

ERROR ESTIMATES FOR CERTAIN PROBABILITY LIMIT THEOREMS¹

BY J. M. SHAPIRO

Ohio State University

1. Summary and Introduction. Consider a sequence of independent random variables $x_1, x_2, \dots, x_k, \dots$ with mean 0 and variance σ_k^2 . Let $S_n = (x_1 + \dots + x_n)/s_n$ where $s_n^2 = \sigma_1^2 + \dots + \sigma_n^2$. The classical forms of the central limit theorem state that, with certain assumptions, the distribution function $F_n(x)$ approaches the Gaussian distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du.$$

Berry [1] and Esseen [3] have studied the behavior of

$$M_n = \sup_{-\infty < x < \infty} |F_n(x) - \Phi(x)|$$

and in their main theorems have obtained bounds on M_n which involve the moments of x_k through the third.

More generally consider a system of random variables $(x_{nk}), k = 1, 2, \dots, k_n; n = 1, 2, \dots$ such that for each n , the variables x_{n1}, \dots, x_{nk_n} are independent. Let $S_n = x_{n1} + \dots + x_{nk_n}$ and again let $F_n(x)$ be the distribution function of S_n . From a well known theorem of Khintchine [5] it follows that if the random variables x_{nk} are infinitesimal (i.e., $\lim_{n \rightarrow \infty} \max_{1 \leq k \leq k_n} P\{|x_{nk}| > \epsilon\} = 0$ for every $\epsilon > 0$) then the class of possible limiting distributions of $F_n(x)$ coincides with the class of infinitely divisible distributions.

Let $F(x)$ be any infinitely divisible distribution function and let $M_n = \sup_{-\infty < x < \infty} |F_n(x) - F(x)|$. In this paper we obtain bounds on M_n in the case where $F(x)$ and the x_{nk} have finite second moments. It is shown that under necessary and sufficient conditions for $F_n(x)$ to approach $F(x)$, the bounds on M_n obtained approach zero as n becomes infinite.

Throughout the paper, given the system (x_{nk}) we shall let $F_{nk}(x), \varphi_{nk}(t), \mu_{nk}$, and σ_{nk}^2 be the distribution function, characteristic function, mean, and variance respectively of x_{nk} , and $F_n(x), \varphi_n(t), \mu_n$, and σ_n^2 have the same meaning for the random variable S_n .

2. Some Preliminary Lemmas. The following lemmas will be used to obtain the general result in the next section.

LEMMA 1. *Let z_1 and z_2 be any two complex numbers such that $|z_1| \leq 1$ and $|z_2| \leq 1$; Then $|z_1 - z_2| \leq |\log z_1 - \log z_2|$.*

Received March 11, 1955.

¹ This paper is the revised form of a Doctor's dissertation accepted by the University of Minnesota, 1954, and presented to the American Mathematical Society December 29, 1954.

This follows from the mean value theorem for complex functions (see [2] page 115).

LEMMA 2. *Let*

$$f(t, x) = (e^{itx} - 1 - itx)^{1/x^2}$$

for real x and t . Then for all x , $|f(t, x)| \leq \frac{1}{2}t^2$ and $|\partial f(t, x)/\partial x| \leq \frac{5}{8}|t|^3$.

This follows from the fact that for real u ,

$$e^{iu} = \sum_{k=0}^{n-1} \frac{(iu)^k}{k!} + \theta \frac{u^n}{n!} \quad \text{where } |\theta| \leq 1.$$

COROLLARY. $|f(t, x) - f(t, y)| \leq \frac{5}{4}|t|^3|x - y|$.

PROOF. We have $f(t, x) = \cos(tx - 1)/x^2 + i \sin(tx - tx)/x^2 \equiv R(t, x) + iI(t, x)$. By the law of the mean we have $|R(t, x) - R(t, y)| \leq \frac{5}{8}|t|^3 \cdot |x - y|$ and the same inequality holds for $|I(t, x) - I(t, y)|$. Thus $|f(t, x) - f(t, y)| \leq \sqrt{2} \cdot \frac{5}{8} \cdot |t|^3 \cdot |x - y| < \frac{5}{4}|t|^3 \cdot |x - y|$.

Now let $F(x)$ be any infinitely divisible distribution function with mean μ , variance σ^2 , and characteristic function $\varphi(t)$. According to Kolmogorov's formula [6] for the characterization of infinitely divisible distributions with finite variance, we know that

$$(2.1) \quad \log \varphi(t) = i\mu t + \int_{-\infty}^{\infty} (e^{itx} - 1 - itx) \frac{1}{x^2} dG(x)$$

where $G(x)$ is a bounded nondecreasing function. If we impose that $G(-\infty) = 0$ and that $G(x)$ is right continuous then the representation of $\log \varphi(t)$ by this formula is unique. (Also if $G(-\infty) = 0$ then $G(+\infty) = \sigma^2$.)

Let $A > 0$ be such that $-A$ and A are continuity points of $G(x)$ and let $0 < \delta \leq 2A$. Define

$$(2.2) \quad m = m(A, \delta) = \left[\frac{2A}{\delta} \right] + 1$$

where $[r]$ is the greatest integer function. Let

$$(2.3) \quad -A = x_0 < x_1 < x_2 < \dots < x_m = A$$

be such that $x_i (i = 0, 1, 2, \dots, m)$ is a continuity point of $G(x)$ and

$$\max_{i=1,2,\dots,m} |x_i - x_{i-1}| < \delta.$$

Let

$$(2.4) \quad \sum_{k=1}^{k_n} \int_{-\infty}^x u^2 dF_{n_k}(u + \mu_{n_k}) = G_n(x)$$

and

$$(2.5) \quad E(n, t, m(A, \delta)) \equiv \frac{5}{4}\delta|t|^3(\sigma_n^2 + \sigma^2) + \frac{t^2}{2} \sum_{i=0}^m |G_n(x_i) - G(x_i)| + \frac{2|t|}{A} [G_n(+\infty) - G_n(A) + G(+\infty) - G(A) + G_n(-A) + G(-A)]$$

with this notation and with $f(t, x)$ defined as in Lemma 2 we have the following lemma.

LEMMA 3. $|\int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)]| \leq E(\nu, t, m(A, \delta))$ for any $A > 0$, $0 < \delta \leq 2A$ and any choice of x_0, x_1, \dots, x_m satisfying (2.3).

PROOF. First let

$$\xi_i = \begin{cases} x_i, & i \text{ odd} \\ x_{i-1}, & i \text{ even} \end{cases} \quad i = 1, \dots, m$$

and consider (for $n = 0, 1, \dots$ with $G(x) \equiv G_0(x)$),

$$\begin{aligned} & \left| \int_{-A}^A f(t, x) dG_n(x) - \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G_n(x_{i-1})] \right| \\ &= \left| \sum_{i=1}^m \int_{x_{i-1}}^{x_i} [f(t, x) - f(t, \xi_i)] dG_n(x) \right| \leq \frac{5}{4}|t|^3 \cdot \delta \cdot [G_n(x_m) - G_n(x_0)] \end{aligned}$$

by the corollary to Lemma 2. Now $G_n(x)$ is nonnegative and nondecreasing and $G_n(+\infty) = \sigma_n^2$ so that $[G_n(x_n) - G_n(x_0)] \leq \sigma_n^2$ where we define $\sigma_0^2 = \sigma^2$. Therefore

$$\left| \int_{-A}^A f(t, x) dG_n(x) - \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G_n(x_{i-1})] \right| \leq \frac{5}{4}|t|^3 \delta \sigma_n^2.$$

Now consider for $n \neq 0$

$$\begin{aligned} & \left| \int_{-A}^A f(t, x) dG_n(x) - \int_{-A}^A f(t, x) dG(x) \right| \\ &= \left| \int_{-A}^A f(t, x) dG_n(x) - \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G_n(x_{i-1})] \right| \\ (2.6) \quad &+ \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G_n(x_{i-1})] - \sum_{i=1}^m f(t, \xi_i)[G(x_i) - G(x_{i-1})] \\ &+ \sum_{i=1}^m f(t, \xi_i)[G(x_i) - G(x_{i-1})] - \int_{-A}^A f(t, x) dG(x) \Big| \\ &\leq \frac{5}{4}|t|^3 \delta (\sigma_n^2 + \sigma^2) + \left| \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G_n(x_{i-1}) - G(x_i) + G(x_{i-1})] \right|. \end{aligned}$$

Now from Lemma 2 $|f(t, x)| \leq \frac{1}{2}t^2$ so that,

$$\begin{aligned} & \left| \sum_{i=1}^m f(t, \xi_i)[G_n(x_i) - G(x_i) + G(x_{i-1}) - G_n(x_{i-1})] \right| \\ (2.7) \quad &\leq \frac{t^2}{2} \left[|G(x_0) - G_n(x_0)| + |G(x_m) - G_n(x_m)| \right. \\ &\quad \left. + 2 \sum_{i=1}^{\lfloor (m-2)/2 \rfloor} |G_n(x_{2i}) - G(x_{2i})| \right]. \end{aligned}$$

Now if we let

$$\xi_i = \begin{cases} x_i, & i \text{ even} \\ x_{i-1}, & i \text{ odd} \end{cases}$$

we see that (2.6) still holds and that (2.7) becomes

$$(2.8) \quad \left| \sum_{i=1}^m f(t, \xi_i) [G_n(x_i) - G(x_i) + G(x_{i-1}) - G_n(x_{i-1})] \right| \\ \leq \frac{t^2}{2} \left[|(G(x_0) - G_n(x_0))| + |G(x_m) - G_n(x_m)| \right. \\ \left. + 2 \sum_{i=1}^{[(m-1)/2]} |G_n(x_{2i-1}) - G(x_{2i-1})| \right].$$

Combining (2.6), (2.7), and (2.8) we find

$$(2.9) \quad \left| \int_{-A}^A f(t, x) dG_n(x) - \int_{-A}^A f(t, x) dG(x) \right| \\ \leq \frac{5}{4}|t|^3 \delta (\sigma_n^2 + \sigma^2) + \frac{t^2}{2} \cdot \sum_{i=0}^m |G_n(x_i) - G(x_i)|.$$

Consider now $\int_{-\infty}^{-A} + \int_A^{\infty} f(t, x) d[G_n(x) - G(x)]$. We note that

$$|f(t, x)| = \left| \frac{i}{x} \int_0^t (e^{itx} - 1) dt \right| \leq \frac{2|t|}{|x|}$$

so that

$$\left| \int_{-A}^{-\infty} + \int_A^{\infty} f(t, x) dG_n(x) \right| \leq 2|t| \int_{-\infty}^{-A} + \int_A^{\infty} \frac{1}{|x|} dG_n(x) \\ \leq \frac{2|t|}{A} [G_n(+\infty) - G_n(A) + G_n(-A)], \quad n = 0, 1, 2, \dots$$

Thus

$$(2.10) \quad \left| \int_{-\infty}^{-A} + \int_A^{\infty} f(t, x) d[G_n(x) - G(x)] \right| \\ \leq \frac{2|t|}{A} [G_n(+\infty) - G_n(A) + G(+\infty) - G(A) + G_n(-A) + G(-A)].$$

By (2.9) and (2.10) we see that

$$\left| \int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)] \right| \leq \frac{5}{4}|t|^3 \delta (\sigma_n^2 + \sigma^2) + \frac{t^2}{2} \sum_{i=0}^m |G_n(x_i) - G(x_{i-1})| \\ + \frac{2|t|}{A} [G_n(+\infty) - G_n(A) + G(+\infty) - G(A) + G_n(-A) + G(-A)] \\ = E(n, t, m(A, \delta)).$$

Q.E.D.

3. The General Result. The bounds we shall obtain on M_n will be derived from bounds on the characteristic functions of $F_n(x)$ and $F(x)$ by using the following two theorems of Esseen [3].

THEOREM 1. Let $M(x)$ be a nondecreasing function, $N(x)$ a real function of bounded variation on the whole real axis such that $N'(x)$ exists and $|N'(x)| \leq B < \infty$, $M(-\infty) = N(-\infty) = 0$ and $N(+\infty) = M(+\infty)$. Let $m(t)$ and $n(t)$ be the corresponding Fourier-Stieltjes transforms and for any $T > 0$ let $\epsilon = \int_{-T}^T |(n(t) - m(t))/t| dt$. Then to every number $k > 1$ there corresponds a finite positive number $c(k)$, only depending on k such that $|N(x) - M(x)| \leq k \cdot \epsilon/2\pi + c(k) \cdot B/T$.

THEOREM 2. Let $M(x)$ be a nondecreasing step function and $N(x)$ a real function of bounded variation on the whole real axis such that

- 1) $M(-\infty) = N(-\infty) = 0$, $M(+\infty) = N(+\infty)$
- 2) If $N(x)$ is discontinuous at $x = x_v (x_v < x_{v+1}, v = 0, \pm 1, \pm 2, \dots)$ there exists a constant $L > 0$ such that $\min(x_{v+1} - x_v) \geq L$,
- 3) $|N'(x)| \leq B < \infty$ everywhere except when $x = x_v (v = 0, \pm 1, \pm 2, \dots)$
- 4) $M(x)$ may be discontinuous only at $x = x_v, (v = 0, \pm 1, \pm 2, \dots)$. Let $m(t)$ and $n(t)$ be the corresponding Fourier Stieltjes transforms and for any $T > 0$ let $\epsilon = \int_{-T}^T |(m(t) - n(t))/t| dt$. Then to every number $k > 1$ there correspond two finite positive constants $c_1(k)$ and $c_2(k)$ only depending on k , such that $|N(x) - M(x)| \leq k(\epsilon/2\pi) + c_1(k) \cdot B/T$, provided that $T \cdot L \geq c_2(k)$.

Now using the notation of (2.1)-(2.4) we define

$$\begin{aligned}
 g(n, m(A, \delta)) &= \left[\frac{1}{3} \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 \right]^{1/5} + \left[\frac{5}{8} \delta (\sigma_n^2 + \sigma^2) \right]^{1/4} \\
 (3.1) \quad &+ \left[\frac{1}{2} \sum_{i=0}^m |G_n(x_i) - G(x_i)| \right]^{1/3} \\
 &+ \left[\frac{4}{A} \{G_n(+\infty) - G_n(A) + G(+\infty) - G(A) + G_n(-A) + G(-A)\} \right. \\
 &\left. + 2|u_n - u| \right]^{1/2}.
 \end{aligned}$$

This leads to the general theorem.

THEOREM 3. Let $F(x)$ be any infinitely divisible distribution function with mean μ and variance σ^2 and with corresponding $G(x)$ given by Kolmogorov's formula (2.1). Let (x_{nk}) be a system of random variables, independent within each row with mean μ_{nk} and variance σ_{nk}^2 . Let $F_n(x)$ be the distribution function of $S_n = x_{n1} + \dots + x_{nk_n}$ and suppose that $dF(x)/dx$ exists and $|dF(x)/dx| \leq B$ for all x . Assume that $\sigma_{nk} \leq 1, k = 1, 2, \dots, k_n$. (The assumption $\sigma_{nk} \leq 1$ is really quite weak as will be seen by Lemma 4.) Then it follows that for any $a > 1$

$$(3.2) \quad M_n = \sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq k(a, B)g(n, m(A, \delta))$$

where $k(a, B)$ is a constant depending only on B and on $a > 1$.

PROOF. For fixed n we first obtain an estimate on $|\log \varphi_n(t) - \log \varphi(t)|$, where $\varphi_n(t)$ and $\varphi(t)$ are the characteristic functions of $F_n(x)$ and $F(x)$ respectively and then use Lemma 1 and Theorem 1 to obtain the bound on M_n . (Since $\varphi(t)$ is the characteristic function of an infinitely divisible distribution, we know that $\log \varphi(t)$ is defined. In the course of the proof it will also become clear that $\log \varphi_n(t)$ is defined.) As in Lemma 2 let $f(t, x) = (e^{itx} - 1 - itx)/x^2$ and define

$$\psi_n(t) = it \sum_{k=1}^{k_n} \mu_{nk} + \int_{-\infty}^{\infty} f(t, x) dG_n(x).$$

Now $|\log \varphi_n(t) - \log \varphi(t)| \leq |\log \varphi_n(t) - \psi_n(t)| + |\psi_n(t) - \log \varphi(t)|$. Let $F'_{nk}(x) = F_{nk}(x + \mu_{nk})$ and let $\varphi'_{nk}(t)$ be the corresponding characteristic function. Let $a_{nk}(t) = \varphi'_{nk}(t) - 1$. Now $\varphi'_{nk}(t) = 1 + \frac{1}{2}\theta\sigma_{nk}^2 t^2$ where $|\theta| \leq 1$ and therefore

$$(3.3) \quad |a_{nk}(t)| = |\theta| \frac{1}{2}\sigma_{nk}^2 t^2$$

Let $T_n = 1/g(n, m(A, \delta))$ and assume in the rest of the proof that $|t| \leq T_n$. Then we see that $|a_{nk}(t)| \leq \frac{4}{5}$ and that

$$\log \varphi'_{nk}(t) = a_{nk}(t) - \frac{a_{nk}^2(t)}{2} + \frac{a_{nk}^3(t)}{3} - \dots$$

so that

$$(3.4) \quad |\log \varphi'_{nk}(t) - a_{nk}(t)| \leq \sum_{r=2}^{\infty} \frac{|a_{nk}(t)|^r}{r} \leq \frac{1}{2} \frac{|a_{nk}(t)|^2}{1 - |a_{nk}(t)|} \leq \frac{5}{2} |a_{nk}(t)|^2.$$

Now we note $\int_{-\infty}^{\infty} f(t, x) dG_n(x) = \sum_{k=1}^{k_n} a_{nk}(t)$, so that

$$(3.5) \quad \psi_n(t) = \sum_{k=1}^{k_n} (it\mu_{nk} + a_{nk}(t)).$$

Also $\varphi'_{nk}(t) = e^{-it\mu_{nk}}\varphi_{nk}(t)$ and thus $\log \varphi_n(t) = \sum_{k=1}^{k_n} (it\mu_{nk} + \log \varphi'_{nk}(t))$. This together with (3.4) and (3.5) shows that

$$\left| \log \varphi_n(t) - \psi_n(t) \right| \leq \sum_{k=1}^{k_n} \left| \log \varphi'_{nk}(t) - a_{nk}(t) \right| \leq \frac{5}{2} \sum_{k=1}^{k_n} |a_{nk}(t)|^2$$

But from (3.3) we see $|a_{nk}(t)| \leq \frac{1}{2}\sigma_{nk}^2 t^2$ so that

$$(3.6) \quad |\log \varphi_n(t) - \psi_n(t)| \leq \frac{5}{2} \cdot \frac{t^4}{4} \sum_{k=1}^{k_n} \sigma_{nk}^4 \leq \frac{5}{8} t^4 \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2$$

where σ_n^2 is the variance of S_n . Now $\log \varphi(t) = i\mu t + \int_{-\infty}^{\infty} f(t, x) dG(x)$. Thus

$$|\psi_n(t) - \log \varphi(t)| \leq |t| \cdot \left| \sum_{k=1}^{k_n} \mu_{nk} - \mu \right| + \left| \int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)] \right|.$$

Applying Lemma 3 we see

$$|\psi_n(t) - \log \varphi(t)| \leq |t| \cdot \left| \sum_{k=1}^{k_n} \mu_{nk} - \mu \right| + E(n, t, m(A, \delta))$$

and using (3.6) we have

$$|\log \varphi_n(t) - \log \varphi(t)| \leq \left\{ \frac{5}{8} t^4 \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 + |t| \cdot |\mu_n - \mu| + E(n, t, m(A, \delta)) \right\} \\ \equiv h(t, n, m(A, \delta)).$$

Now using Lemma 1 we have $|\varphi_n(t) - \varphi(t)| \leq h(t, n, m(A, \delta))$ for $|t| \leq T_n$. In order to apply Esseen's Theorem 1 we consider

$$\int_{-T_n}^{T_n} \left| \frac{\varphi_n(t) - \varphi(t)}{t} \right| dt \leq 2 \int_0^{T_n} \frac{h(t, n, m(A, \delta))}{|t|} dt \leq g(n, m(A, \delta)).$$

Now applying Theorem 1 we see that for any $a > 1$,

$$\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq \frac{a}{2\pi} g(n, m(A, \delta)) + c(a)B \cdot \frac{1}{T_n} = k(a, B)g(n, m(A, \delta))$$

where $k(a, B) = a/2\pi + c(a)B$. Q.E.D.

We shall now examine, under suitable conditions the behavior of $g(n, m(A, \delta))$ as n becomes infinite. To this end we state Theorem 4 (c.f. [4]) which gives the condition for the distribution functions $F_n(x)$ to converge to a limiting distribution and also gives the form of the limiting distribution.

THEOREM 4. *Suppose that the random variables $(x_{nk} - \mu_{nk})$ are infinitesimal. Then a necessary and sufficient condition for the convergence of the distribution functions of sums $S_n = x_{n1} + \dots + x_{nk_n}$ of independent random variables with finite variance to a limiting distribution function with finite variance, and the convergence of the variances of these sums to the variance of the limiting distribution is that there exist a bounded (non-decreasing) function $G(u)$ (with $G(-\infty) = 0$), and a real constant μ such that*

$$1) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{-\infty}^u x^2 dF_{nk}(x + \mu_{nk}) = G(u) \text{ at all continuity points of } G(u),$$

$$2) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{-\infty}^{\infty} x^2 dF_{nk}(x + \mu_{nk}) = \int_{-\infty}^{\infty} dG(u) = G(+\infty) \text{ and}$$

$$3) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \mu_{nk} = \mu.$$

The characteristic function of the limiting distribution is given by Kolmogorov's formula (2.1) using the constant μ and the function $G(u)$ just determined.

Motivated by this theorem we shall assume

$$(3.7) \quad \begin{cases} \text{a) } (x_{nk} - \mu_{nk}) \text{ infinitesimal,} \\ \text{b) } \lim_{n \rightarrow \infty} F_n(x) = F(x) \text{ (at all continuity points of } F(x)) \text{ and} \\ \text{c) } \lim_{n \rightarrow \infty} \sigma_n^2 = \sigma^2 \end{cases}$$

and we see by Theorem 4 that, in the notation of (2.4),

$$(3.8) \quad \begin{cases} 1) & \lim_{n \rightarrow \infty} G_n(x) = G(x) \text{ at all continuity points of } G(x), \\ 2) & \lim_{n \rightarrow \infty} G_n(+\infty) = G(+\infty) \text{ and} \\ 3) & \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \mu_{nk} = \mu \end{cases}$$

where $G(x)$ and μ are the corresponding G and μ of Kolmogorov's formula (2.1) associated with the infinitely divisible distribution function $F(x)$.

We have the following lemma.

LEMMA 4. *If the system (x_{nk}) satisfies (3.7) then $\lim_{n \rightarrow \infty} \max_{1 \leq k \leq k_n} \sigma_{nk}^2 = 0$.*

PROOF. We have

$$\max_{1 \leq k \leq k_n} \sigma_{nk}^2 = \max_{1 \leq k \leq k_n} \int_{-\infty}^{\infty} x^2 dF_{nk}(x + \mu_{nk}).$$

Let $\epsilon > 0$ be given and let $c > 0$ and $-c$ be continuity points of $G(x)$ such that

$$(3.9) \quad |G(c) - G(+\infty)| < \frac{\epsilon}{7} \text{ and } |G(-c)| < \frac{\epsilon}{7}.$$

Now

$$\begin{aligned} \max_{1 \leq k \leq k_n} \int_{-\infty}^{\infty} x^2 dF_{nk}(x + \mu_{nk}) &\leq \max_{1 \leq k \leq k_n} \int_{|x| \leq \sqrt{\epsilon/7}} x^2 dF_{nk}(x + \mu_{nk}) \\ &+ \max_{1 \leq k \leq k_n} \int_{\sqrt{\epsilon/7} < |x| < c} x^2 dF_{nk}(x + \mu_{nk}) + \max_{1 \leq k \leq k_n} \int_{|x| > c} x^2 dF_{nk}(x + \mu_{nk}) \\ &\leq \frac{\epsilon}{7} + c^2 \max_{1 \leq k \leq k_n} P\{|x_{nk} - \mu_{nk}| > \sqrt{\epsilon/7}\} + G_n(-c) + G_n(+\infty) - G_n(c). \end{aligned}$$

By (3.7) and (3.8) we may take N so large that $n > N$ implies

$$\begin{aligned} \max_{1 \leq k \leq k_n} P\{|x_{nk} - \mu_{nk}| > \sqrt{\epsilon/7}\} &\leq \epsilon/7c^2, \\ |G_n(-c) - G(-c)| &< \epsilon/7, \quad |G_n(+\infty) - G(+\infty)| < \epsilon/7 \\ &\text{and } |G(c) - G_n(c)| < \epsilon/7. \end{aligned}$$

Thus we see using (3.9) that

$$\begin{aligned} \max_{1 \leq k \leq k_n} \sigma_{nk}^2 &\leq \frac{2\epsilon}{7} + |G_n(-c) - G(-c)| + |G(-c)| + |G_n(+\infty) - G(+\infty)| \\ &+ |G(+\infty) - G(c)| + |G(c) - G_n(c)| \leq \epsilon \text{ for } n > N. \end{aligned}$$

Q.E.D.

With the notation of Theorem 3 we have the following lemma.

LEMMA 5. *If the random variables (x_{nk}) satisfy (3.7) then for fixed A and δ , $\lim_{n \rightarrow \infty} \sum_{i=0}^{m(A, \delta)} |G_n(x_i) - G(x_i)| = 0$, and $\lim_{n \rightarrow \infty} |\mu_n - \mu| = 0$. This follows from 1 and 3 of (3.8).*

We see from Lemmas 4 and 5 that the first, third and last part of the fourth term of $g(n, m(A, \delta))$, (see (3.1)), approach zero as n becomes infinite under (3.7). Intuitively we think of δ being small and A being large so that $g(n, m(A, \delta))$ will be small. We shall formulate this idea precisely. Let $1 > \delta > 0$ be such that $\pm (1/\delta)^{1/2}$ are continuity points of $G(x)$, ($G(x)$ is arbitrary but fixed), and consider $g(n, m(1/\delta^{1/2}, \delta))$, (i.e. let $A = 1/\delta^{1/2}$) and let $m(1/\delta^{1/2}, \delta) = m(\delta)$. If $1 > \delta_n > 0$ is any sequence of constants then under Theorem 3, we know that

$$(3.10) \quad \sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq k(a, B)g(n, m(\delta_n)).$$

The foregoing discussion leads to the following result which we state as a theorem.

THEOREM 5. *If $F(x)$ and the random variables (x_{nk}) satisfy (3.7) and the hypothesis of Theorem 3, then there exists a sequence $\{1 > \delta_n > 0\}$, $\delta_n \rightarrow 0$ such that (3.10) holds and such that $\lim_{n \rightarrow \infty} g(n, m(\delta_n)) = 0$.*

PROOF. We know that (3.10) holds for any sequence $1 > \delta_n > 0$ of constants. By Lemma 5 we see that if $\pm(1/\delta)^{1/2}$ are continuity points of $G(x)$ then

$$\sum_{i=0}^{m(\delta)} |G_n(x_i) - G(x_i)| \rightarrow 0$$

as $n \rightarrow \infty$. Clearly we can find a sequence $\delta_n \rightarrow 0$, $\pm(1/\delta_n)^{1/2}$ continuity points of $G(x)$ such that $\sum_{i=0}^{m(\delta_n)} |G_n(x_i) - G(x_i)| \rightarrow 0$. But then using Lemmas 4 and 5 and this sequence $\{\delta_n\}$ we see that $g(n, m(\delta_n)) \rightarrow 0$ as n becomes infinite.

A result analogous to that given by Theorem 4 of Berry [1] is contained in the following corollary to Theorems 3 and 5.

COROLLARY. *Under the hypothesis of Theorem 5 if n is so large that*

$$\left| \sum_{k=1}^{k_n} \int_{-\infty}^{x_i} u^2 dF_{nk}(u + \mu_{nk}) - G(x_i) \right| \leq \delta^{3/4},$$

$$\max_{1 \leq k \leq k_n} \sigma_{nk}^2 \leq \delta^{5/4} \quad \text{and} \quad |\mu_n - \mu| \leq \delta^{1/2}$$

then there exists a finite positive number K , independent of n and δ such that $M_n \leq K\delta^{1/4}$.

Now of course Theorems 3 and 5 require $dF(x)/dx$ to exist so that in particular $F(x)$ is continuous. By use of Theorem 2 we can obtain a theorem weakening the condition of continuity but which will require $F_n(x)$ to be a step function. (As a special case where we require $F(x)$ itself to be a step function we get a stronger result due to the fact that $dF(x)/dx = 0$ wherever it exists). Using the notation of Theorem 3 we have the following theorem.

THEOREM 6. *Let $F(x)$ be an infinitely divisible distribution function with two moments such that if $F(x)$ has discontinuities at x_v ,*

$$(x_v < x_{v+1}, v = 0, \pm 1, \pm 2, \dots),$$

then there exists a constant L such that $\min(x_{v+1} - x_v) \geq L$. Suppose that $dF(x)/dx$ exists everywhere except at $x = x_v, v = 0, \pm 1, \pm 2, \dots$ and $|dF(x)/dx| \leq B, x \neq x_v$. Then if $F_n(x)$ is a step function whose only possible discontinuities are $x = x_v, v = 0, \pm 1, \pm 2, \dots$ and if $\sigma_{nk} \leq 1, k = 1, 2, \dots, k_n$, it follows that

for any $a > 1 \sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq k(a, B)g(n, m(A, \delta))$ provided that $L/g(n, m(A, \delta)) \geq c_2(a)$ where $c_2(a)$ is the constant determined in Theorem 2 and $k(a, B)$ has the same meaning as in Theorem 3.

The proof of this theorem is the same as that of Theorem 3 except that we use Theorem 2 instead of Theorem 1.

Now if we define $g_1(n, m(A, \delta)) = \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 + \delta(\sigma_n^2 + \sigma^2) + \sum_{i=0}^m |G_n(x_i) - G(x_i)| + \{G_n(+\infty) - G_n(A) + G(+\infty) - G(A) + G_n(-A) + G(-A)\}/A + |\mu_n - \mu|$ we have the following theorem.

THEOREM 7. *Let $F(x)$ be an infinitely divisible distribution function with two moments and further let it be a step function with discontinuities at $x = x_\nu (x_\nu < x_{\nu+1}, \nu = 0, \pm 1, \pm 2, \dots)$. We assume there exists a constant L such that $\min(x_{\nu+1} - x_\nu) \geq L$. Then if $F_n(x)$ is a step function whose only possible discontinuities are $x = x_\nu, \nu = 0, \pm 1, \pm 2, \dots$, it follows that for any $a > 1$ there exists a constant $k(a)$ depending only on a such that $\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq k(a)g_1(n, m(A, \delta))$ provided that $\max_{1 \leq k \leq k_n} \sigma_{nk} \leq L/c_2(a)$.*

PROOF. Clearly the essential difference in this theorem and Theorem 6 is in the absence of the roots in the expression $g_1(n, m(A, \delta))$. The reason for this is that B of Theorem 2 may be replaced by zero here. With this in mind if we define $T_n = T = c_2(a)/L$ a proof analogous to that of Theorem 3 will hold here as well.

We remark that both Theorems 6 and 7 can be extended just as Theorem 3 was and that the remarks and lemmas following Theorem 3 hold here as well.

4. Specialized theorems for the cases where the limiting distribution is Gaussian or Poisson. In the special cases where the limiting distribution is Gaussian or Poisson the results of the last section may be simplified. For the Gaussian distribution,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du,$$

the G of Kolmogorov's formula (2.1) is given by

$$(4.1) \quad G(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

and for the Poisson distribution, $F(x) = \sum_{0 \leq k \leq x} e^{-\lambda} \lambda^k / k!$ we have

$$(4.2) \quad G(x) = \begin{cases} 0, & x < 1 \\ \lambda, & x \geq 1. \end{cases}$$

The simple nature of the G 's in both of these cases is the reason the results may be simplified.

4a. The Gaussian distribution. In contrast to (3.1) we define (for any $\epsilon > 0$)

$$g_2(n, \epsilon) = \left[\frac{1}{3} \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 \right]^{1/5} + \left[\frac{1}{9} \max(\sigma_n^2, 1) \cdot \epsilon \right]^{1/4} + \left[\sum_{k=1}^{k_n} \int_{|x| \geq \epsilon} x^2 dF_{nk}(x + \mu_{nk}) + \frac{1}{2} |\sigma_n^2 - 1| \right]^{1/3} + |2\mu_n|^{1/2}.$$

We have the following theorem.

THEOREM 8. *Let $\{\epsilon_n > 0\}$ be any sequence of constants. If $\sigma_{nk} \leq 1, k = 1, \dots, k_n$, then for any $a > 1 \sup_{-\infty < x < \infty} |F_n(x) - \Phi(x)| \leq k(a)g_2(n, \epsilon_n)$ where $k(a)$ is a constant depending only on $a > 1$.*

PROOF. The proof here follows the same lines as the proof of Theorem 3. Suffice it to say that in estimating $|\int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)]|$, where $G(x)$ is given by (4.1), we consider

$$\int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)] = \int_{-\infty}^{-\epsilon_n} f(t, x) d[G_n(x) - G(x)] + \int_{\epsilon_n}^{\infty} f(t, x) d[G_n(x) - G(x)]$$

instead of using Lemma 3. Using integration by parts on $\int_{-\infty}^{-\epsilon_n}$ and noticing that on $(-\infty, -\epsilon_n]$ and $[\epsilon_n, +\infty)$, $G_n(x) - G(x)$, is increasing we obtain

$$\left| \int_{-\infty}^{\infty} f(t, x) d[G_n(x) - G(x)] \right| \leq t^2 \sum_{k=1}^{k_n} \int_{|x| \geq \epsilon_n} x^2 dF_{nk}(x + \mu_{nk}) + \frac{t^2}{2} |\sigma_n^2 - 1| + \frac{5}{3}|t|^3 \max[\sigma_n^2, 1] \cdot \epsilon_n.$$

Now if we assume (3.7) where $F(x)$ is the Gaussian distribution $\Phi(x)$, it follows that for any $\epsilon > 0$.

$$(4.3) \quad \begin{cases} 1) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|x| \geq \epsilon} x^2 dF_{nk}(x + \mu_{nk}) = 0, \\ 2) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|x| < \epsilon} x^2 dF_{nk}(x + \mu_{nk}) = 1 \quad \text{and} \\ 3) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \mu_{nk} = 0. \end{cases}$$

Thus, using an argument similar to that used in Theorem 5 we see that if we assume (3.7) that there exists a sequence $\{\epsilon_n > 0\}, \epsilon_n \rightarrow 0$ such that $\lim_{n \rightarrow \infty} g_2(n, \epsilon_n) = 0$.

In order to see more precisely how $g_2(n, \epsilon_n)$ behaves, we shall consider, under appropriate assumptions, finding explicitly a sequence $\{\epsilon_n\}$ which will make $g_2(n, \epsilon_n)$ approach zero as n becomes infinite. We assume that the random variables of the system (x_{nk}) satisfy (3.7) (and therefore (4.3)) so that in particular by Lemma 4 we have $\lim_{n \rightarrow \infty} \max_{1 \leq k \leq k_n} \sigma_{nk}^2 = 0$. Also assume that there exists a $p > 1$ such that

$$(4.4) \quad \max_{1 \leq k \leq k_n} \frac{1}{\sigma_{nk}^2} \left(\int_{-\infty}^{\infty} (x^2)^p dF_{nk}(x + \mu_{nk}) \right)^{1/p}$$

is bounded in n , and let q be determined by $1/p + 1/q = 1$. Then it follows that if we take

$$(4.5) \quad \epsilon_n = \left[\max_{1 \leq k \leq k_n} \sigma_{nk} \right]^{8/(8+3q)}$$

that $g_2(n, \epsilon_n) \rightarrow 0$ as $n \rightarrow \infty$. In fact the only term in $g_2(n, \epsilon_n)$ that needs study is $\sum_{k=1}^{k_n} \int_{|x| \geq \epsilon_n} x^2 dF_{nk}(x + \mu_{nk})$. But by Holders inequality this is

$$\leq \sum_{k=1}^{k_n} \left(\int_{|x| \geq \epsilon_n} (x^2)^p dF_{nk}(x + \mu_{nk}) \right)^{1/p} \left(\int_{|x| \geq \epsilon_n} dF_{nk}(x + \mu_{nk}) \right)^{1/q},$$

and using Tchebycheff's inequality we see this last expression is

$$\leq B \cdot \sigma_n^2 \left(\frac{\max_{1 \leq k \leq k_n} \sigma_{nk}^2}{\epsilon_n} \right)^{1/q}$$

where B is the bound on the expression in (4.4). Using (4.5) we obtain

$$\sum_{k=1}^{k_n} \int_{|x| \geq \epsilon_n} x^2 dF_{nk}(x + \mu_{nk}) \leq B \cdot \sigma_n^2 [\max_{1 \leq k \leq k_n} \sigma_{nk}]^{6/(8+3q)}.$$

Thus using Theorem 8 we have proven the following theorem.

THEOREM 9. *If $\sigma_{nk} \leq 1, k = 1, 2, \dots, k_n$ and if the random variables satisfy (4.4) then*

$$\sup_{-\infty < x < \infty} |F_n(x) - \Phi(x)| \leq k(a) \left\{ \left[\frac{1}{3} \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 \right]^{1/5} + |2\mu_n|^{1/2} + \left| \frac{1}{2} (\sigma_n^2 - 1) \right|^{1/3} \right. \\ \left. + \left[\left(\frac{1}{9} \max_{1 \leq k \leq k_n} (\sigma_n^2, 1) \right)^{1/4} + [B\sigma_n^2]^{1/3} \right] [\max_{1 \leq k \leq k_n} \sigma_{nk}]^{2/(8+3q)} \right\}.$$

We remark that if the (x_{nk}) are not assumed to have any moments and if the limiting distribution is Gaussian or Poisson it is possible, using the method of truncation to obtain an error estimate on $\sup_{-\infty < x < \infty} |F_n(x) - F(x)|$ and to show that the estimate approaches zero under conditions analogous to (3.7). (To show that the estimate approaches zero it is necessary to know (among other things) that $\lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|x| > \tau} dF_{nk}(x) = 0$ for some $\tau > 0$. This is not necessarily true if the limiting distribution is not Gaussian or Poisson.)

4b. *The Poisson distribution.* Define (for any $\epsilon > 0$)

$$g_3(n, \epsilon) = \left\{ \sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 + |\mu_n - \lambda| + \sum_{k=1}^{k_n} \int_{|x-1| \geq \epsilon} x^2 dF_{nk}(x + \mu_{nk}) \right. \\ \left. + |\sigma_n^2 - \lambda| + \max(\sigma_n^2, \lambda) \cdot \epsilon \right\}$$

where λ is the parameter of the Poisson distribution. We have the following theorem analogous to Theorem 8.

THEOREM 10. *Let $F(x)$ be the Poisson distribution and assume that $F_n(x)$ is a step function whose only possible discontinuities are at $x = 0, 1, 2, \dots$. Then it follows that for every $a > 1$ there exists a constant $d(a)$ depending only on a such that $\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq d(a)g_3(n, \epsilon_n)$ provided that $\max_{1 \leq k \leq k_n} \sigma_{nk} \leq 1/c_2(a)$ where $c_2(a)$ is the constant determined in Theorem 2.*

The proof of this theorem is essentially the same as the proof of Theorems 3 and 8 and will be omitted. We remark however that in place of T_n in Theorem 3 we let $T = c_2(a)$ and restrict $|t| \leq T$.

Now if we assume (3.7) with $F(x)$ the Poisson distribution, it follows that

$$(4.6) \quad \left\{ \begin{array}{l} 1) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|z-1| \geq \epsilon} x^2 dF_{nk}(x + \mu_{nk}) = 0 \\ 2) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|z-1| < \epsilon} x^2 dF_{nk}(x + \mu_{nk}) = \lambda \\ 3) \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \mu_{nk} = \lambda \end{array} \right.$$

and the same type of remarks following Theorem 8 hold here as well. In particular, under (3.7), we see that there exists a sequence $\{\epsilon_n > 0\}$ such that $\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq d(a)g_3(n, \epsilon_n)$ and $g_3(n, \epsilon_n)$ approaches zero as $n \rightarrow \infty$. We could also consider finding an explicit sequence $\{\epsilon_n\}$ such that $g_3(n, \epsilon_n)$ approaches zero as n becomes infinite. In fact if $\{\epsilon_n > 0, \epsilon_n \rightarrow 0\}$ is such that for some $\eta, 0 < \eta < 1$, and for some $p > 1$.

$$\max_{1 \leq k \leq k_n} \frac{1}{\sigma_{nk}^2} \left(\int_{\substack{|z-1| \geq \epsilon_n \\ |z| \geq \eta}} x^{2p} dF_{nk}(x + \mu_{nk}) \right)^{1/p}$$

is bounded in n then (under (3.7) and hence (4.6)) $g_3(n, \epsilon_n)$ approaches 0. This follows the same way as the proof of Theorem 9.

As we have said, the simple nature of the $G(x)$'s defined by (4.1) and (4.2) is the reason the error terms for the Gaussian and Poisson distribution are simplified compared to the general case. Evidently the same type of arguments used in this section could be used for other limiting distributions, provided that the corresponding $G(x)$ is of this simple form. Let

$$(4.7) \quad G(x) = \begin{cases} a^2, & x \geq b \\ 0, & x < b \end{cases} \quad a \neq 0.$$

(If $a = 0$ this $G(x)$ corresponds to the unitary distribution.) This leads to the following theorem.

THEOREM 11. *Let X be a random variable with infinitely divisible distribution function $F(x)$. Suppose that $F(x)$ has mean μ and finite variance a^2 .² Let the $G(x)$ of Kolmogorov's formula (2.1) be given by (4.7). Then if $b = 0$, $F(x)$ is the Gaussian distribution, and if $b \neq 0$, the random variable $x' = (x - \mu + a^2/b^2)/b$ is Poisson distributed with parameter a^2/b^2 .*

This theorem follows readily from an examination of the characteristic functions of X and X' .

5. An example. We now consider a specific example, that is a specific system of random variables (x_{nk}) . The system we define here is the system considered in

² If X does not have any moments, then a similar theorem holds using the Lévy-Khintchine formula for the representation of infinitely divisible distributions [7].

the classical Poisson theorem, that is x_{nk} is determined by

$$P\{x_{nk} = 1\} = \frac{\lambda}{n}$$

$$P\{x_{nk} = 0\} = 1 - \frac{\lambda}{n} \quad \text{where } \lambda = 0, \left(\frac{\lambda}{n} \leq 1\right),$$

$k = 1, 2, \dots, n$. We define S_n in the usual way so that $S_n = x_{n1} + \dots + x_{nn}$. It is well known that the distribution functions $F_n(x)$ of S_n approach the Poisson distribution with parameter λ . Using Theorem 10 we consider $\sup_{-\infty < x < \infty} |F_n(x) - F(x)|$. We note that $\mu_{nk} = \lambda/n$ and $\sigma_{nk}^2 = \lambda(1 - \lambda/n)/n$. Assume that $n > 2\lambda$ and define $\epsilon_n = \lambda r/n$ where $2 > r > 1$. Now consider the terms involved in $g_s(n, \epsilon_n)$:

$$\sigma_n^2 \max_{1 \leq k \leq k_n} \sigma_{nk}^2 = \frac{\lambda^2}{n} \left(1 - \frac{\lambda}{n}\right)^2, \quad |\mu_n - \lambda| = 0,$$

$$|\sigma_n^2 - \lambda| = \frac{\lambda^2}{n}, \quad \max(\sigma_n^2, \lambda) \cdot \epsilon_n = \frac{\lambda^2 r}{n}$$

and

$$\sum_{k=1}^{k_n} \int_{|x-1| \geq \epsilon_n} x^2 dF_{nk}(x + \mu_{nk}) = n \cdot \left(\frac{\lambda}{n}\right)^2 \cdot \left(1 - \frac{\lambda}{n}\right) = \frac{\lambda^2}{n} \left(1 - \frac{\lambda}{n}\right).$$

Thus (for $n \geq c_2(a) \cdot \lambda$) we see (since r may be taken arbitrarily close to 1)

$$\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \leq d(a) \left[\frac{4\lambda^2}{n} - \frac{3\lambda^3}{n^2} + \frac{\lambda^4}{n^3} \right] \leq D \cdot \frac{1}{n}$$

where D is a constant.

We remark that although in the above example the system (x_{nk}) is explicitly given, by use of the theorems given here bounds can be obtained on M_n without actually knowing specifically the system x_{nk} involved. Finally we remark that analogous results to those presented here could be obtained by considering sums $S_n^* = x_{n1} + \dots + x_{nk_n} - A_n$ in place of $S_n = x_{n1} + \dots + x_{nk_n}$ where the A_n are constants.

REFERENCES³

- [1] A. C. BERRY, "The accuracy of the Gaussian approximation to the sum of independent variates," *Trans. Amer. Math. Soc.*, Vol. 49 (1941), pp. 449-463.
- [2] L. BIEBERBACH, *Lehrbuch Der Funktionentheorie*, Vol. 1, Chelsea Pub. Co., New York, 1945.
- [3] C. G. ESSEEN, "Fourier analysis of distribution functions," *Acta Math.*, Vol. 77 (1945) pp. 1-125.
- [4] B. V. GNEDENKO AND A. N. KOLMOGOROV, *Limit Distributions for Sums of Independent Random Variables*, translation by K. L. Chung, Addison-Wesley, 1954.
- [5] A. KHINTCHINE, "Zur theorie der unbeschränkt teilbaren verteilungsgestze," *Rec. Math. [Mat. Sbornik]*, N.S. 2 (1937), pp. 79-119.
- [6] A. KOLMOGOROV, "Sulla forma generale di una processo stocastico omogeneo," *Rend. Accad. Lincei*, Vol. 15 (1932), pp. 805-808, 866-869.
- [7] P. LÉVY, *Theorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937.

³ The material referred to in references [3], [5], [6], and [7] may be found in [4].