The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 241–259

RATES OF WEAK CONVERGENCE AND ASYMPTOTIC EXPANSIONS FOR CLASSICAL CENTRAL LIMIT THEOREMS[1]

BY R. N. BHATTACHARYA

University of California, Berkeley

0. Introduction and summary. Let $Q_n(n = 1, 2, \cdots)$, Q be probability measures on the Borel σ-field of R^k. The sequence $\{Q_n\}$ *converges weakly* to Q if for every real-valued, bounded, almost surely (Q) continuous function g on R^k the convergence

$$\lim_n \int g\, dQ_n = \int g\, dQ \tag{0.1}$$

holds (cf. [9], Chapter 1). Such functions g are called *Q-continuous.* If the indicator function I_A of the set A is Q-continuous, then A is also called *Q-continuous.* If $\mathscr{F}$ is a class of Q-continuous functions g over which the convergence (0.1) is uniform for every sequence $\{Q_n\}$ converging weakly to Q, then $\mathscr{F}$ is called a *Q-uniformity class.* A class of sets is called a *Q-uniformity class* if the indicator functions of the sets of this class form a Q-uniformity class of functions. A systematic study of Q-uniformity (in separable metric spaces) was initiated by Ranga Rao [21], who obtained a number of nice results. His studies were carried further in a very useful manner by Billingsley and Topsøe [10].

In this article the error of normal approximation $|\int g\, d(Q_n - \Phi)|$ is estimated for arbitrary Φ-continuous g, Φ being the k-dimensional standard normal distribution and Q_n the distribution of the appropriately normalised nth partial sum of a sequence of independent k-dimensional random vectors $\{X^{(r)}; r = 1, 2, \cdots\}$. The classical central limit theorems assert weak convergence of $\{Q_n\}$ to Φ under certain moment conditions. It is shown here (Theorem 1, Section 3) that for an arbitrary real-valued, bounded, measurable g on R^k one has

$$\left|\int g\, d(Q_n - \Phi)\right| \leqq c(k, \delta)\omega_g(R^k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} + \int \omega_g(S(x, \varepsilon_n))\, d\Phi(x), \tag{0.2}$$

where δ is any positive number; $\rho_{3+\delta,n}$ is defined by (1.4), and

$$\omega_g(A) = \sup\{|g(x) - g(y)|;\, x, y \in A\},\quad S(x, \varepsilon) = \{y; |x - y| < \varepsilon\}, \tag{0.3}$$

$$\varepsilon_n = c(k)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}} \log n,$$

$c(k)$, $c(k, \delta)$ being positive constants depending only on their respective arguments. If ε_n goes to zero as n goes to infinity, then the right side of (0.2) goes to zero for every Φ-continuous g. For the rest of this section let us assume that $\{\rho_{3+\delta,n}\}$ is bounded. By (0.2), if $\int \omega_g(S(x, \varepsilon))\, d\Phi(x) = O(\varepsilon)$ as ε goes to zero, then the error of approximation is $O(n^{-\frac{1}{2}} \log n)$. One may also use (0.2) to obtain uniform upper

Received April 28, 1969; revised February 22, 1970.

[1] Much of this work was carried out in partial fulfillment of the Ph.D. requirement in the University of Chicago. This research was partially supported by U.S. Army Grant DA AROD-31-124 G-816.

bounds for errors of approximation over arbitrary Φ-uniformity classes. In particular,

$$\sup\{|\int g\, d(Q_n-\Phi)|;\ g\in\mathscr{F}_1(\Phi;c,d,\varepsilon_0)\} = O(n^{-\frac{1}{2}}\log n), \tag{0.4}$$

where

$$\mathscr{F}_1(\Phi;c,d,\varepsilon_0) = \{g;\omega_g(R^k)\leqq c,\ \int\omega_g(S(\cdot,\varepsilon))\,d\Phi\leqq\varepsilon \text{ for } 0<\varepsilon\leqq\varepsilon_0\}, \tag{0.5}$$

c, d, ε_0 being arbitrary positive constants. For the largest translation-invariant subclass of this class a sharper bound $O(n^{-\frac{1}{2}})$, which is best possible, was obtained in [3], [5], (cf. [4]). A similar result (in the i.i.d. case) has been independently obtained by Von Bahr [22]. As applications one obtains precise bounds for many interesting classes of sets and functions. However, there are Φ-continuous functions and Φ-uniformity classes of functions for which the technique used in [5] or [22] is not effective. An example in Section 3 shows that there are Borel sets A such that the upper bound for $|Q_n(A)-\Phi(A)|$ as provided by [5] (Theorem 1) is $O(1)$, while (0.2) provides the bound $O(n^{-\frac{1}{2}}\log n)$. A modification of the bound (0.2) when applied to $g = I_A$ for an arbitrary Borel set A enables one to show that the Prokhorov distance between Q_n and Φ is $O(n^{-\frac{1}{2}}\log n)$. It is not known whether the factor $\log n$ in the expression for ε_n in (0.3) (and, hence, in (0.4) and in the estimate of Prokhorov's distance) may be dispensed with or not. However, under Cramér's condition (3.42) $\log n$ may be replaced by one.

The remaining theorems are proved for the i.i.d. case, partly for the sake of simplicity and partly because of the non-availability in the existing literature of complete proofs for some of the expansions related to the characteristic function of Q_n in the non-identically distributed case. Theorem 2 provides an asymptotic expansion for $\int g\, dQ_n$ with a remainder term which is $o(n^{-(s-2)/2})$ uniformly over all g in $\mathscr{F}_1{}^*(\Phi;\, c,\, d,\, \varepsilon_0)$, the largest translation-invariant subclass of $\mathscr{F}_1(\Phi;\, c,\, d,\, \varepsilon_0)$, when $E|X^{(1)}|^s<\infty$ for some integer s not smaller than three and the characteristic function of $X^{(1)}$ obeys Cramér's condition (3.42)$'$. Applications to the class $\mathscr{C}$ of all measurable convex sets, the class $L(c, d)$ (see (3.54)) of bounded Lipschitzian functions, etc., are immediate. Theorem 3 gives an asymptotic expansion for $\int g\, dQ_n$ for a very special class of functions g when no restriction like (3.42)$'$ is imposed. Theorem 4 provides some classes of functions g (under varying restrictions on the distribution of $X^{(1)}$) for which the error of approximation $|\int g\, d(Q_n-\Phi)|$ is of the order $O(n^{-1})$.

Section 1 introduces notation to be used throughout the article. Section 2 provides basic lemmas for proving the results (outlined above) of Section 3.

1. Notation. All probability measures here are defined over the Borel σ-field $\mathscr{B}^k$, unless otherwise specified. Let $\{X^{(r)} = (X_1{}^{(r)}, \cdots, X_k{}^{(r)});\ r = 1, 2, \cdots\}$ be a sequence of independent random vectors in R^k, the rth vector $X^{(r)}$ having distribution $Q^{(r)}$ and characteristic function $f^{(r)}$. It will be assumed that

$$\begin{aligned} E(X_i{}^{(r)}) &= 0, \qquad i = 1,\cdots,k;\quad r = 1,2,\cdots,\\ \operatorname{Cov} X^{(r)} &= D^{(r)}, \qquad r = 1,2,\cdots, \end{aligned} \tag{1.1}$$

$D^{(r)}$ being a positive definite covariance matrix. The same symbol will be used for a linear operator on R^k and its matrix relative to the standard Euclidean basis. Thus Bx denotes the image of x under the map B. Let B_n, B_n' denote a non-singular matrix and its transpose, respectively, such that

$$B_n'B_n = n(\textstyle\sum_{r=1}^{n} D^{(r)})^{-1}. \tag{1.2}$$

Let

$$Y_n = n^{-\frac{1}{2}}B_n \textstyle\sum_{r=1}^{n} X^{(r)}; \qquad Q_n(A) = \text{Probability}(Y_n \in A), \qquad A \in \mathscr{B}^k; \tag{1.3}$$
$$f_n(t) = \textstyle\prod_{r=1}^{n} f^{(r)}(n^{-\frac{1}{2}}B_n't), \qquad t \in R^k.$$

Thus Q_n is the distribution of Y_n and f_n its characteristic function. The covariance matrix of Y_n is, by (1.2), the identity matrix. We write

$$\mu_s^{(r)} = \textstyle\sum_{j=1}^{k} E(X_j^{(r)})^s, \qquad \mu_{s,n} = (\sum_{r=1}^{n} \mu_s^{(r)})/n,$$
$$\beta_s^{(r)} = \textstyle\sum_{j=1}^{k} E|X_j^{(r)}|^s, \qquad \beta_{s,n} = (\sum_{r=1}^{n} \beta_s^{(r)})/n, \tag{1.4}$$
$$\rho_{s,n} = (\textstyle\sum_{r=1}^{n} E|B_n X^{(r)}|^s)/n, \qquad \lambda_{s,n}(u) = (\sum_{r=1}^{n} \lambda_s^{(r)}(u))/n, \qquad u \in R^k,$$

where $|x| = (\sum_{j=1}^{k} x_j^2)^{\frac{1}{2}}$, $(x, y) = \sum_{j=1}^{k} x_j y_j$ for $x = (x_1, \cdots, x_k)$, and $y = (y_1, \cdots, y_k)$ belonging to R^k, and $\lambda_s^{(r)}(u)$ is the cumulant of order s of the random variable $(u, X^{(r)})$. The expressions in (1.4) are, of course, defined for appropriate values of s. Let $P_j(u), j = 0, 1, 2, \cdots$, be polynomials in $u = (u_1, \cdots, u_k)$ defined purely formally by equating coefficients of $n^{-j/2}$ on both sides of

$$\exp[\textstyle\sum_{j=3}^{\infty} n^{-(j-2)2}\lambda_{j,n}(u)/j!] = \sum_{j=0}^{\infty} n^{-j/2}P_j(u). \tag{1.5}$$

Thus what P_j's are meaningfully defined depends on the set of moments which are assumed finite. One has

$$P_0(u) = 1, \qquad P_1(u) = \lambda_{3,n}(u)/6, \qquad P_2(u) = \lambda_{4,n}(u)/24 + \lambda_{3,n}^2(u)/72. \tag{1.6}$$

We shall denote by $P_j(-\varphi)$ the function on R^k whose Fourier transform has the value $P_j(it)\exp(-|t|^2/2)$ at t. Note that since the relation

$$(it_1)^{s_1} \cdots (it_k)^{s_k} \exp(-|t|^2/2) = (-1)^{s_1+\cdots+s_k} \int \exp[i(t,x)] \frac{\partial^{s_1+\cdots+s_k}}{\partial x_1^{s_1} \cdots \partial x_k^{s_k}} \varphi(x)\,dx, \tag{1.7}$$

where φ is the standard normal density, $\varphi(x) = (2\pi)^{-k/2}\exp(-|x|^2/2)$, holds for every k-tuple of nonnegative integers (this may be proved by repeated integration by parts), $P_j(-\varphi)(x)$ may be obtained from the expression $P_j(it)\exp(-|t|^2/2)$ by replacing each term $(it_1)^{s_1} \cdots (it_k)^{s_k} \exp(-|t|^2/2)$ by $(-\partial/\partial x_1)^{s_1} \cdots (-\partial/\partial x_k)^{s_k}\varphi(x)$. The finite signed measure with density $P_j(-\varphi)$ is denoted by $P_j(-\Phi)$. The distribution function corresponding to the measure Φ will also be denoted by Φ.

The topological boundary of any subset A of R^k will be denoted by ∂A. Also the *ε-neighborhood* A^ε of A is defined by

$$A^\varepsilon = \{x; |x-y| < \varepsilon \text{ for some } y \text{ in } A\}, \qquad \varepsilon > 0. \tag{1.8}$$

For ease of reference the definitions (0.3) of the sphere $S(x, \varepsilon)$ and the oscillation $\omega_g(A)$ are repeated here.

$$S(x,\varepsilon) = \{y; y \in R^k, |x-y| < \varepsilon\}, \qquad x \in R^k, \varepsilon > 0; \tag{1.9}$$

$$\omega_g(A) = \sup\{|g(x)-g(y)|; x, y \in A\}, \qquad A \subset R^k,$$

g being a given real-valued function on R^k. Often times we shall deal with the *oscillation function* $\omega_g(S(x, \varepsilon))$ on R^k into the nonnegative reals for a given positive ε.

A probability measure P will be said to have *support* in a set B if $P(B) = 1$.

CONVENTION. Throughout c's will denote constants, either absolute or depending on the indicated arguments.

2. Some lemmas. We shall prove five lemmas in this section. Lemma 1 gives a type of inequality first obtained by Cramér [11] (page 72, Lemma 2).

LEMMA 1. *If* $\rho_{3+\delta,n} < \infty$ *for some* δ, $0 < \delta \leqq 1$, *and* $|t| \leqq n^{\frac{1}{2}}/(2\rho_{3+\delta,n}^{1/(3+\delta)})$, *then*

$$|f_n(t)-(1+n^{-\frac{1}{2}}P_1(it))\exp(-|t|^2/2)| \tag{2.1}$$
$$\leqq (5/2)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-(1+\delta)/2}(|t|^{3+\delta}+|t|^{3(1+\delta)})\exp(-|t|^2/2).$$

PROOF. We first prove the lemma for $k=1$. In this case

$$B_n = \mu_{2,n}^{-\frac{1}{2}}, \qquad \rho_{3+\delta,n} = \mu_{2,n}^{-(3+\delta)/2}\beta_{3+\delta,n}. \tag{2.2}$$

Let

$$g(r,t) = f^{(r)}(n^{-\frac{1}{2}}\mu_{2,n}^{-\frac{1}{2}}t), \qquad U(r,t) = g(r,t)-1. \tag{2.3}$$

By Taylor expansion (cf. [18], page 199),

$$g(r,t) = 1-\mu_2^{(r)}t^2/(2n\mu_{2,n})+\mu_3^{(r)}(it)^3/(6n^{\frac{3}{2}}\mu_{2,n}^{\frac{3}{2}}) \tag{2.4}$$
$$+\theta 2^{1-\delta}\beta_{3+\delta}^{(r)}|t|^{3+\delta}/[(1+\delta)(2+\delta)(3+\delta)n^{(3+\delta)/2}\mu_{2,n}^{(3+\delta)/2}],$$

where θ is used here and elsewhere for a complex number, not always the same, of magnitude not exceeding one. In the given range of t,

$$|U(r,t)| < \tfrac{3}{16}, \qquad |\log(1+U(r,t))-U(r,t)| \leqq |U(r,t)|^2. \tag{2.5}$$

Hence from (2.4) one gets

$$\log g(r,t) = -\mu_2^{(r)}t^2/(2n\mu_{2,n})+(\tfrac{1}{6})\mu_3^{(r)}(it)^3/(n\mu_{2,n})^{\frac{3}{2}}$$
$$+\theta(\tfrac{1}{3})\beta_{3+\delta}^{(r)}|t|^{3+\delta}/(n\mu_{2,n})^{(3+\delta)/2}+\theta[\mu_2^{(r)}t^2/(2n\mu_{2,n}) \tag{2.6}$$
$$+(\tfrac{1}{6})\beta_3^{(r)}|t|^3/(n\mu_{2,n})^{\frac{3}{2}}+(\tfrac{1}{3})\beta_{3+\delta}^{(r)}|t|^{3+\delta}/(n\mu_{2,n})^{(3+\delta)/2}]^2.$$

We now note that for $2 \leqq s \leqq 3+\delta$ one has

$$\beta_s^{(r)} \leqq (\beta_{3+\delta}^{(r)})^{s/(3+\delta)}, \qquad \textstyle\sum_{r=1}^n (\beta_s^{(r)})^2 \leqq (n\beta_{3+\delta,n})^{2s/(3+\delta)}. \tag{2.7}$$

Summing both sides of (2.6) over $r = 1, \cdots, n$, and using (2.7) one obtains after elementary calculations (note that for any three complex numbers a, b, c, $|a+b+c|^2 \leqq 3(|a|^2+|b|^2+|c|^2)$)

$$\log f_n(t) = -t^2/2+\mu_{3,n}(it)^3/(6n^{\frac{1}{2}}\mu_{2,n}^{\frac{3}{2}}) + \theta(\tfrac{29}{24})\beta_{3+\delta,n}|t|^{3+\delta}/(n^{(1+\delta)/2}\mu_{2,n}^{(3+\delta)/2}) = -t^2/2+V, \tag{2.8}$$

say, so that

$$f_n(t)-\exp(-t^2/2) = [\exp(V)-1]\exp(-t^2/2). \tag{2.9}$$

Clearly,

$$\exp(V)-1 = V+(\theta/2)|V|^2\exp(|V|). \tag{2.10}$$

Also simple calculations show that in the given range of t

$$|V| < \tfrac{1}{5}, \qquad |V|^2 \leqq \beta_{3+\delta,n}^{3(1+\delta)/(3+\delta)}|t|^{3(1+\delta)}/(n^{(1+\delta)/2}\mu_{2,n}^{3(1+\delta)/2}). \tag{2.11}$$

Using (2.10) and (2.11) in (2.9) one obtains

$$|f_n(t)-[1+n^{-\frac{1}{2}}\mu_{3,n}\mu_{2,n}^{-\frac{3}{2}}(it)^3/6]\exp(-t^2/2)| \leqq (\tfrac{5}{2})\beta_{3+\delta,n}^{3(1+\delta)/(3+\delta)}\mu_{2,n}^{-3(1+\delta)/2}n^{-(1+\delta)/2}(|t|^{3+\delta}+|t|^{3(1+\delta)})\exp(-|t|^2/2) \tag{2.12}$$

for $|t| \leqq n^{\frac{1}{2}}/(2\rho_{3+\delta,n}^{1/(3+\delta)})$. This proves the lemma for $k = 1$. Note that in this case $P_1(it) = \lambda_{3,n}(it)/6 = \mu_{3,n}\mu_{2,n}^{-\frac{3}{2}}(it)^3/6$. For the general case, define, for a non-zero t in R^k,

$$Z^{(r)} = (t, B_n X^{(r)})/|t|. \tag{2.13}$$

Then $\{Z^{(r)}\}$ is a sequence of independent random variables centered at expectations, and

$$(\textstyle\sum_{r=1}^n E(Z^{(r)})^2)/n = 1, \qquad (i|t|)^3(\sum_{r=1}^n E(Z^{(r)})^3)/(6n) = P_1(it), \qquad (\sum_{r=1}^n E|Z^{(r)}|^s)/n \leqq \rho_{s,n}, \qquad s \geqq 0. \tag{2.14}$$

Applying (2.1) (for $k = 1$) to the characteristic function g_n, say, of $(\sum_{r=1}^n Z^{(r)})/n^{\frac{1}{2}}$ and using (2.14) one gets the inequality

$$|g_n(v)-[1+n^{-\frac{1}{2}}(iv)^3(\textstyle\sum_{r=1}^n E(Z^{(r)})^3)/(6n)]\exp(-v^2/2)| \leqq (\tfrac{5}{2})\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)}n^{-(1+\delta)/2}(|v|^{3+\delta}+|v|^{3(1+\delta)})\exp(-v^2/2) \tag{2.15}$$

for $|v| \leqq n^{\frac{1}{2}}/(2\rho_{3+\delta,n}^{1/(3+\delta)})$. Take $v = |t|$ in (2.15). Since $g_n(|t|) = f_n(t)$, (2.15) reduces to (2.1). □

LEMMA 2. *If* $\rho_{3+\delta,n} < \infty$ *for some* δ, $0 < \delta \leqq 1$, *and* $|t| \leqq n^{\frac{1}{2}}/(4\rho_{3+\delta,n}^{3/(3+\delta)})$, *then*

$$|f_n(t)| \leqq \exp(-|t|^2/3). \tag{2.16}$$

PROOF. For $k = 1$ and $\delta = 0$ this is a result of Cramér [11] (page 75, Lemma 3). Since $\rho_{3,n} \leqq \rho_{3+\delta,n}^{3/(3+\delta)}$ for $\delta \geqq 0$, (2.16) holds in the given range of t, for the case $k = 1$. The multi-dimensional case is proved by applying the one-dimensional inequality to the characteristic function g_n defined above. □

The next two lemmas estimate the effect of smoothing by convolution. We denote by μ^+, μ^-, $|\mu|$, the positive, negative, and total variations, respectively, of a finite signed measure μ ($\mu = \mu^+ - \mu^-$, $|\mu| = \mu^+ + \mu^-$). The symbol '$*$' denotes the operation of convolution. For a bounded, real-valued function g on R^k and a positive number ε, we define (see (1.9))

$$g^{s,\varepsilon}(x) = \sup\{g(y); y \in S(x,\varepsilon)\}, \tag{2.17}$$
$$g^{i,\varepsilon}(x) = \inf\{g(y); y \in S(x,\varepsilon)\}.$$

Note that $g^{s,\varepsilon}$ is lower semi-continuous and $g^{i,\varepsilon}$ is upper semi-continuous because of the (readily verified) equalities

$$\{x; g^{s,\varepsilon}(x) > c\} = \bigcup\{S(x,\varepsilon); g(x) > c\}, \qquad c \in R^1; \tag{2.18}$$
$$g^{i,\varepsilon} = -(-g)^{s,\varepsilon}.$$

In particular, $g^{s,\varepsilon}$, $g^{i,\varepsilon}$, $\omega_g(S(\cdot,\varepsilon)) = g^{s,\varepsilon} - g^{i,\varepsilon}$ are all Borel measurable (whether or not g is measurable).

LEMMA 3. *Let G_ε be a probability measure with support in $S(0,\varepsilon)$, P an arbitrary probability measure, and Q a finite signed measure. For a real-valued, bounded, Borel measurable function g on R^k, define*

$$\gamma(\varepsilon) = \max\{\textstyle\int g^{s,\varepsilon}\, d(P-Q) * G_\varepsilon, -\int g^{i,\varepsilon}\, d(P-Q) * G_\varepsilon\}, \tag{2.19}$$
$$\tau(\varepsilon) = \max\{\textstyle\int (g^{s,2\varepsilon} - g)\, dQ^+, \int (g - g^{i,2\varepsilon})\, dQ^+\}.$$

Then, for every positive ε,

$$|\textstyle\int g\, d(P-Q)| \leqq \gamma(\varepsilon) + \tau(\varepsilon). \tag{2.20}$$

PROOF. By definitions (2.19),

$$\begin{aligned}
\gamma(\varepsilon) &\geqq \textstyle\int g^{s,\varepsilon}\, d(P-Q) * G_\varepsilon \\
&= \textstyle\int_{|x|<\varepsilon} [\int g^{s,\varepsilon}(y+x)\, d(P-Q)(y)]\, dG_\varepsilon(x) \\
&= \textstyle\int_{|x|<\varepsilon} [\int g^{s,\varepsilon}(y+x)\, dP(y) - \int g(y)\, dQ(y) \\
&\qquad\qquad - \textstyle\int (g^{s,\varepsilon}(y+x) - g(y))\, dQ(y)]\, dG_\varepsilon(x) \\
&\geqq \textstyle\int_{|x|<\varepsilon} [\int g(y)\, dP(y) - \int g(y)\, dQ(y) - \int (g^{s,\varepsilon}(y+x) - g(y))\, dQ^+(y)]\, dG_\varepsilon(x) \\
&\geqq \textstyle\int_{|x|<\varepsilon} (\int g\, d(P-Q))\, dG_\varepsilon(x) - \int_{|x|<\varepsilon} [\int (g^{s,2\varepsilon}(y) - g(y))\, dQ^+(y)]\, dG_\varepsilon(x) \\
&= \textstyle\int g\, d(P-Q) - \int (g^{s,2\varepsilon} - g)\, dQ^+ \geqq \int g\, d(P-Q) - \tau(\varepsilon).
\end{aligned} \tag{2.21}$$

Similarly,

$$\begin{aligned}
-\gamma(\varepsilon) &\leqq \textstyle\int g^{i,\varepsilon}\, d(P-Q) * G_\varepsilon = \int_{|x|<\varepsilon} [\int g^{i,\varepsilon}(y+x)\, d(P-Q)(y)]\, dG_\varepsilon(x) \\
&= \textstyle\int_{|x|<\varepsilon} [\int g^{i,\varepsilon}(y+x)\, dP(y) - \int g(y)\, dQ(y) \\
&\qquad\qquad + \textstyle\int (g(y) - g^{i,\varepsilon}(y+x))\, dQ(y)]\, dG_\varepsilon(x) \\
&\leqq \textstyle\int g\, d(P-Q) + \int (g - g^{i,2\varepsilon})\, dQ^+ \leqq \int g\, d(P-Q) + \tau(\varepsilon).
\end{aligned} \tag{2.22}$$

If $\int g\,d(P-Q) \geqq 0$, (2.21) yields (2.20); if $\int g\,d(P-Q) < 0$, then (2.20) follows from (2.22). □

COROLLARY. *Under the hypothesis of Lemma* 3, *the following inequality holds*:

$$(2.23)\qquad \left|\int g\,d(P-Q)\right| \leqq \left|\int g\,d(P-Q)*G_\varepsilon\right| + \int \omega_g(S(\cdot,\varepsilon))\,d\left|(P-Q)*G_\varepsilon\right| + \int \omega_g(S(\cdot,2\varepsilon))\,d\left|Q\right|.$$

If, further $(P-Q)*G_\varepsilon(R^k) = 0$, *then*

$$(2.24)\qquad \left|\int g\,d(P-Q)\right| \leqq \omega_g(R^k)\left|(P-Q)*G_\varepsilon\right|(R^k) + \int \omega_g(S(\cdot,2\varepsilon))\,d\left|Q\right|.$$

PROOF. It is easy to see that

$$(2.25)\qquad \gamma(\varepsilon) \leqq \left|\int g\,d(P-Q)*G_\varepsilon\right| + \int \omega_g(S(\cdot,\varepsilon))\,d\left|(P-Q)*G_\varepsilon\right|,$$

and that

$$(2.26)\qquad \tau(\varepsilon) \leqq \int \omega_g(S(\cdot,2\varepsilon))\,d\left|Q\right|.$$

Using these estimates in Lemma 3 one gets (2.23). If $(P-Q)*G_\varepsilon(R^k) = 0$, then from the definition of $\gamma(\varepsilon)$ it follows that

$$(2.27)\qquad \gamma(\varepsilon) \leqq \omega_g(R^k)\left|(P-Q)*G_\varepsilon\right|(R^k).$$

Inequalities (2.26), (2.27) yield (2.24). □

The next lemma is similar to Lemma 3 in content. Given any probability measure G we denote by G_ε the distribution of the random vector εX, X having distribution G. In this notation G_ε of Lemma 3 may be regarded as arising from a G with support in the unit sphere. Given any real-valued function g on R^k we denote by g_u the translate of g by u, i.e.,

$$(2.28)\qquad g_u(x) = g(x+u).$$

For a given probability measure G, and a constant α' satisfying

$$(2.29)\qquad \tfrac{1}{2} < \alpha' < 1,$$

one can find a constant α such that

$$(2.30)\qquad \int_{|x|<\alpha\varepsilon} dG_\varepsilon(x) = \int_{|x|<\alpha} dG(x) \geqq \alpha'.$$

LEMMA 4. *Let P be a probability measure and Q a finite signed measure. For a real-valued, bounded, Borel measurable function g on R^k, define*

$$(2.31)\qquad \gamma_1(\varepsilon) = \sup\left\{\max\left(\left|\int g_u^{s,\alpha\varepsilon}\,d(P-Q)*G_\varepsilon\right|, \left|\int g_u^{i,\alpha\varepsilon}\,d(P-Q)*G_\varepsilon\right|\right); u\in R^k\right\},$$

$$\tau_1(\varepsilon) = \sup\left[\max\left\{\int (g_u^{s,2\alpha\varepsilon}-g_u)\,d\left|Q\right|, \int (g_u-g_u^{i,2\alpha\varepsilon})\,d\left|Q\right|\right\}; u\in R^k\right],$$

where G is any probability measure, α is chosen to satisfy (2.30). *Then one has, for every positive* ε,

$$(2.32)\qquad \left|\int g\,d(P-Q)\right| \leqq (2\alpha'-1)^{-1}[\gamma_1(\varepsilon)+\tau_1(\varepsilon)].$$

This lemma and the following corollary are proved in [5] (Lemma 8 and relation (2.27)).

COROLLARY. *Under the hypothesis of Lemma* 4 *one has*

$$(2.33)\quad \left|\int g\,d(P-Q)\right| \leqq (2\alpha'-1)^{-1}[\sup\{|\int g_u\,d(P-Q)*G_\varepsilon|+\int \omega_{g_u}(S(\cdot,\alpha\varepsilon))\cdot d\,|(P-Q)*G_\varepsilon|;u\in R^k\}+\sup\{\int \omega_{g_u}(S(\cdot,2\alpha\varepsilon))\,d\,|Q|;u\in R^k\}].$$

Lastly, we shall need the following lemma.

LEMMA 5. *There exists a probability measure* H_1 *with support in* $S(0,1)$ *and having a characteristic function* ζ *satisfying*

$$(2.34)\qquad |\zeta(t)| \leqq \alpha(k)\exp(-|t|^{\frac{1}{2}}),\qquad t\in R^k.$$

PROOF. By a result of Ingham [17], there exists a probability measure H on $\mathscr{B}^1$ such that

$$(2.35)\qquad \int_{-k^{-1/2}}^{k^{-1/2}} dH = 1,\qquad \left|\int \exp(itx)\,dH(x)\right| \leqq \alpha(k)\exp(-|t|^{\frac{1}{2}}),\qquad t\in R^1.$$

Let H_1 be the product measure on $(R^k, \mathscr{B}^k)$, each coordinate measure being H. □

3. Main results. We continue to use the notation of Section 1.

THEOREM 1. *If* $\rho_{3+\delta,n} < \infty$ *for some* $\delta > 0$, *then for any bounded, Borel measurable function* g *on* R^k, *the inequality*

$$(3.1)\quad \left|\int g\,d(Q_n-\Phi)\right| \leqq c(k,\delta)\omega_g(R^k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)}n^{-\frac{1}{2}}+\int \omega_g(S(\cdot,\varepsilon_n))\,d\Phi$$

holds with $\varepsilon_n = c(k)\rho_{3+\delta,n}^{3/(3+\delta)}n^{-\frac{1}{2}}\log n$.[2]

PROOF. Without loss of generality we assume $0<\delta\leqq 1$. Let Z be a random vector with distribution H_1 of Lemma 5. Let H_η denote the distribution of ηZ for positive η. In Lemma 3 take $P = Q_n$, $Q = \Phi$, $\varepsilon = p\eta$, $G_\varepsilon = H_\eta{}^{*p}$, where p is a positive integer and $H_\eta{}^{*p}$ is the p-fold convolution of H_η. Since $(Q_n-\Phi)*G_\varepsilon(R^k)=0$, one has, by (2.24) (corollary to Lemma 3),

$$(3.2)\qquad \left|\int g\,d(Q_n-\Phi)\right| \leqq \omega_g(R^k)\,|(Q_n-\Phi)*G_\varepsilon|\,(R^k)+\int \omega_g(S(\cdot,2\varepsilon))\,d\Phi.$$

Now

$$(3.3)\quad |(Q_n-\Phi)*G_\varepsilon|\,(R^k) \leqq |(Q_n-\Phi-n^{-\frac{1}{2}}P_1(-\Phi))*G_\varepsilon|\,(R^k)+n^{-\frac{1}{2}}\,|P_1(-\Phi)|\,(R^k).$$

One can show (cf. [5], Lemma 7) that

$$(3.4)\qquad |P_1(-\Phi)|\,(R^k) \leqq h(k)\rho_{3,n}.$$

We now estimate $|\mu_n|\,(R^k)$, where

$$(3.5)\qquad \mu_n = (Q_n-\Phi-n^{-\frac{1}{2}}P_1(-\Phi))*G_\varepsilon.$$

[2] One may take $\delta = 0$ if $k = 1$ or 2 (see [5]). This is true for all k if $\{X^{(r)}\}$ is i.i.d. This and some other recent results in the i.i.d. case are contained in the author's article "Recent results on refinements of the central limit theorem" in the forthcoming *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*

For $r > 0$,

(3.6) $$|\mu_n|(R^k) = |\mu_n|(S(0,r)) + |\mu_n|(R^k - S(0,r)).$$

Now

(3.7) $$|\mu_n|(R^k - S(0,r)) \leqq (Q_n * H_\eta^{*p} + \Phi * H_\eta^{*p})(R^k - S(0,r)) + n^{-\frac{1}{2}}|P_1(-\Phi)| * H_\eta^{*p}(R^k).$$

We shall later choose r, p, and η so as to satisfy

(3.8) $$r > 2p\eta,$$

and, consequently,

(3.9) $$H_\eta^{*p}(R^k - S(0, r/2)) = 0.$$

Therefore,

(3.10) $$Q_n * H_\eta^{*p}(R^k - S(0,r)) \leqq Q_n(R^k - S(0, r/2)),$$
$$\Phi * H_\eta^{*p}(R^k - S(0,r)) \leqq \Phi(R^k - S(0, r/2)).$$

Now it is easy to show that

(3.11) $$\Phi(R^k - S(0, r/2)) \leqq (c_1(k)\exp(-r^2/8k))/r.$$

Also one can show (cf. [5], relation (2.48)) by using the Berry–Esseen theorem (cf. [14], page 43, Theorem 1) that

(3.12) $$Q_n(R^k - S(0, r/2)) \leqq (c_1(k)\exp(-r^2/8k))/r + c_2(k)\rho_{3,n}n^{-\frac{1}{2}}.$$

Using these estimates and (3.4) in (3.7) one obtains

(3.13) $$|\mu_n|(R^k - S(0,r)) \leqq (2c_1(k)\exp(-r^2/8k))/r + c_2(k)\rho_{3,n}n^{-\frac{1}{2}} + h(k)\rho_{3,n}n^{-\frac{1}{2}}.$$

It remains to estimate $|\mu_n|(S(0,r))$. Now μ_n has an integrable Fourier transform ξ_n given by

(3.14) $$\xi_n(t) = (f_n(t) - [1 + n^{-\frac{1}{2}}P_1(it)]\exp(-|t|^2/2))\zeta^p(\eta t), \qquad t \in R^k,$$

where ζ is the characteristic function of H_1. By the Fourier inversion theorem μ_n has a density q_n given by

(3.15) $$q_n(x) = (2\pi)^{-k}\int \exp[-i(t,x)] \cdot \xi_n(t)\,dt, \qquad x \in R^k.$$

Hence

(3.16) $$|\mu_n|(S(0,r)) = \int_{|x|<r} |q_n(x)|\,dx \leqq c_3(k)r^k(I_1 + I_2 + I_3),$$

where

$$I_1 = \int_{\{|t| \leqq n^{1/6}/(2\rho_{3+\delta,n}^{1/(3+\delta)})\}} |f_n(t) - [1 + n^{-\frac{1}{2}}P_1(it)]\exp(-|t|^2/2)|\,dt,$$

(3.17) $$I_2 = \int_{\{|t| > n^{1/6}/(2\rho_{3+\delta,n}^{1/(3+\delta)})\}} |f_n(t)\zeta^p(\eta t)|\,dt,$$

$$I_3 = \int_{\{|t| > n^{1/6}/(2\rho_{3+\delta,n}^{1/(3+\delta)})\}} |1 + n^{-\frac{1}{2}}P_1(it)|\exp(-|t|^2/2)\,dt.$$

By Lemma 1,

$$I_1 \leqq c_4(k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-(1+\delta)/2}. \tag{3.18}$$

Since one may assume that

$$n^{\frac{1}{6}}/(2\rho_{3+\delta,n}^{1/(3+\delta)}) \geqq 1 \tag{3.19}$$

(in the contrary case (3.1) is trivially true), one easily obtains

$$I_3 \leqq c_5(k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-(1+\delta)/2}. \tag{3.20}$$

By Lemma 2,

$$I_2 \leqq \int_{\{|t|>n^{1/6}/(2\rho_{3+\delta,n}^{1/(3+\delta)})\}} \exp(-|t|^2/3)\,dt + \int_{\{|t|\geqq n^{1/2}/(4\rho_{3+\delta,n}^{3/(3+\delta)})\}} |f_n(t)\zeta^p(\eta_t)|\,dt. \tag{3.21}$$

The first integral is smaller than

$$c_6(k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-(1+\delta)/2}, \tag{3.22}$$

and the second is, by Lemma 5, smaller than

$$I_4 = \int_{\{|t|\geqq n^{1/2}/(4\rho_{3+\delta,n}^{3/(3+\delta)})\}} \alpha^p(k)\exp(-|\eta t|^{\frac{1}{4}} p)\,dt. \tag{3.23}$$

We now choose

$$\eta = 16(\log\alpha(k)+k)^2\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}}, \qquad p = [\log n]+1, \tag{3.24}$$

where $[x]$ denotes the integer part of x. Elementary calculations now yield

$$I_4 \leqq c_7(k) n^{-k/2}\log^{-2k} n. \tag{3.25}$$

Using estimates (3.22) and (3.25) in (3.21) one obtains

$$I_2 \leqq c_6(k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-(1+\delta)/2} + c_7(k) n^{-k/2}\log^{-2k} n. \tag{3.26}$$

The estimates (3.18), (3.20) and (3.26), when used in (3.16), give

$$|\mu_n|(S(0,r)) \leqq c_8(k,\delta)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} r^k n^{-\frac{1}{2}}\log^{-2k} n. \tag{3.27}$$

Now choose

$$r = (8k\log(n+1))^{\frac{1}{2}}. \tag{3.28}$$

Then (3.13) and (3.27) give

$$|\mu_n|(R^k) \leqq c_9(k,\delta)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}}. \tag{3.29}$$

Finally, making use of (3.29) in (3.2) one obtains (via (3.3) and (3.4)) the desired inequality (3.1). Note that $\varepsilon = p\eta \doteqdot c(k).\ \rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}}\log n$ by the choice of η and p in (3.24). □

REMARK 1. If g in Theorem 1 is such that

$$\int \omega_g(S(\cdot, \varepsilon))\, d\Phi = O(\varepsilon), \qquad \varepsilon \to 0, \tag{3.30}$$

then

$$\left|\int g\, d(Q_n - \Phi)\right| \leqq c(k,\delta)\omega_g(R^k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} + c_{10}(k,g)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}} \log n. \tag{3.31}$$

For the Φ-uniformity class $\mathscr{F}_1(\Phi; c, d, \varepsilon_0)$ defined by (0.5) one has

$$\begin{aligned} &\sup\left\{\left|\int g\, d(Q_n - \Phi)\right|; g \in \mathscr{F}_1(\Phi; c, d, \varepsilon_0)\right\} \\ &\qquad \leqq c(k,\delta) c \rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} + c(k)(d + c/\varepsilon_0)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}} \log n. \end{aligned} \tag{3.32}$$

The term involving ε_0 is introduced to take care of those integers for which $\varepsilon_n > \varepsilon_0$; note that $\left|\int g\, d(Q_n - \Phi)\right| \leqq c$ for all g in the class $\mathscr{F}_1(\Phi; c, d, \varepsilon_0)$. If we denote by $\mathscr{A}_1(\Phi; d, \varepsilon_0)$ the class of all Borel sets A satisfying (cf. [5], Section 1)

$$\Phi((\partial A)^\varepsilon) \leqq d\varepsilon, \qquad 0 < \varepsilon \leqq \varepsilon_0, \tag{3.33}$$

∂A denoting the boundary of A, then the class of all indicator functions of sets in this class is contained in $\mathscr{F}_1(\Phi; 1, d, \varepsilon_0)$.
Hence

$$\begin{aligned} &\sup\left\{\left|Q_n(A) - \Phi(A)\right|; A \in \mathscr{A}_1(\Phi; d, \varepsilon_0)\right\} \\ &\qquad \leqq c(k,\delta)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} + c(k)(d + 1/\varepsilon_0)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}} \log n. \end{aligned} \tag{3.34}$$

For suitable c and d these two classes include most functions and sets of interest. But (3.1) provides an upper bound for every Φ-continuous g. By a variant of a characterization of uniformity classes of functions due to Billingsley and Topsøe [8] (also see [5]), $\mathscr{F}$ is a Φ-uniformity class of functions if and only if

$$\begin{aligned} &\text{(i)} \quad \sup\{\omega_g(R^k); g \in \mathscr{F}\} < \infty, \\ &\text{(ii)} \quad \limsup_{\varepsilon \downarrow 0} \left\{\int \omega_g(S(\cdot, \varepsilon))\, d\Phi; g \in \mathscr{F}\right\} = 0. \end{aligned} \tag{3.35}$$

Hence (3.1) provides effective uniform upper bounds to errors of normal approximation over arbitrary Φ-uniformity classes.

REMARK 2. An error bound different from (3.1) is given in [5] (Theorem 1). According to this

$$\begin{aligned} \left|\int g\, d(Q_n - \Phi)\right| &\leqq c'(k,\delta)\omega_g(R^k)\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} \\ &\quad + c'(k) \sup\left\{\int \omega_{g_u}(S(\cdot, \alpha_n))\, d\Phi; u \in R^k\right\}, \end{aligned} \tag{3.36}$$

where $\alpha_n = c''(k)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}}$. Although (3.36) provides a precise upper bound $O(n^{-\frac{1}{2}})$ (if $\{\rho_{3+\delta,n}\}$ is bounded) for several interesting classes of functions and sets (cf. [5], [20]), we shall show by an example now that there are Borel sets A for which $\Phi((\partial A)^\varepsilon) = O(\varepsilon)$ as ε goes to zero, while $\sup\{\Phi((\partial(A-u))^\varepsilon); u \in R^k\} = 1$ for every positive ε; thus for such a set A, (3.36) is useless, while (3.1) provides an upper bound $O(\rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} \log n)$.

EXAMPLE. In R^1 let

$$A = \bigcup_{r=1}^{\infty} \bigcup_{i=1}^{[(r-1)/2]} \{[r+2i/r, r+(2i+1)/r]\}, \tag{3.37}$$

where $[(r-1)/2]$ is the integer part of $(r-1)/2$. It is easy to see that for every positive ε

$$\sup\{\Phi((\partial(A-u))^{\varepsilon}); u \in R^k\} = 1, \quad \text{but} \tag{3.38}$$

$$\Phi((\partial A)^{\varepsilon}) \leqq \varepsilon(2\pi)^{-\frac{1}{2}} \textstyle\sum_{r=1}^{\infty} r \exp(-r^2/2) = d\varepsilon, \quad \text{say.} \tag{3.39}$$

REMARK 3. It is clear that if g has a compact support, then one may take $\delta = 0$, in Theorem 1. For this case one will have to replace (3.2) by (see (2.23))

$$\begin{aligned} \left|\int g\, d(Q_n - \Phi)\right| \leqq \left|\int g\, d(Q_n - \Phi) * G_\varepsilon\right| + \int \omega_g(S(\cdot,\varepsilon))\, d\left|(Q_n - \Phi) * G_\varepsilon\right| \\ + \int \omega_g(S(\cdot, 2\varepsilon))\, d\Phi, \end{aligned} \tag{3.40}$$

and (3.27) by (use Lemma 1 and Lemma 2 with $\delta = 0$)

$$|\mu_n|\,(S(0,r)) \leqq c_{11}(k,g)\rho_{3,n} n^{-\frac{1}{2}}, \tag{3.41}$$

where r is such that g vanishes outside $S(0,r)$. A similar remark applies to the inequality (3.36). It is not known however, whether the factor $\log n$ in the expression for ε_n in Theorem 1 may be removed or not. If the sequence of characteristic functions $\{f^{(r)}\}$ obeys *Cramér's condition*: *for all positive* η

$$\sup\{|f^{(r)}(t)|; |t| > \eta, r = 1, 2, \cdots\} < 1, \tag{3.42}$$

then even without the factor $\log n$ in ε_n (i.e., with $p = 1$ in (3.24)) one may easily show that I_2, defined by (3.17), is of the smaller order of $1/n$ as n goes to infinity, so that one obtains (3.1) with $\varepsilon_n = c(k)\rho_{3+\delta,n}^{3/(3+\delta)} n^{-\frac{1}{2}}$. Consequently, $\log n$ may be removed from the expressions (3.31), (3.32), and (3.34). For the independent and identically distributed case (3.42) is equivalent to:

$$\limsup_{|t|\to\infty} |f^{(1)}(t)| < 1. \tag{3.42$'$}$$

REMARK 4. There are several well-known metrics which metrize the topology of weak convergence of probability measures on $(R^k, \mathscr{B}^k)$. We mention here the Lévy distance d_L, the Prokhorov distance d_P, and the bounded Lipschitzian distance d_{BL} defined by

$$\begin{aligned} d_L(Q,Q') &= \inf\{\varepsilon; \varepsilon > 0, F_Q(x - \varepsilon e) - \varepsilon \leqq F_{Q'}(x) \\ &\qquad \leqq F_Q(x + \varepsilon e) + \varepsilon \text{ for all } x \text{ in } R^k\}, \\ d_P(Q,Q') &= \inf\{\varepsilon; \varepsilon > 0, Q(A) \leqq Q'(A^{\varepsilon}) + \varepsilon \text{ and} \\ &\qquad Q'(A) \leqq Q(A^{\varepsilon}) + \varepsilon \text{ for all Borel sets } A\}, \\ d_{BL}(Q,Q') &= \sup\{|\textstyle\int g\, d(Q - Q')|; g, |g(x) - g(y)| \leqq |x - y| \text{ for all} \\ &\qquad x, y \text{ in } R^k, \omega_g(R^k) \leqq 1\}, \end{aligned} \tag{3.43}$$

where F_Q, $F_{Q'}$ are the distribution functions corresponding to Q, Q', respectively, and $e = (1, 1, \cdots, 1)$ is the unit vector in R^k. The fact that d_L metrizes the topology of weak convergence of probability measures in R^1 is proved in [16] (page 33, Theorem 1); the proof for R^k is entirely analogous. A proof of the corresponding assertion for d_P (in a separable metric space) may be found in [9] (page 237–238). Dudley [12] (Theorem 12) shows that d_{BL} also metrizes this topology (in a separable metric space). We now estimate these distances between Q_n and Φ. Note that in view of Bergström's extension (cf. [1]) to R^k of the Berry–Esseen theorem, one has

(3.44) $$d_L(Q_n, \Phi) \leqq \sup\{|F_{Q_n}(x) - \Phi(x)|; x \in R^k\} = O(n^{-\frac{1}{2}}), \qquad n \to \infty,$$

if $\{\rho_3^{(r)}\}$ is a bounded sequence. This estimate is precise. The estimate (3.44) may also be obtained under weaker hypotheses (cf. [5], (3.19)). It has been shown in [5] (Section 3, Application 2) that under the hypothesis of Theorem 1,

(3.45) $$d_{BL}(Q_n, \Phi) = O(n^{-\frac{1}{2}}), \qquad n \to \infty.$$

To estimate d_P we note that (cf. [13], Proposition 1)

(3.46) $$d_P(Q, Q') = \inf\{\varepsilon; \varepsilon > 0, Q(A) \leqq Q'(A^\varepsilon) + \varepsilon \text{ for all Borel sets } A\}.$$

Also, note that, by (2.21), for every bounded Borel measurable g,

(3.47) $$\int g\, d(Q-Q') \leqq \int g^{s,\varepsilon} d(Q-Q') * G_\varepsilon + \int (g^{s,2\varepsilon} - g)\, dQ',$$

for probability measures Q, Q'. Taking $g = I_A$, where A is a Borel set, one obtains

(3.48) $$Q(A) - Q'(A) \leqq (Q-Q') * G_\varepsilon(A^\varepsilon) + Q'(A^{2\varepsilon} - A).$$

Now take $Q = Q_n$, $Q' = \Phi$, and G_ε as in Theorem 1. From the estimates obtained in the course of proving Theorem 1, one now has, for every Borel set A,

(3.49) $$Q_n(A) - \Phi(A) \leqq c_{12}(k, \delta) \rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} + \Phi(A^{\varepsilon_n} - A),$$

where ε_n is as in Theorem 1. Let $c_{13}(k, \delta) = \max\{c(k), c_{12}(k, \delta)\}$.
Then

(3.50) $$Q_n(A) - \Phi(A) \leqq \varepsilon_n' + \Phi(A^{\varepsilon_n'} - A),$$ where

(3.51) $$\varepsilon_n' = c_{13}(k, \delta) \rho_{3+\delta,n}^{3(1+\delta)/(3+\delta)} n^{-\frac{1}{2}} \log n.$$

It now follows from (3.46) and (3.50) that

(3.52) $$d_P(Q_n, \Phi) \leqq \varepsilon_n'.$$

The next theorem provides an asymptotic expansion for a large class of functions under Cramér's condition (3.42)'. For a given triplet (c, d, ε_0) of positive numbers, we define

(3.53) $$\begin{aligned}\mathscr{F}_1^*(\Phi; c, d, \varepsilon_0) &= \{g; g_u \in \mathscr{F}_1(\Phi; c, d, \varepsilon_0) \text{ for all } u \in R^k\} \\ &= \{g; \omega_g(R^k) \leqq c, \textstyle\int \omega_{g_u}(S(\cdot, \varepsilon))\, d\Phi \leqq d\varepsilon \text{ for all } \varepsilon \text{ in} \\ &\qquad (0, \varepsilon_0] \text{ and for all } u \text{ in } R^k\}.\end{aligned}$$

Thus $\mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)$ is the largest translation-invariant subclass of $\mathscr{F}_1(\Phi; c, d, \varepsilon_0)$. We consider a few examples.

EXAMPLES. Let $\mathscr{C}$ be the class of all measurable convex sets of R^k. For a suitable constant d depending on k (cf. [20], Appendix A, or [5], Application 1), the class of all indicator functions of members of $\mathscr{C}$ is contained in $\mathscr{F}_1{}^*(\Phi; 1, d, \varepsilon_0)$ for every positive ε_0. The bounded Lipschitz class $L(c, d)$ of all functions g satisfying

$$(3.54)\qquad \omega_g(R^k) \leqq c, \qquad |g(x)-g(y)| \leqq d\,|x-y| \text{ for all } x \text{ and } y \text{ in } R^k,$$

is contained in $\mathscr{F}_1{}^*(\Phi; c, 2d, \varepsilon_0)$. The class of all indicator functions of sets in R^2, each with a boundary contained in a rectifiable curve of length not exceeding a given number l, is contained in $\mathscr{F}_1{}^*(\Phi; 1, 2l+4, 1)$ in R^2 (cf. [10], Section 9, Example 7).

The proof of the theorem below makes use of Lemma 4 and an important estimate of Bikjalis [8]. We shall consider only the identically distributed case. The symbol Φ_ε will denote the distribution of εZ, where Z_1 has distribution Φ.

THEOREM 2. *Let $\{X^{(r)}\}$ be a sequence of independent and identically distributed k-dimensional random vectors each with a zero mean vector, a covariance matrix I (the $k \times k$ identity matrix), and a finite moment $\rho_s = E|X^{(1)}|^s$, s being an integer not smaller than three. If Q_n denotes the distribution of $(\sum_{r=1}^n X^{(r)})/n^{\frac{1}{2}}$, then for every triplet of positive numbers (c, d, ε_0),*

$$(3.55)\qquad \sup\{|\textstyle\int g\, d[Q_n - \sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)]|\,;\, g \in \mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)\} = o(n^{-(s-2)/2}), \qquad n \to \infty,$$

provided $X^{(1)}$ obeys Cramér's condition: $\limsup_{|t|\to\infty} |f^{(1)}(t)| < 1$.

PROOF. It has been shown by Bikjalis [8] (relations (15), (16), (22), (28), (29) combined) that

$$(3.56)\qquad |[Q_n - \textstyle\sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)] * \Phi_\varepsilon|\,(R^k) = o(n^{-(s-2)/2}), \qquad n \to \infty$$

for a suitable ε satisfying

$$(3.57)\qquad \varepsilon = o(n^{-(s-2)/2}), \qquad n \to \infty.$$

By Lemma 4, for every bounded Borel measurable g,

$$(3.58)\qquad |\textstyle\int g\, d[Q_n - \sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)]| \leqq (2\alpha' - 1)^{-1}[\gamma_1(\varepsilon) + \tau_1(\varepsilon)],$$

where, denoting by ν_n the signed measure $\sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)$, one has

$$(3.59)\qquad \begin{aligned} \gamma_1(\varepsilon) &= \sup\{\max(|\textstyle\int g_u{}^{s,\alpha\varepsilon}\, d(Q_n - \nu_n) * \Phi_\varepsilon|)\,;\, u \in R^k\} \\ &\leqq \omega_g(R^k)\,|(Q_n - \nu_n) * \Phi_\varepsilon|\,(R^k) = o(n^{-(s-2)/2}), \qquad n \to \infty, \end{aligned}$$

by (3.56). The inequality in (3.59) holds because

$$(3.60)\qquad (Q_n - \nu_n)(R^k) = 0.$$

Also,

$$\tau_1(\varepsilon) = \sup[\max\{\int(g_u^{s,2\alpha\varepsilon} - g_u)\, d|\nu_n|, \int(g_u - g_u^{i,2\alpha\varepsilon})\, d|\nu_n|\}; u \in R^k], \tag{3.61}$$

so that, for $g \in \mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)$, one has

$$\begin{aligned}\tau_1(\varepsilon) &\leqq \sup\{\int \omega_{g_u}(S(\cdot, 2\alpha\varepsilon))\, d(\Phi + \textstyle\sum_{j=1}^{s-2} n^{-j/2}|P_j(-\Phi)|); u \in R^k\} \\ &\leqq 2\, d\alpha\varepsilon + \sup\{\int \omega_{g_u}(S(\cdot, 2\alpha\varepsilon))\, d(\textstyle\sum_{j=1}^{s-2} n^{-j/2}|P_j(-\Phi)|); u \in R^k\}.\end{aligned} \tag{3.62}$$

Now (see the definition of $P_j(-\Phi)$ in Section 1) for $r = (3s \log n)^{\frac{1}{2}}$,

$$\begin{aligned}&\int_{S(0,r)} \omega_{g_u}(S(\cdot, 2\alpha\varepsilon))\, d(\textstyle\sum_{j=1}^{s-2} n^{-j/2}|P_j(-\Phi)|) \\ &\qquad \leqq c_{14}(k) n^{-\frac{1}{2}} \rho_s (1 + r^{3(s-2)}) \int \omega_{g_u}(S(\cdot, 2\alpha\varepsilon))\, d\Phi \leqq c_{15}(k) \rho_s\, d\varepsilon,\end{aligned} \tag{3.63}$$

and

$$\begin{aligned}&\int_{R^k - S(0,r)} \omega_{g_u}(S(\cdot, 2\alpha\varepsilon))\, d(\textstyle\sum_{j=1}^{s-2} n^{-j/2}|P_j(-\Phi)|) \\ &\qquad \leqq \omega_g(R^k) \textstyle\sum_{j=1}^{s-2} n^{-j/2} |P_j(-\Phi)|\,(R^k - S(0, r)) = o(n^{-s/2}), \qquad n \to \infty.\end{aligned} \tag{3.64}$$

Hence

$$\tau_1(\varepsilon) = o(n^{-(s-2)/2}), \qquad n \to \infty. \tag{3.65}$$

One now obtains (3.55) by using (3.59) and (3.65) in (3.58). □

REMARK 1. As an application of Theorem 2 one has

$$\sup\{|Q_n(C) - \textstyle\sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)(C)|; C \in \mathscr{C}\} = o(n^{-(s-2)/2}), \qquad n \to \infty, \tag{3.66}$$

under the hypothesis of Theorem 2. This is an improvement on a previous result of Ranga Rao [20] (Theorem 5.3.2). A result similar to Theorem 2 for indicator functions has been proved by Von Bahr [22] (Theorem 3(b)) under more restrictive assumptions. When applied to distribution functions (i.e., taking the supremum in (3.66) over the subclass of infinite rectangles), one obtains a previous result due to Bikjalis [8] (Teorema 2) and Von Bahr [22] (Theorem 2). The present author has now been able to prove that one may replace $\mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)$ by the larger class $\mathscr{F}_1(\Phi; c, d, \varepsilon_0)$ in Theorem 2. The proof, however, is based on somewhat different techniques, and will be given elsewhere.

REMARK 2. If $X^{(1)}$ has a distribution with a non-zero absolutely continuous (with respect to Lebesgue measure) component, then Cramér's condition is satisfied. However, in this case the following much stronger inequality holds under the same moment conditions as in Theorem 2:

$$|Q_n - \textstyle\sum_{j=0}^{s-2} n^{-j/2} P_j(-\Phi)|\,(R^k) = o(n^{-(s-2)/2}), \qquad n \to \infty. \tag{3.67}$$

For $k = 1$, this was proved simultaneously by Petrov [19] (Theorem 5) and Bikjalis [6] (Teorema 1). For arbitrary k it has been proved by Bikjalis [8] (Teorema 3).

Theorem 2 obviously holds (and so does Theorem 1) if we allow g to be complex-valued, and redefine $\mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)$ appropriately. In particular, if one takes $g(x) = \exp[i(t, x)]$, $x \in R^k$, then one obtains

$$|f_n(t) - [\textstyle\sum_{j=0}^{s-2} n^{-j/2} P_j(it)] \exp(-|t|^2/2)| = o(n^{-(s-2)/2}), \qquad n \to \infty, \tag{3.68}$$

for all t in R^k. However, much more refined expansions of the characteristic function $f_n(t)$ are available (cf. [16], page 204, Theorem 1 for $k = 1$, and [20], Theorem 5.4.1 for arbitrary k), under the same moment conditions as in Theorem 2, and without the assumption of Cramér's condition (such expansions are, in fact, used to prove (3.56); note also that Lemma 1 of Section 1 is an expansion of this kind). It is, therefore, natural to seek out functions g for which asymptotic expansions hold whatever be the type of distribution of $X^{(1)}$. The theorem below is only a preliminary result in this direction. It does not imply (3.68).

THEOREM 3. *Let $\{X^{(r)}\}$ be a sequence of independent and identically distributed k-dimensional random vectors, each with a zero mean vector, a covariance matrix I (the $k \times k$ identity matrix), and a finite moment $\rho_s = E|X^{(1)}|^s$ for some integer s not smaller than three. Let Q_n be the distribution of $(\sum_{r=1}^n X^{(r)})/n^{\frac{1}{2}}$. Then for a real-valued, integrable function g on R^k whose Fourier transform Ψ satisfies*

$$\int |t|^{s-2}|\Psi(t)|\,dt < \infty, \tag{3.69}$$

the asymptotic expansion

$$\int g\,d[Q_n - \textstyle\sum_{j=0}^{s-2} n^{-j/2}P_j(-\Phi)] = o(n^{-(s-2)/2}), \qquad n \to \infty \tag{3.70}$$

holds.

PROOF. We need the following lemma (cf. [20], Theorem 5.4.1).

LEMMA 6. *For $|t| \leqq n^{\frac{1}{2}}(1/8s)\rho_s^{-3/s}$, the characteristic function f_n of Q_n satisfies*

$$\begin{aligned} &|f_n(t) - (\textstyle\sum_{j=0}^{s-2} n^{-j/2}P_j(it))\exp(-|t|^2/2)| \\ &\quad \leqq c_{16}(k,s)\,\delta(n)n^{-(s-2)/2}\rho_s^{3(s-2)/s}(|t|^s + |t|^{3(s-2)})\exp(-|t|^2/4), \end{aligned} \tag{3.71}$$

where $\delta(n)$ goes to zero as n goes to infinity. Now by Parseval's formula

$$\begin{aligned} &\int g\,d[Q_n - \textstyle\sum_{j=0}^{s-2} n^{-j/2}P_j(-\Phi)] \\ &\quad = (2\pi)^{-k}\int \Psi(-t)[f_n(t) - (\textstyle\sum_{j=0}^{s-2} n^{-j/2}P_j(it))\exp(-|t|^2/2)]\,dt. \end{aligned} \tag{3.72}$$

The integral on the right is estimated first over the region $\{|t| \leqq n^{\frac{1}{2}}(1/8s)\rho_s^{-3/s}\}$ by Lemma 6. This is of the order $o(n^{-(s-2)/2})$. Over the complement of this region it is bounded above in absolute value by

$$\begin{aligned} &(2\pi)^{-k}\int_{\{|t| > n^{1/2}(1/8s)\rho_s^{-3/s}\}} |\Psi(t)|\,dt \\ &\quad + (2\pi)^{-k}\int_{\{|t| > n^{1/2}(1/8s)\rho_s^{-3/s}\}} |\textstyle\sum_{j=0}^{s-2} n^{-j/2}P_j(it)|\exp(-|t|^2/2)\,dt. \end{aligned} \tag{3.73}$$

The first integral is $o(n^{-(s-2)/2})$ because of (3.69), while the second integral is $o(n^{-(s-2)/2})$ because of the presence of the exponential term. □

REMARK. Condition (3.69) implies that g has bounded, continuous derivatives of all orders up to (and including) $s-2$.

EXAMPLES. The functions $g_1(x) = \exp[-\sum_{i,j=1}^k a_{ij}(x_i - m_i)(x_j - m_j)]$ where $A = ((a_{ij}))$ is a positive definite symmetric matrix and $m = (m_1, \cdots, m_k)$ is a given vector in R^k, $g_2(x) = \prod_{i=1}^{k}(1+x_i^2)^{-1}$, and any function whose Fourier transform has a compact support (e.g., $\prod_{i=1}^{k}(\sin^2 x_i)/x_i^2)$ meet the requirement (3.69).

Before stating the final result let us observe that one cannot expect the error $|\int g\, d(Q_n-\Phi)|$ to be of order $o(n^{-\frac{1}{2}})$ if the second term $n^{-\frac{1}{2}}\int g\, dP_1(-\Phi)$ in the asymptotic expansion (whenever appropriate) does not vanish. It does vanish, however, if g is symmetric (about zero) in each co-ordinate for every set of values of the remaining co-ordinates, in which case we shall say that g is *symmetric.* It also vanishes if the third order moments of $X^{(1)}$ (i.e., $E(X_i^{(1)}X_j^{(1)}X_l^{(1)})$ for all i, j, l) vanish (e.g., if $X^{(1)}$ and $-X^{(1)}$ have the same distribution). Esseen [14] (Theorem 1, page 92) has shown that the error is of the order $O(n^{-k/(k+1)})$ uniformly over all indicator functions of spheres centered at the origin provided $E|X^{(1)}|^4$ is finite. Theorem 4 below provides some classes of functions (under varying restrictions on $\{X^{(r)}\}$) for which the error of normal approximation is $O(n^{-1})$.

THEOREM 4. *Let $\{X^{(r)}\}$ be a sequence of independent and identically distributed k-dimensional random vectors each being centered at expectation, and having the covariance matrix I and a finite fourth moment $\rho_4 = E|X^{(1)}|^4$. Let g be a real-valued, bounded, Borel measurable function on R^k. Let also the following hypothesis* (H) *hold: either g is symmetric, or all the third order moments of $X^{(1)}$ are equal to those of Φ. Then each of the conditions* (a), (b), (c) *below implies*

$$|\int g\, d(Q_n-\Phi)| = O(n^{-1}), \qquad n\to\infty. \tag{3.74}$$

(a) *g is integrable and has an integrable Fourier transform Ψ satisfying*

$$\int_{\{|t|>c\}} |t|\cdot|\Psi(t)|\, dt = O(c^{-1}), \qquad c\to\infty. \tag{3.75}$$

(b) *Cramér's condition* (i.e., $\limsup_{|t|\to\infty} |f^{(1)}(t)| < 1$) *holds, and*

$$\sup\{\int \omega_{g_u}(S(\cdot,\varepsilon))\, d\Phi; u\in R^k\} = O(\varepsilon), \qquad \varepsilon\downarrow 0. \tag{3.76}$$

(c) *The distribution of the random vector $X^{(1)}$ has a non-zero absolutely continuous component.*

PROOF. (a) By the hypothesis (H) and Parseval's formula,

$$\begin{aligned}\int g\, d(Q_n-\Phi) &= \int g\, d(Q_n-\Phi-n^{-\frac{1}{2}}P_1(-\Phi))\\ &= (2\pi)^{-k}\int \Psi(-t)[f_n(t)-(1+n^{-\frac{1}{2}}P_1(it))\exp(-|t|^2/2)]\, dt.\end{aligned} \tag{3.77}$$

Over the region $\{|t| \leq n^{\frac{1}{2}}(1/32)\rho_4^{-\frac{3}{4}}\}$ the last integral is of the order $O(n^{-1})$, by Lemma 6. Over the complement of this region it is bounded above in absolute value by

$$\begin{aligned}&(2\pi)^{-k}\int_{\{|t|>n^{1/2}(1/32)\rho_4^{-3/4}\}} |\Psi(t)|\, dt\\ &\quad+(2\pi)^{-k}\int_{\{|t|>n^{1/2}(1/32)\rho_4^{-3/4}\}} |1+n^{-\frac{1}{2}}P_1(it)|\exp(-|t|^2/2)\, dt.\end{aligned} \tag{3.78}$$

The first integral is of the order $O(n^{-1})$ because of (3.75), and the second integral is of the order $o(n^{-1})$ because of the presence of the exponential term.

(b) In this case (3.74) is a consequence of Theorem 2 and the first equality in (3.77).

(c) In this case (3.74) follows from the first equality in (3.77) and Remark 2 following Theorem 2. □

REMARK 1. If one assumes, instead of (H), that all moments of order s and less (s is an integer larger than two) of $X^{(1)}$ coincide with the corresponding moments of Φ, then under the condition (b) one has, by Theorem 2,

$$\sup\{|\int g\, d(Q_n-\Phi)|; g\in\mathscr{F}_1{}^*(\Phi; c, d, \varepsilon_0)\} = o(n^{-(s-2)/2}), \tag{3.79}$$

since the polynomials $P_j(it)$ (see (1.5) and remember that every cumulant of Φ of order three or more is zero), and, hence, the corresponding signed measures $P_j(-\Phi)$, vanish identically for $j = 1, 2, \cdots, s-2$. If, in this case, (c) holds, then (by (3.67),

$$|Q_n-\Phi|(R^k) = o(n^{-(s-2)/2}). \tag{3.80}$$

REMARK 2. All the results in this article may be stated for convergence to a normal distribution Φ_Σ with an arbitrary positive definite covariance matrix Σ. This may be done directly by using expansions of characteristic functions in terms of the characteristic function of such a normal distribution (cf. Bikjalis [7]), or by noting that (cf. [5], Section 4) if $\{Q_n\}$ converges weakly to Φ_Σ, then for every bounded, measurable g, one has

$$\int g\, d(Q_n-\Phi_\Sigma) = \int gT^{-1}\, d(P_n-\Phi), \tag{3.81}$$

where $\{P_n\}$ converges weakly to Φ, and $gT^{-1}(x) = g(T^{-1}(x))$, T being a linear operator satisfying

$$T'T = \Sigma^{-1}. \tag{3.82}$$

Acknowledgment. The author is deeply indebted to Professor Patrick Billingsley for guidance and encouragement. Thanks are also due to Professors V. Petrov, V. Strassen, W. J. Hall, and the referee for helpful comments.

REFERENCES

[1] BERGSTRÖM, H. (1949). On the central limit theorem in the case of not equally distributed random variables. *Skand. Aktuarietidskr.* **32** 37–62.

[2] BERRY, A. C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Am. Math. Soc.* **48** 122–136.

[3] BHATTACHARYA, R. N. (1967). Berry–Esseen bounds for the multi-dimensional central limit theorem. Ph.D. dissertation, Univ. of Chicago.

[4] BHATTACHARYA, R. N. (1968). Berry–Esseen bounds for the multi-dimensional central limit theorem. *Bull. Amer. Math. Soc.* **75** 285–287.

[5] BHATTACHARYA, R. N. (1970). Rates of weak convergence for the multi-dimensional central limit theorem. *Teor. Veroyatnost. i Primenen.* **15** 69–85.

[6] BIKJALIS, A. (1964). On the refinement of the remainder term in the multi-dimensional central limit theorem (Russian). *Litovsk. Mat. Sb.* **4** 153–158.

[7] BIKJALIS, A. (1968). On multivariate characteristic functions (in Russian). *Litvosk. Mat. Sb.* **8** 21–39.

[8] BIKJALIS, A. (1968). Asymptotic expansions of the distribution function and the density function for sums of independent identically distributed random vectors (in Russian). *Litovsk. Mat. Sb.* **8** 405–422.

[9] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[10] BILLINGSLEY, P. and TOPSØE, F. (1967). Uniformity in weak convergence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **7** 1–16.

[11] CRAMÉR, H. (1963). *Random Variables and Probability Distributions* (2nd ed.). Cambridge Univ. Press.

[12] DUDLEY, R. M. (1966). Convergence of Baire measures. *Studia Math.* **27** 251–268.

[13] DUDLEY, R. M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563–1572.

[14] ESSEEN, C. G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125.

[15] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications.* **2** Wiley, New York.

[16] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables* (English translation by K. L. Chung). Addison–Wesley, Cambridge.

[17] INGHAM, A. E. (1934). A note on Fourier transforms. *J. London Math. Soc.* **9** 29–32.

[18] LOÈVE, M. (1963). *Probability Theory,* (3rd ed.). Van Nostrand, Princeton.

[19] PETROV, V. V. (1964). On local limit theorems for sums of independent random variables. (English translation in *Theor. Probability Appl.* **9** 312–320.)

[20] RANGA RAO, R. (1960). Some problems in probability theory. D. Phil. thesis, Calcutta Univ.

[21] RANGA RAO, R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.

[22] VON BAHR, B. (1967). Multi-dimensional integral limit theorems. *Arkiv for Matematik* **7** 71–88.