}| \leqq \lambda n^{1/2}, S_n = 0\right)$$
$$\leqq \Psi(\lambda) + Q n^{-3/5}(1 + \lambda^{-6/5})$$

by Lemma 2.

Together we have

$$\begin{aligned} \Psi((1-\epsilon)\lambda) - Q n^{-3/5}(1 + \lambda^{-6/5}) &\leqq P(S^*_{n-1} \leqq \lambda n^{1/2}, S_n = 0) \\ &\leqq \Psi(\lambda) + Q n^{-3/5}(1 + \lambda^{-6/5}). \end{aligned} \tag{4.3}$$

This being true no matter what the X_j's are provided that their distribution $G(x)$ satisfies (i) and (ii) of §2, we can estimate the Ψ's in (4.3) by evaluating approximately the probability in the middle of (4.3) for a special case. We shall make use of the classical Bernoullian distribution and a combinatorial formula given by Bachelier.

5. LEMMA 3. *Let* $\{X_j\}$, $j=1, \cdots, n$, *be independent random variables having the following distribution*

$$X_j = \begin{cases} +1 & \textit{with probability } 1/2 \\ -1 & \textit{with probability } 1/2. \end{cases} \tag{5.1}$$

Suppose that n is even, and

$$(A_0 \lg n)^{-1} \leqq \lambda(n) \leqq A_0 \lg n. \tag{5.2}$$

Then as $n \to \infty$

$$P(S_n^* \leqq \lambda(n)n^{1/2}, S_n = 0) = \left(\frac{2}{\pi n}\right)^{1/2} \Phi(\lambda_n) + O\left(\frac{(\lg n)^{11}}{n}\right) \tag{5.3}$$

where the O term does not depend on $\lambda(n)$ *if the constant* A_0 *in* (5.2) *is fixed.*

Proof. If n is even and b is an integer, the formula of Bachelier [10, pp. 252–253] may be written as follows:

$$P(S_n^* < b, S_n = 0) = \frac{1}{2^n}\left\{\binom{n}{n/2} + 2\sum_{1 \leqq j \leqq n/2b} (-1)^j \binom{n}{n/2 - jb}\right\}.$$

Applying Stirling's formula with a remainder term we find after some routine calculations that if $0 \leqq j \leqq (\lg n)^2$, $b = \lambda_n n^{1/2}$

$$\frac{1}{2^n}\binom{n}{n/2 - jb} = \left(\frac{2}{\pi n}\right)^{1/2} e^{-2j^2\lambda_n^2}\left\{1 + O\left(\frac{j^3\lambda_n^3}{n^{1/2}}\right)\right\}.$$

On the other hand, by a well known estimate concerning the binomial distribution, we have

$$\sum_{(\lg n)^2 < j \leqq n^{1/2}/2\lambda_n} \frac{1}{2^n}\binom{n}{n/2 - jb} \leqq e^{-A\lambda_n^2(\lg n)^4} \leqq e^{-A(\lg n)^2}$$

which is of a smaller order of magnitude than any negative power of n. Thus

$$\begin{aligned} P(S_n^* < \lambda_n n^{1/2}, S_n = 0) &= \left(\frac{2}{\pi n}\right)^{1/2}\left\{1 + 2\sum_{j=1}^{\infty} (-1)^j e^{-2j^2\lambda_n^2} \right. \\ &\left. + O\left(\frac{(\lg n)^8}{n^{1/2}}\lambda_n^3\right)\right\} + O\left(\frac{1}{n}\right) = \left(\frac{2}{\pi n}\right)^{1/2}\Phi(\lambda_n) + O\left(\frac{(\lg n)^{11}}{n^{1/2}}\right). \end{aligned} \tag{5.4}$$

Notice that here the O-term depends on the A_0 in (5.2).

This is the desired result (5.3) except that we have to replace the strict "$<$" in the probability on the left by "$\leqq$" and also remove the restriction that $\lambda_n n^{1/2}$ is an integer. Now if we replace $\lambda_n n^{1/2}$ by $\lambda_n n^{1/2}+\theta=\rho_n n^{1/2}$ we have

$$\rho_n - \lambda_n = \theta_n^{-1/2}.$$

Differentiating $\Phi(\lambda)$ we have

$$\Phi'(\lambda) = 4\lambda \sum_{j=-\infty}^{\infty} (-1)^j j^2 e^{-2j^2\lambda^2} = -4\lambda\left(\sum_{j^2 \leqq 2^{-1}\lambda^{-2}} + \sum_{j^2 > 2^{-1}\lambda^{-2}}\right).$$

The maximum of $x^2 e^{-2x^2\lambda^2}$ being $2^{-1}e^{-1}\lambda^{-2}$ the first sum is $\leqq A\lambda^{-3}$. The second sum is dominated by its first term, hence $\leqq A\lambda^{-2}$. Thus

$$|\Phi'(\lambda)| \leqq A\left(\frac{1}{\lambda} + \frac{1}{\lambda^2}\right). \tag{5.5}$$

It follows that

$$|\Phi(\rho_n) - \Phi(\lambda_n)| \leqq A\left(\frac{1}{\lambda_n} + \frac{1}{\lambda_n^2}\right)\frac{1}{n^{1/2}} \leqq A\,\frac{(\lg n)^2}{n^{1/2}}$$

by (5.2). Comparing (5.4) with the same formula with ρ_n replacing λ_n we see that

$$P(S_{n-1}^* \leqq \lambda_n n^{1/2}, S_n = 0) - P(S_{n-1}^* < [\lambda_n n^{1/2}], S_n = 0) \leqq A\,\frac{(\lg n)^2}{n}$$

where $\lambda_n n^{1/2}$ need not be an integer now. This and (5.4) establish (5.3).

The distribution (5.1) does not satisfy (i) of §4. To remedy this we put

$$Z = \frac{X_1 + X_2 + X_3 + X_4}{2} \tag{5.6}$$

where the form X's are independent random variables having the distribution (5.1). Then Z has a distribution satisfying (i) and (ii) of §4. Now let $Z_j, j=1, \cdots, n$, be n independent random variables each distributed as Z in (5.6). Form

$$W_n = \sum_{j=1}^{n} Z_j, \qquad W_n^* = \max_{1\leqq k\leqq n} |W_k|.$$

It is easy to see that we have, the S's referring to partial sums of the X's in (5.1),

$$P(S_{4n-1}^* \leqq 2x, S_{4n} = 0) \leqq P(W_{n-1}^* \leqq x, W_n = 0) \leqq P(S_{4n-1}^* \leqq 2x + 3, S_{4n} = 0).$$

By the same reasoning as in the last paragraph we see that if $x=\lambda_n n^{1/2}$ the difference between the extreme terms in the last inequality is

$$\leqq \frac{A}{n}\left(\frac{1}{\lambda_n}+\frac{1}{\lambda_n^2}\right).$$

Therefore we obtain

$$(5.7)\qquad \begin{aligned} P(W_{n-1}^* \leqq \lambda_n n^{1/2},\, W_n = 0) &= P(S_{4n-1}^* \leqq 2\lambda_n n^{1/2},\, S_{4n} = 0) + O\left(\frac{(\lg n)^2}{n}\right) \\ &= \left(\frac{1}{2\pi n}\right)^{1/2} \Phi(\lambda_n) + O\left(\frac{(\lg n)^{11}}{n}\right) \end{aligned}$$

by (5.3).

Since the distribution of the Z's satisfies (i) and (ii) of §4, we can substitute (5.7) for the middle term in (4.3). Thus we obtain

$$\begin{aligned} \Psi((1-\epsilon_n)\lambda_n) - Qn^{-3/5}(1+\lambda^{-6/5}) &\leqq \frac{1}{(2\pi n)^{1/2}}\Phi(\lambda_n) \\ &\leqq \Psi(\lambda_n) + Qn^{-3/5}(1+\lambda^{-6/5}), \end{aligned}$$

the error term in (5.7) being absorbed into the one in (4.3).

Hence

$$\Psi(\lambda_n) \leqq \frac{1}{(2\pi n)^{1/2}}\Phi\left(\frac{\lambda_n}{1-\epsilon_n}\right) + Qn^{-3/5}(1+\lambda^{-6/5})$$

$$\Psi((1-\epsilon_n)\lambda_n) \geqq \frac{1}{(2\pi n)^{1/2}}\Phi((1-\epsilon_n)\lambda_n) - Qn^{-3/5}(1+\lambda^{-6/5});$$

substituting these into (4.3) we obtain

$$(5.8)\qquad \begin{aligned} \Phi((1-\epsilon_n)\lambda_n) - Qn^{-3/5}(1+\lambda^{-6/5}) &\leqq (2\pi n)^{1/2} P(S_{n-1}^* \leqq \lambda_n n^{1/2},\, S_n = 0) \\ &\leqq \Phi\left(\frac{\lambda_n}{1-\epsilon_n}\right) + Qn^{-3/5}(1+\lambda^{-6/5}). \end{aligned}$$

Now from (5.5) we obtain

$$\begin{aligned} \Phi\left(\frac{\lambda_n}{1-\epsilon_n}\right) - \Phi(\lambda_n) &\leqq A\epsilon_n\lambda_n\left(\frac{1}{\lambda_n}+\frac{1}{\lambda_n^2}\right) \\ &\leqq \frac{A(\lg n)^{1/2}}{n^{1/10}}\left(\frac{1}{\lambda_n}+\frac{1}{\lambda_n^2}\right) \end{aligned}$$

by (3.1). The same estimate holds for $\Phi(\lambda_n)-\Phi((1-\epsilon_n)\lambda_n)$. Hence observing that $\lambda_n^{-6/5}$ is between λ_n^{-1} and λ_n^{-2} we can write (5.8) as

$$(5.9)\qquad \begin{aligned}&(2\pi n)^{1/2}P(S_{n-1}^* \leqq \lambda_n n^{1/2}, S_n = 0)\\ &\qquad = \Phi(\lambda_n) + Qn^{-1/10}\{1 + (\lg n)^{1/2}(\lambda_n^{-1} + \lambda_n^{-2})\}.\end{aligned}$$

This result (5.9) holds in general for any independent random variables X_j, $j=1, \cdots, n$, whose common distribution function $G(x)$ satisfies (i) and (ii) of §4. In particular it holds for the Y_j's defined in (2.1). Referring to (2.2) and changing Q into A we have therefore

$$(5.10)\qquad (2\pi n)^{1/2}P_n = \Phi(\lambda_n) + An^{-1/10}\{1 + (\lg n)^{-1/2}(\lambda_n^{-1} + \lambda_n^{-2})\}.$$

Using Stirling's formula in (2.3) we obtain

$$P(d_n \leqq \lambda_n n^{1/2}) = (2\pi n)^{1/2}\left(1 + \frac{\theta}{12n}\right)P_n$$

which reduces to (1.5) by means of (5.10). Theorem 1 is proved.

We state two simple corollaries of Theorem 1 which we shall need in the next section.

COROLLARY 1. *There exists a universal constant $C>0$ such that for every n we have*

$$P(d_n \leqq n^{1/2}) \geqq C.$$

This serves the purpose of Tchebycheff's inequality and follows already from Kolmogoroff's theorem without the remainder term. For all finite n the probability is a positive number; for all $n \geqq n_0(\epsilon)$ the probability is $\geqq \Phi(1)-\epsilon$ where ϵ may be taken to be any positive number $<\Phi(1)$.

COROLLARY 2. *For the range*

$$\lambda(n) \sim C_0 \lg_2 n$$

we have

$$C_1 e^{-2\lambda^2(n)} \leqq P(d_n \geqq \lambda(n)n^{1/2}) \leqq C_2 e^{-2\lambda^2(n)}$$

where the C's are positive constants, C_1 and C_2 depending on C_0.

In this range $1-\Phi(\lambda)$ is dominated by the term $2e^{-2\lambda^2}$ and the remainder term is negligible in comparison.

6. If $n<m$ we define $(m-n)F_{m,n}(x)$ to be the number of X_k, $n<k\leqq m$, whose values do not exceed x. Recalling the definition of $nF_n(x)$ in §1 we have the relation

$$(6.1)\qquad nF_n(x) + (m-n)F_{n,m}(x) = mF_m(x).$$

Given $\eta>0$ let $\alpha>1$ be such that

$$\frac{1+\eta}{\alpha^2} \geqq 1+\frac{\eta}{2}.$$

Put $n_k=[\alpha^k]$.

Suppose that $n_k \leqq n < n_{k+1}$. Consider the events

$$E_n \qquad d_n = \sup_x \left| n(F_n(x) - F(x)) \right| \geqq (1+\eta)(2^{-1}n \lg_2 n)^{1/2}$$

and

$$E_{n,n_{k+1}} \qquad \sup_x \left| (n_{k+1}-n)(F_{n,n_{k+1}}(x) - F(x)) \right| \leqq (n_{k+1}-n)^{1/2}.$$

By (6.1) the two events E_n and $E_{n,n_{k+1}}$ together imply

$$\sup_x \left| n_{k+1}(F_{n_{k+1}}(x) - F(x)) \right| \geqq (1+\eta)(2^{-1}n \lg_2 n)^{1/2} - (n_{k+1}-n)^{1/2}$$

$$(6.2) \qquad = (1+\eta)(2^{-1}n_{k+1} \lg_2 n_{k+1})^{1/2}\left\{\left(\frac{n_k \lg_2 n_k}{n_{k+1}\lg_2 n_{k+1}}\right)^{1/2} - \frac{1}{1+\eta}\left(\frac{n_{k+1}-n_k}{2^{-1}n_{k+1}\lg_2 n_{k+1}}\right)^{1/2}\right\}.$$

As $k\to\infty$ the expression within the braces tends to α^{-1}, thus is not less than α^{-2} for sufficiently large k. Then (6.2) implies the following event:

$$F_{n_{k+1}} \qquad \sup \left| n_{k+1}(F_{n_{k+1}}(x) - F(x)) \right| \geqq \left(1+\frac{\eta}{2}\right)(2^{-1}n_{k+1}\lg_2 n_{k+1})^{1/2}$$

by the choice of α.

From Corollary 1 to Theorem 1 we have

$$P(E_{n,n_{k+1}}) \geqq C.$$

This together with the fact that $E_{n,n_{k+1}}$ is independent of all E_k for $k<n$ gives

$$P\left(\sum_{n=n_k}^{n_{k+1}-1} E_n\right) \leqq C^{-1}P(F_{n_{k+1}})$$

(see Feller [4, p. 394]). According to Corollary 2 to Theorem 1,

$$P(F_{n_{k+1}}) \leqq C_2(\lg n_k)^{-(1+\eta/2)} \leqq C_3 k^{-(1+\eta/2)}$$

where $C_3>0$ is a constant depending only on α. Hence

$$\sum_k P\left(\sum_{n=n_k}^{n_{k+1}-1} E_n\right) \leqq C^{-1}\sum_k P(F_{n_{k+1}}) < \infty.$$

It follows from the Borel-Cantelli lemma that $P(E_n$ occurs i.o.$)=0$. In other words,

$$P\left(\limsup_{n\to\infty} \frac{d_n}{(2^{-1}n \lg_2 n)^{1/2}} < 1 + \eta\right) = 1. \tag{6.3}$$

Now let $\beta > 0$ be an integer such that

$$\left(1 - \frac{\eta}{2}\right)\left(\frac{\beta - 1}{\beta}\right)^{1/2} - \left(\frac{1}{\beta}\right)^{1/2} \geqq 1 - \frac{\eta}{2}.$$

Consider the events

$$H_k \qquad \sup_x \left| \beta^k(F_{\beta^k}(x) - F(x)) \right| \leqq (1 - \eta)(2^{-1}\beta^k \lg_2 \beta^k)^{1/2},$$

$$H_{k,k+1} \qquad \sup_x \left| (\beta^{k+1} - \beta^k)(F_{\beta^k,\beta^{k+1}}(x) - F(x)) \right|$$
$$\geqq \left(1 - \frac{\eta}{2}\right)(2^{-1}(\beta^{k+1} - \beta^k) \lg_2 (\beta^{k+1} - \beta^k))^{1/2}.$$

By (6.1) the two events H_k and $H_{k,k+1}$ together imply

$$\begin{aligned} &\sup_x \left| \beta^{k+1}(F_{\beta^{k+1}}(x) - F(x)) \right| \\ &\geqq \left(1 - \frac{\eta}{2}\right)(2^{-1}(\beta^{k+1} - \beta^k) \lg_2 (\beta^{k+1} - \beta^k))^{1/2} \\ &\quad - (1 - \eta)(2^{-1}\beta^k \lg_2 \beta^k)^{1/2} \\ &\geqq (2^{-1}\beta^{k+1} \lg_2 \beta^{k+1})^{1/2} \left\{ \left(1 - \frac{\eta}{2}\right)\left(\frac{\beta^{K+1} - \beta^k \lg_2 (\beta^{k+1} - \beta^k)}{\beta^{k+1} \lg_2 \beta^{k+1}}\right)^{1/2} \right. \\ &\quad \left. - \left(\frac{\beta^k \lg_2 \beta^k}{\beta^{k+1} \lg_2 \beta^{k+1}}\right)^{1/2} \right\}. \end{aligned} \tag{6.4}$$

As $k \to \infty$ the expression within the braces tends to

$$\left(1 - \frac{\eta}{2}\right)\left(\frac{\beta - 1}{\beta}\right)^{1/2} - \left(\frac{1}{\beta}\right)^{1/2}.$$

Thus for sufficiently large k, say for $k \geqq k_0$, (6.4) implies

$$\overline{H}_{k+1} \qquad \sup_x \left| \beta^{k+1}(F_{\beta^{k+1}}(x) - F(x)) \right| \geqq (1 - \eta)(2^{-1}\beta^{k+1} \lg_2 \beta^{k+1})^{1/2}$$

by the choice of β.

Since $H_{k,k+1}$ is independent of H_j for $j \leqq k$, it follows that

$$P\left(\prod_{j=K_0}^{k} H_k\right) P(H_{k,k+1}) \leqq P\left(\prod_{j=k_0}^{K} H_j \cdot \overline{H}_{k+1}\right)$$

or

$$P\left(\prod_{k=k_0}^{k} H_j\right) \leqq P\left(\prod_{k=k_0}^{k} H_j\right)(1 - P(H_{k,k+1})).$$

From Corollary 2 to Theorem 1 we have

$$P(H_{k,k+1}) \geqq C_1(\lg (\beta^{k+1} - \beta^k))^{-(1-\eta/2)} \geqq C_4 K^{-(1-\eta/2)}$$

where $(C_4 > 1)$ is a constant depending only on β. Thus

$$P\left(\prod_{k=k_0}^{\infty} H_j\right) \leqq \prod_{j=k_0}^{\infty} (1 - C_4 k^{-(1-\eta/2)}) = 0.$$

This means $P(\overline{H}_k$ occurs i.o.$) = 1$. In other words,

$$P\left(\limsup_{n\to\infty} \frac{d_n}{(2^{-1} n \lg_2 n)^{1/2}} > 1 - \eta\right) = 1. \tag{6.5}$$

Since η is arbitrary, (6.3) and (6.5) establish Theorem 2.

References

1. A. Kolmogoroff, *Sulla determinazione empirica di une legge di distribuzione*, Giornale dell'Istituto Italiano degli Attuari vol. 4 (1933) pp. 83–91.

2. N. Smirnoff, *Sin les écarts de la courbe de distribution empirique*, Rec. Math. (Mat. Sbornik) N.S. vol. 6 (1939) pp. 3–26.

3. W. Feller, *On the Kolmogoroff-Smirnoff limit theorems*, Ann. Math. Statist. vol. 19 (1948) pp. 177–189.

4. ———, *The general form of the so-called law of the iterated logarithm*, Trans. Amer. Math. Soc. vol. 54 (1943) pp. 373–402.

5. A. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Berlin, Springer, 1933.

6. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. vol. 52 (1946) pp. 292–302.

7. C. G. Esseen, *Fourier analysis of distribution functions*, Acta Math. vol. 77 (1945) pp. 1–125.

8. K. L. Chung, *On the maximum partial sums of sequences of independent random variables*, Trans. Amer. Math. Soc. vol. 64 (1948) pp. 205–233.

9. A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. Amer. Math. Soc. vol. 49 (1941) pp. 122–136.

10. L. Bachelier, *Calcul des probabilités*, vol. 1, Paris, 1912.

11. P. Lévy, *Théorie de l'addition des variables aléatoires*, Paris, Gautier-Villars, 1937.

Princeton University,
Princeton, N. J.
Cornell University,
Ithaca, N. Y.