The Annals of Probability
1977, Vol. 5, No. 1, 1-27

SPECIAL INVITED PAPER

REFINEMENTS OF THE MULTIDIMENSIONAL CENTRAL LIMIT THEOREM AND APPLICATIONS[1]

BY R. N. BHATTACHARYA

University of Arizona

This is an expository survey of recent developments in the field of rates of convergence and asymptotic expansions in the context of the multidimensional central limit theorem. A number of applications are discussed. One of them deals with normal approximations and expansions of distribution functions of a class of statistics which includes functions of sample moments.

0. Introduction and summary. The problem of estimating the error of normal approximation in the central limit theorem is an old one. Among important early work we cite Liapounov [28], Cramér [19], [20], Khinchin [26], Berry [6], Esseen [21], and Bergström [5]. The present article emphasises those developments which have taken place since the appearance of Ranga Rao's work [36], [37]. Since the detailed proofs given in the literature may often appear to be long and somewhat cumbersome, the statements of results in this survey are generally accompanied by sketches of main ideas underlying the proofs.

In order to motivate the discussion we consider a sequence of probability measures $\{Q_n : n \geqq 1\}$ on R^k converging weakly to a probability measure Q. This means

$$\int_{R^k} f\,dQ_n \to \int f\,dQ \qquad n \to \infty \tag{0.1}$$

for every real-valued, bounded Borel measurable function f on R^k whose points of discontinuity form a Q-null set. Equivalently, (0.1) holds if f is bounded and the *oscillation*

$$\omega_f(x : \varepsilon) \equiv \sup\{|f(y) - f(z)| : y, z \in B(x : \varepsilon)\} \tag{0.2}$$

of f on the open ball $B(x : \varepsilon)$ with center x and radius ε goes to zero as $\varepsilon \downarrow 0$ for almost every $x(Q)$. In turn this means that (0.1) holds if

$$\begin{aligned} \omega_f(R^k) &\equiv \sup\{|f(y) - f(z)| : y, z \in R^k\} < \infty\,, \\ \bar{\omega}_f(\varepsilon : Q) &\equiv \int_{R^k} \omega_f(x : \varepsilon) Q(dx) \downarrow 0 \qquad \text{as} \quad \varepsilon \downarrow 0\,. \end{aligned} \tag{0.3}$$

We say $\omega_f(R^k)$ is the *total oscillation of f and $\bar{\omega}_f(\varepsilon : Q)$* is the *average modulus of oscillation of f with respect to Q*.

Received December 22, 1975.

[1] This research was supported by the National Science Foundation Grant MPS 75-07549.

AMS 1970 *subject classification.* Primary 60F05.

Key words and phrases. Normal approximation, asymptotic expansions, Cramér-Edgeworth polynomials, average modulus of oscillation, lattice points, delta method.

A variant of a theorem due to Billingsley and Topsøe [14] says that in order that the convergence (0.1) be uniform over a class of functions $\mathscr{F}$, irrespective of the sequence $\{Q_n: n \geqq 1\}$ converging weakly to Q, it is necessary as well as sufficient that (0.3) holds uniformly over $\mathscr{F}$. Hence if $\{Q_n: n \geqq 1\}$ converges weakly to Q, then it should be possible to bound $|\int f\,dQ_n - \int f\,dQ|$ by an expression which depends on f only via $\omega_f(R^k)$ and the function $\varepsilon \to \bar{\omega}_f(\varepsilon: Q)$. Note that if I_A denotes the indicator function of a Borel set A, then

$$\text{(0.4)} \qquad \begin{aligned} \omega_{I_A}(R^k) &= 1 \qquad \text{if} \quad A \neq R^k, \quad A \neq \phi, \\ \bar{\omega}_{I_A}(\varepsilon: Q) &= Q((\partial A)^\varepsilon), \end{aligned}$$

where ∂A is the *topological boundary of* A and $(\partial A)^\varepsilon$ is the *set of all points at distances less than* ε *from* ∂A. One also defines another related average modulus of oscillation, namely,

$$\text{(0.5)} \qquad \omega_f^*(\varepsilon: Q) \equiv \sup_{y \in R^k} \bar{\omega}_{f_y}(\varepsilon: Q),$$

where f_y is the *translate of* f by y, i.e.,

$$\text{(0.6)} \qquad f_y(x) = f(x+y) \qquad x, y \in R^k.$$

Again note that if A is a Borel set, then

$$\text{(0.7)} \qquad \omega_{I_A}^*(\varepsilon: Q) = \sup_{y \in R^k} Q((\partial(A+y))^\varepsilon),$$

where $A + y = \{x + y: x \in A\}$. Let now $\{X_n: n \geqq 1\}$ be a sequence of i.i.d. k-dimensional random vectors each with mean zero, covariance $I = ((\delta_{ij}))$ ($k \times k$ *identity matrix*), and finite sth absolute moment for some integer $s \geqq 3$. Let Q_n denote the distribution of $n^{-\frac{1}{2}}(X_1 + \cdots + X_n)$, and let Φ be the standard normal distribution on R^k. Theorem 1.7 estimates the error $\int f\,dQ_n - \int f\,d\Phi$ in terms of $\omega_f(R^k)$ and $\omega_f^*(\varepsilon_n: \Phi)$ where $\varepsilon_n = O(n^{-\frac{1}{2}})$. Theorem 1.5 provides an asymptotic expansion of $\int f\,dQ_n$ with an error term $o(n^{-(s-2)/2})$ for all f satisfying

$$\text{(0.8)} \qquad \omega_f(R^k) < \infty, \qquad \bar{\omega}_f(\varepsilon: \Phi) = o((-\log \varepsilon)^{-(s-2)/2}) \qquad \varepsilon \downarrow 0,$$

if Cramér's condition (1.36) holds. The condition (0.8) is very mild. Both these theorems have appropriate extensions to unbounded f. In case X_1 has a density, Theorem 1.2 provides an asymptotic expansion for the density of Q_n. Under the assumption that X_1 has a nonzero absolutely continuous component, Theorem 1.3 implies that the variation norm of the difference between Q_n and its asymptotic expansion is $o(n^{-(s-2)/2})$. If X_1 has a lattice distribution then a precise expansion of the point masses of Q_n (Theorem 2.1) may be used in conjunction with a multidimensional generalization of the Euler–Maclaurin sum formula to yield an asymptotic expansion of probabilities of rectangles properly aligned with the lattice (Theorem 2.2). This restriction on the type of sets for which one has computable expansions in the lattice case is rather severe. The source of difficulty here lies fairly deep. An indication of the nature of this problem is afforded by a discussion of its relationship with the lattice point problem of analytic number theory (Section 3). The usefulness of Theorem 3.1 and its extension

(3.12) to special convex sets A may be viewed in this context. Section 4 is devoted to another application. Here one finds precise error bounds and asymptotic expansions of distribution functions of a class of statistics. This class includes those statistics which are functions of sample moments. With the help of the expansions of Section 4 we are also able to resolve an old conjecture concerning the validity of the formal Edgeworth expansion using the so-called *delta method* for computation of approximate moments and cumulants. To keep the presentation simple, proofs (of the results of Section 4) are merely outlined leaving the details to a future publication. The final section briefly discusses some other applications.

The recent monograph [11] gives a comprehensive account of the theory of rates of convergence and asymptotic expansions in the context of the central limit theorem. Details of proofs of the results in Sections 1 and 2 (excepting Lemma 1.4) may be found there. However, applications are not dealt with in [11]. The present article is intended not only to provide an easy access to some of the main results of the theory but also to introduce the reader to some areas of fruitful applications. Bearing statistical applications (especially, robustness) in mind, an attempt has been made to specify (see, e.g., remarks following (1.50) and (1.64)) the nature of dependence of the error in asymptotic expansions not only on the function, whose integral one approximates, but also on the underlying distribution. In addition, Lemma 1.4 serves to clarify the role of Cramér's condition (1.36) in applications.

For ease of reference we list here some of the main notation used in this article. We deal with sequences of i.i.d. random vectors $\{X_n : n \geqq 1\}$ (or $\{Y_n : n \geqq 1\}$, $\{Z_n : n \geqq 1\}$). The nth normalized partial sum is $n^{-\frac{1}{2}}(X_1 + \cdots + X_n)$ if $EX_1 = 0$; its distribution is Q_n and characteristic function $\hat{Q}_n$. The Fourier transform of a function f is $\hat{f}$, and the Fourier–Stieltjes transform of a finite signed measure G is $\hat{G}$. The standard normal distribution on R^k is Φ and its density is ϕ, while Φ_V, ϕ_V denote the distribution and density of a normal random vector with mean zero and covariance V. Thus $\Phi = \Phi_I$, where I is the identity matrix. The Cramér–Edgeworth polynomials $\tilde{P}_r$, $r \geqq 1$, are defined by (1.16), (1.21). The function $P_r(-\phi_V)$ is defined by (1.24) (on replacing ϕ by the more general ϕ_V). In other words, $P_r(-\phi_V)$ is the function (ϕ_V times a polynomial) whose Fourier transform is $(\tilde{P}_r \cdot \hat{\Phi}_V)(t) = \tilde{P}_r(it) \exp\{-\frac{1}{2}\langle t, Vt\rangle\}$. The signed measure having density $P_r(-\phi_V)$ is $P_r(-\Phi_V)$.

1. The Cramér–Edgeworth expansions and rates of convergence. Consider a sequence of independent and identically distributed (i.i.d.) random vectors $\{X_n = (X_n^{(1)}, \cdots, X_n^{(k)}) : n \geqq 1\}$ with values in R^k and common distribution Q_1. Unless otherwise specified we assume (without essential loss of generality)

$$EX_1 = 0\,, \qquad \operatorname{Cov} X_1 = I\,, \tag{1.1}$$

where EX_1, $\operatorname{Cov} X_1$ are, respectively, the *mean vector* and *covariance matrix of* X_1, and I is the $k \times k$ *identity matrix*. Let $\nu = (\nu^{(1)}, \cdots, \nu^{(k)})$ denote a *multiindex*, i.e., a k-tuple of nonnegative integers, and write

$$|\nu| = \nu^{(1)} + \cdots + \nu^{(k)}, \qquad \nu! = \nu^{(1)}!\,\nu^{(2)}! \cdots \nu^{(k)}!, \tag{1.2}$$
$$x^\nu = (x^{(1)})^{\nu^{(1)}} \cdots (x^{(k)})^{\nu^{(k)}} \qquad x = (x^{(1)}, \cdots, x^{(k)}) \in R^k .$$

The νth *moment of* X_1 (or of Q_1) is

$$\mu_\nu = EX_1^\nu = \int_{R^k} x^\nu Q_1(dx), \tag{1.3}$$

provided the integral is convergent. For a positive integer s the sth *absolute moment of* X_1 (or of Q_1) is

$$\rho_s = E\|X_1\|^s = \int_{R^k} \|x\|^s Q_1(dx), \tag{1.4}$$

where $\|\cdot\|$ is *Euclidean norm*. If G is a finite signed measure on (the Borel sigma field of) R^k, then the *Fourier–Stieltjes transform* (or *characteristic function* (ch.f.) in case G is a probability measure) *of* G is

$$\hat{G}(t) = \int_{R^k} \exp\{i\langle t, x\rangle\} G(dx) \qquad t \in R^k, \tag{1.5}$$

where $\langle\ ,\ \rangle$ denotes *Euclidean inner product*. Since a Taylor expansion and (1.1) yields

$$|\hat{Q}_1(t) - 1| \leqq \frac{\|t\|^2}{2} \qquad t \in R^k, \tag{1.6}$$

the range of $\hat{Q}_1$ on the unit ball $\{\|t\| < 1\}$ is contained in the disc $D(1 : \frac{1}{2}) \equiv \{z \in \mathbb{C} : |z - 1| < \frac{1}{2}\}$ of the complex plane. Since *log*, the principal branch of the logarithm, is analytic in $D(1 : \frac{1}{2})$, $\log \hat{Q}_1$ has continuous derivatives of all orders up to s (if $\rho_s < \infty$) in a neighborhood of the origin. The νth *cumulant of* X_1 (or of Q_1) is

$$\chi_\nu = i^{-|\nu|}(D^\nu \log \hat{Q}_1)(0), \tag{1.7}$$

provided $\rho_{|\nu|} < \infty$. Here D^ν is the νth derivative, i.e.,

$$D^\nu = D_1^{\nu^{(1)}} \cdots D_k^{\nu^{(k)}}, \tag{1.8}$$

where D_j denotes differentiation with respect to the jth coordinate variable. Since $\chi_\nu = \mu_\nu = 0$ if $|\nu| = 1$, a comparison of the Taylor-expansions

$$\hat{Q}_1(t) = 1 + \sum_{2 \leqq |\nu| \leqq s} \frac{\mu_\nu}{\nu!}(it)^\nu + o(\|t\|^s), \tag{1.9}$$
$$\log \hat{Q}_1(t) = \sum_{2 \leqq |\nu| \leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu + o(\|t\|^s) \qquad t \to 0,$$

leads to the formal identity

$$\sum_{2 \leqq |\nu| < \infty} \frac{\chi_\nu}{\nu!}(it)^\nu = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}\left[\sum_{2 \leqq |\nu| < \infty} \frac{\mu_\nu}{\nu!}(it)^\nu\right]^m. \tag{1.10}$$

By equating coefficients of $(it)^\nu$ from the two sides of (1.10) one may express cumulants in terms of moments. In particular, if $s \geqq 3$ then (1.10) implies

$$\chi_\nu = \mu_\nu \qquad \text{if } |\nu| = 2, 3. \tag{1.11}$$

Now the distribution of $X_1 + \cdots + X_n$ is the *n-fold convolution* Q_1^{*n} and its ch.f. is $\hat{Q}_1^n$. Let Q_n denote the *distribution of* $n^{-\frac{1}{2}}(X_1 + \cdots + X_n)$. The νth cumulant of Q_n is $i^{-|\nu|}$ times

(1.12) $$D^\nu(\log \hat{Q}_n)(0) = D^\nu[\log \hat{Q}_1^n(\cdot/n^{\frac{1}{2}})](0) = n^{-(|\nu|-2)/2}\chi_\nu i^{|\nu|},$$

if $\rho_{|\nu|} < \infty$. This relation makes an asymptotic expansion of $\hat{Q}_n$ in powers of $n^{-\frac{1}{2}}$ possible. To see this assume $\rho_s < \infty$ for some $s \geqq 2$ and use the second relation in (1.9) to obtain

(1.13) $$\begin{aligned}\log \hat{Q}_n(t) &= -\frac{||t||^2}{2} + \sum_{3\leqq|\nu|\leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu n^{-(|\nu|-2)/2} + n \cdot o(||t/n^{\frac{1}{2}}||^s)\\ &= -\frac{||t||^2}{2} + \sum_{3\leqq|\nu|\leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu n^{-(|\nu|-2)/2} + o(n^{-(s-2)/2}), \quad n \to \infty.\end{aligned}$$

Hence for all $t \in R^k$

(1.14) $$\begin{aligned}\hat{Q}_n &= \exp\{-||t||^2/2\} \cdot \exp\left\{\sum_{3\leqq|\nu|\leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu n^{-(|\nu|-2)/2}\right\}\\ &\quad \times [1 + o(n^{-(s-2)/2})] \qquad n \to \infty.\end{aligned}$$

If one takes $s = 2$ in (1.14) and uses the Cramér–Lévy continuity theorem ([20], page 106) then one arrives at the classical multidimensional central limit theorem: *If* $\rho_2 < \infty$, *then* $\{Q_n : n \geqq 1\}$ *converges weakly to the standard normal distribution* Φ. If $s \geqq 3$, then expanding the second exponential in (1.14) and collecting together terms involving the same power of $n^{-\frac{1}{2}}$ one has

(1.15) $$\exp\left\{\sum_{3\leqq|\nu|\leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu n^{-(|\nu|-2)/2}\right\} = 1 + \sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(it) + o(n^{-(s-2)/2}).$$

More precisely, replacing $n^{-\frac{1}{2}}$ by the real variable u one obtains a Taylor expansion of the exponential as a function of u. The sum $1 + \sum_{r=1}^{s-2} u^r\tilde{P}_r(it)$ is this Taylor expansion, i.e.,

(1.16) $$\frac{d^r}{du^r}\left[\exp\left\{\sum_{3\leqq|\nu|\leqq s} \frac{\chi_\nu}{\nu!}(it)^\nu u^{|\nu|-2}\right\}\right](0) = r!\, \tilde{P}_r(it).$$

Combining (1.14) and (1.15) one gets

(1.17) $$\hat{Q}_n(t) = \exp\left\{-\frac{||t||^2}{2}\right\} \cdot [1 + \sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(it)] + o(n^{-(s-2)/2}).$$

By carefully estimating the remainders in the two Taylor expansions (1.13) and (1.15) one may obtain the following result.

THEOREM 1.1. *Suppose* $\rho_s < \infty$ *for some integer* $s \geqq 3$. *There exist two positive constants* $c_1(k, s)$, $c_2(k, s)$ *depending only on* k *and* s *such that if* $||t|| \leqq c_1(k, s)n^{\frac{1}{2}} \div \rho_s^{1/(s-2)}$ *then*

(1.18) $$\begin{aligned}&|D^\nu[\hat{Q}_n(t) - \{1 + \sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(it)\} \exp\{-||t||^2/2\}]|\\ &\qquad \leqq \frac{\delta_n}{n^{(s-2)/2}}[||t||^{s-|\nu|} + ||t||^{3(s-2)+|\nu|}] \exp\{-||t||^2/4\} \qquad |\nu| \leqq s,\end{aligned}$$

where $\delta_n \to 0$ *as* $n \to \infty$ *and* $\delta_n \leqq c_2(k, s)\rho_s$ *for all* n.

The first result of this type was obtained for $k = 1$ by Cramér [20] (page 72). Many authors have refined Cramér's result and the present version is proved in [11] (Theorem 9.12).

To obtain a more computable expression for $\tilde{P}_r(it)$ it is convenient to define

$$(1.19)\qquad \chi_r(t) = r! \sum_{|\nu|=r} \frac{\chi_\nu}{\nu!} t^\nu , \qquad \chi_r(it) = i^r\chi_r(t) = r! \sum_{|\nu|=r} \frac{\chi_\nu}{\nu!} (it)^\nu .$$

It is not difficult to check that $\chi_r(t)$ is the rth cumulant of the random variable $\langle t, X_1\rangle$. Now (1.15) reduces to

$$(1.20)\qquad \exp\left\{ \sum_{3\le r\le s} \frac{\chi_r(it)}{r!} n^{-(r-2)/2}\right\} = 1 + \sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(it) + o(n^{-(s-2)/2}) .$$

From this one obtains

$$(1.21)\qquad \tilde{P}_r(it) = \sum_{m=1}^{r} \frac{1}{m!} \left\{\sum{}^* \frac{\chi_{j_1+2}(it)}{(j_1+2)!} \frac{\chi_{j_2+2}(it)}{(j_2+2)!} \cdots \frac{\chi_{j_m+2}(it)}{(j_m+2)!}\right\}$$

where the summation $\sum^*$ is over all m-tuples of positive integers $(j_1, \cdots, j_m)$ satisfying

$$(1.22)\qquad \sum_{l=1}^{m} j_l = r .$$

For example,

$$(1.23)\qquad \begin{aligned} \tilde{P}_1(it) &= \frac{\chi_3(it)}{3!} = \frac{i^3}{3!}\chi_3(t) = \sum_{|\nu|=3} \frac{\chi_\nu}{\nu!}(it)^\nu , \\ \tilde{P}_2(it) &= \frac{\chi_4(it)}{4!} + \frac{\chi_3^2(it)}{2!\,(3!)^2} , \\ \tilde{P}_3(it) &= \frac{\chi_5(it)}{5!} + \frac{\chi_4(it)\chi_3(it)}{4!\,3!} + \frac{\chi_3^3(it)}{(3!)^4} . \end{aligned}$$

For a smooth function f rapidly decreasing at infinity the function $t \to (it)^\nu \hat{f}(t)$ is the Fourier transform of $(-1)^{|\nu|}D^\nu f$. Hence the function $t \to \tilde{P}_r(it)\exp\{-||t||^2/2\}$ is the Fourier transform of the function

$$(1.24)\qquad P_r(-\phi)(x) \equiv \tilde{P}_r(-D)\phi(x) \qquad x \in R^k ,$$

where ϕ is the *standard normal density*, i.e.,

$$(1.25)\qquad \phi(x) = (2\pi)^{-k/2}\exp\{-||x||^2/2\} \qquad x \in R^k ,$$

and $\tilde{P}_r(-D)$ is the differential operator obtained by formally replacing $(it)^\nu$ by $(-D)^\nu = (-1)^{|\nu|}D^\nu$ (for each multiindex ν) in the polynomial expression (1.21) for $\tilde{P}_r(it)$. For example,

$$(1.26)\qquad \begin{aligned} P_1(-\phi)(x) = &-\tfrac{1}{6}\sum_{l=1}^{k} E(X_1^{(l)})^3[3x^{(l)} - (x^{(l)})^3] \\ &- \tfrac{1}{2}\sum_{1\le l\ne m\le k} E[(X_1^{(l)})^2X_1^{(m)}][x^{(m)} - x^{(m)}(x^{(l)})^2] \\ &+ \sum_{1\le l<m<p\le k} E(X_1^{(l)}X_1^{(m)}X_1^{(p)})x^{(l)}x^{(m)}x^{(p)} \\ &\qquad x = (x^{(1)}, \cdots, x^{(k)}) \in R^k . \end{aligned}$$

The finite signed measure having density $P_r(-\phi)$ will be denoted by $P_r(-\Phi)$.

Note that $\phi + \sum_{r=1}^{s-2} n^{-r/2}P_r(-\phi)$ is a polynomial times ϕ, and that (1.18) implies (on taking the derivative at $t = 0$)

$$(1.27)\qquad \int_{R^k} x^\nu Q_n(dx) = (-i)^{|\nu|}(D^\nu \hat{Q}_n)(0) = \int_{R^k} x^\nu[\phi(x) + \textstyle\sum_{r=1}^{s-2} n^{-r/2}P_r(-\phi)(x)]\,dx \qquad 0 \leqq |\nu| \leqq s\,.$$

However, the relations (1.27) do not uniquely determine this polynomial (multiple of ϕ). The reason for this is that the polynomial is of degree $3(s-2) > s$, if $s > 3$.

Let us assume now that $\hat{Q}_1$ is integrable. Then $\hat{Q}_n$ is integrable and Q_n has a density q_n. By Fourier inversion

$$(1.28)\qquad h_\nu(x) \equiv x^\nu[q_n(x) - \phi(x) - \textstyle\sum_{r=1}^{s-2} n^{-r/2}P_r(-\phi)(x)] = (2\pi)^{-k} \int_{R^k} \exp\{-i\langle t, x\rangle\}\hat{h}_\nu(t)\,dt \qquad x \in R^k\,,$$

where

$$(1.29)\qquad \hat{h}_\nu(t) = (-i)^{|\nu|}D^\nu\left[\hat{Q}_n(t) - \{1 + \textstyle\sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(it)\}\exp\left\{-\frac{||t||^2}{2}\right\}\right].$$

By Theorem 1.1

$$(1.30)\qquad \int_{\{||t|| \leqq c_1(k,s)n^{\frac{1}{2}}/\rho_s^{1/(s-2)}\}} |\hat{h}_\nu(t)|\,dt = o(n^{-(s-2)/2})\,.$$

Also, since $|\hat{Q}_1(t)| < 1$ for $t \neq 0$ and $|\hat{Q}_1(t)| \to 0$ as $||t|| \to \infty$ (by the Riemann–Lebesgue lemma)

$$(1.31)\qquad \delta \equiv \sup\{|\hat{Q}_1(t)|;\ ||t|| > c_1(k, s)/\rho_s^{1(s-2)}\} < 1\,.$$

By repeated use of the Leibnitz rule for differentiation of a product of functions it may be shown that

$$(1.32)\qquad |D^\nu \hat{Q}_n(t)| = |D^\nu \hat{Q}_1^n(t/n^{\frac{1}{2}})| \leqq n^{|\nu|/2}\rho_{|\nu|}|\hat{Q}_1(t/n^{\frac{1}{2}})|^{n-|\nu|}\,.$$

Therefore,

$$(1.33)\qquad \begin{aligned} &\int_{\{||t|| > c_1(k,s)n^{\frac{1}{2}}/\rho_s^{1/(s-2)}\}} |D^\nu \hat{Q}_n(t)|\,dt \\ &\quad \leqq n^{|\nu|/2}\rho_{|\nu|}\delta^{n-|\nu|-1} \int_{R^k} |\hat{Q}_1(t/n^{\frac{1}{2}})|\,dt \\ &\quad = n^{(|\nu|+k)/2}\delta^{n-|\nu|-1} \int_{R^k} |\hat{Q}_1(t)|\,dt = o(n^{-(s-2)/2})\,. \end{aligned}$$

Since the remaining terms in $\hat{h}_\nu$ possess an exponential factor it follows from (1.28), (1.30), and (1.33) that

$$(1.34)\qquad \sup_{x \in R^k} |x^\nu[q_n(x) - \phi(x) - \textstyle\sum_{r=1}^{s-2} n^{-r/2}P_r(-\phi)(x)]| = o(n^{-(s-2)/2}) \qquad 0 \leqq |\nu| \leqq s\,.$$

Taking $\nu = 0$ in (1.34) one arrives at a *uniform local expansion* of q_n. It may be noted that the proof undergoes only minor modification if one assumes that $|\hat{Q}_1|^m$ is integrable for some $m \geqq 1$. It is also fairly simple to show that the last condition is equivalent to saying that Q_1^{*m} has a bounded density for some positive integer m. Thus one has

THEOREM 1.2. *Assume $\rho_s < \infty$ for some integer $s \geqq 2$. In order that for sufficiently large n the distribution Q_n may have a density q_n satisfying* (1.34) *it is necessary as well as sufficient that Q_1^{*m} has a bounded density for some positive integer m.*

One dimensional versions of this theorem may be found in Gnedenko and Kolmogorov [22] (page 228) and in Petrov [32]. The present version is proved in [11] (Theorems 19.1, 19.2). If $s > k + 1$, then on integration over R^k the relation (1.34) yields an estimate $o(n^{-(s-2)/2})$ for the *variation norm* $\|Q_n - \Phi - \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)\|$. However, for this there is a better result. We denote by $|G|$ the *total variation* (measure) *of a finite signed measure* G.

THEOREM 1.3. *Suppose* $\rho_s < \infty$ *for some integer* $s \geqq 2$. *In order that the relation*

$$(1.35) \qquad \int_{R^k} (1 + \|x\|^s) |Q_n - \Phi - \textstyle\sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)|(dx) = o(n^{-(s-2)/2})$$

may hold it is necessary as well as sufficient that Q_1^{*m} *has a nonzero absolutely continuous component for some positive integer* m.

If the integrand $(1 + \|x\|^s)$ in (1.35) is replaced by 1, then one arrives at the variation norm estimate mentioned above. This estimate is due to Bikjalis [13]. The present stronger result is useful (e.g., in estimating moments of a function of $\bar{X} = n^{-1}(X_1 + \cdots + X_n)$) and a detailed proof is given in [11] (Theorem 19.5). It should be noted that the hypothesis of Theorem 1.3 is less restrictive than that of Theorem 1.2. Thus if one is interested only in estimating $\int f \, dQ_n$ for Borel measurable functions f (or probabilities $Q_n(B)$ for Borel sets B), then Theorem 1.3 is a more useful result than Theorem 1.2.

A hypothesis less restrictive than those used in the preceding theorems was introduced by Cramér ([20], page 82). This is the so-called *Cramér's condition*:

$$(1.36) \qquad \limsup_{\|t\| \to \infty} |\hat{Q}_1(t)| < 1 \, .$$

In view of the Riemann–Lebesgue lemma, it Q_1 has a nonzero absolutely continuous component then $\hat{Q}_1$ satisfies (1.36). There are, however, many singular measures satisfying Cramér's condition. The following lemma provides a class of examples which are used in Section 4.

LEMMA 1.4. *Let* X *be a random vector with values in* R^m *whose distribution has a nonzero absolutely continuous component* H (*relative to Lebesgue measure on* R^m). *Let* f_i, $1 \leqq i \leqq k$, *be Borel measurable real-valued functions on* R^m. *Assume that there exists an open ball* B *of* R^m *in which the density of* H *is positive almost everywhere and in which* f_i*'s are continuously differentiable. If in* B *the functions* 1, $f_1, \cdots, f_k$ *are linearly independent, then the distribution* Q_1 *of* $(f_1(X), \cdots, f_k(X))$ *satisfies Cramér's condition* (1.36).

PROOF. Let $\theta_0 = (\theta_0^{(1)}, \cdots, \theta_0^{(k)}) \in R^k$, $\theta_0 \neq 0$. The assumption of linear independence implies that there is a j $(1 \leqq j \leqq m)$ and an $x_0 \in B$ such that $(\sum_{l=1}^k \theta_0^{(l)} D_j f_l)(x_0) \neq 0$. Without loss of generality we may take $j = 1$ and assume that $(\sum_{l=1}^k \theta^{(l)} D_1 f_l)(x) > \delta > 0$ for all $x \in B$ and all θ in the open ball $B(\theta_0 : \varepsilon)$ with center θ_0 and radius $\varepsilon > 0$. Here δ is an appropriate positive number. Consider the function

$$g(\theta, x) = (\theta, x') \, ,$$

$$x' = (\textstyle\sum_{l=1}^k \theta^{(l)} f_l(x), x^{(2)}, \cdots, x^{(m)}) \qquad x = (x^{(1)}, \cdots, x^{(m)}) \in R^m \, ,$$

on $B(\theta_0: \varepsilon) \times B$. Since the Jacobian of this map is $\sum_{l=1}^k \theta^{(l)} D_1 f_l$, which is positive on $B(\theta_0: \varepsilon) \times B$, one may use the inverse function theorem to assert (by reducing ε and B is necessary) that g defines a diffeomorphism between $B(\theta_0: \varepsilon) \times B$ and its image under g. It follows that for each $\theta \in B(\theta_0: \varepsilon)$ the map $g_\theta(x) = (\sum_{l=1}^k \theta^{(l)} f_l(x), x^{(2)}, \cdots, x^{(m)})$ is a diffeomorphism between B and $g_\theta(B)$. Let H_0 denote the restriction of H to B. Then the measure $H_0 \circ g_\theta^{-1}$ induced on $g_\theta(B)$ by the map g_θ has a density given by

$$h_\theta(z) = \frac{h(g_\theta^{-1}(z))}{\sum_{l=1}^k \theta^{(l)}(D_1 f_l)(g_\theta^{-1}(z))} \qquad z \in g_\theta(B),$$

where h is the density of H_0. Extend h to all of R^m by setting it equal to zero outside $g_\theta(B)$. Then for all $z \in R^m$, $z \notin \partial g_{\theta_0}(B)$, $h_\theta(z) \to h_{\theta_0}(z)$ as $\theta \to \theta_0$. Write

$$h_{\theta,1}(z^{(1)}) = \int_{R^{m-1}} h_\theta(z)\, dz^{(2)} \cdots dz^{(m)} \qquad z^{(1)} \in R^1.$$

Since the m-dimensional Lebesgue measure of $\partial g_{\theta_0}(B)$ is zero,

$$(1.37) \qquad \int_{R^1} |h_{\theta,1}(u) - h_{\theta_0,1}(u)|\, du \leqq \int_{R^m} |h_\theta(z) - h_{\theta_0}(z)|\, dz \to 0 \qquad \theta \to \theta_0.$$

Now suppose (1.36) does not hold. Then there exists a sequence $\{t_n: n \geqq 1\}$ such that $||t_n|| \to \infty$ and

$$(1.38) \qquad |\hat{Q}_1(t_n)| \to 1 \quad \text{as} \quad n \to \infty.$$

Let $\theta_n = t_n/||t_n||$. Restricting to a subsequence if necessary, we assume that $\{\theta_n\}$ converges to some θ_0. Let G_n be the distribution of the random variable $\sum_{l=1}^k \theta_n^{(l)} f_l(X)$ for $n = 0, 1, 2, \cdots$. Write

$$(1.39) \qquad G_n = G_{n,1} + G_{n,2},$$

where $G_{n,2}$ has density $h_{\theta_n,1}$. Then

$$(1.40) \qquad |\hat{Q}_1(t_n)| = |\hat{G}_n(||t_n||)| \leqq |\hat{G}_{n,1}(||t_n||)| + |\hat{G}_{n,2}(||t_n||)|.$$

But $||G_{n,2} - G_{0,2}|| \to 0$ as $n \to \infty$ by (1.37). Hence $\hat{G}_{n,2}(u)$ converges to $\hat{G}_{0,2}(u)$ uniformly in u. Also by the Riemann–Lebesgue lemma $\hat{G}_{0,2}(||t_n||) \to 0$ as $n \to \infty$. Therefore, $|\hat{G}_{n,2}(||t_n||)| \to 0$. Using this and (1.40) one has

$$\limsup_{n\to\infty} |\hat{Q}_1(t_n)| \leqq \limsup_{n\to\infty} |\hat{G}_{n,1}(||t_n||)| \leqq \limsup_{n\to\infty} ||G_{n,1}|| = 1 - \int_{R^1} h_{\theta_0,1}(u)\, du < 1.$$

This contradicts (1.38). □

To appreciate the significance of this result take $m = 1$, $k > 1$. Then $x \to (f_1(x), \cdots, f_k(x))$ is a curve in R^k, and the distribution Q_1 of the random vector $(f_1(X), \cdots, f_k(X))$ is clearly singular (with respect to Lebesgue measure on R^k).

It has been shown by Yurunskii [44] that if the f_i's in Lemma 1.4 are analytic then there exists an integer m such that Q_1^{*m} has a nonzero absolutely continuous component.

It is clear that if Q_1^{*n} is, for all n, singular (with respect to Lebesgue measure on R^k), then there exists a Borel set A such that $Q_n(A) = 1$ for all n and $\Phi(A) = 0$.

Thus convergence in variation norm is ruled out and we fall back on weak convergence. Recall the definitions of the total oscillation $\omega_f(R^k)$ and the average modulus of oscillation $\bar{\omega}_f(\varepsilon : \Phi)$ (see (0.2), (0.3)). The following expansion holds.

THEOREM 1.5. *If $\rho_s < \infty$ for some integer $s \geqq 3$ and $\hat{Q}_1$ satisfies Cramér's condition* (1.36), *then for every real-valued, bounded, Borel measurable function f on R^k one has*

$$(1.41)\qquad |\textstyle\int_{R^k} f d[Q_n - \Phi - \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)]| \leqq \dfrac{\delta_n}{n^{(s-2)/2}} \omega_f(R^k) + \bar{\omega}_f(e^{-dn} : \Phi) ,$$

where $\delta_n \to 0$ as $n \to \infty$, d is a positive constant, and the quantities δ_n and d do not depend on f.

A detailed proof of this theorem may be found in [8] (Theorem 4.3). However, the main ideas underlying the proof may be stated rather simply. In the present case $\hat{Q}_n$ is not necessarily integrable. Therefore, one chooses a kernel probability measure K whose support is contained in the closed unit ball of R^k, and whose characteristic function satisfies

$$(1.42)\qquad |D^\nu \hat{K}(t)| = O(\exp\{-||t||^{\frac{1}{2}}\}) \qquad ||t|| \to \infty ,$$

for all multiindices ν. The existence of such a kernel follows from a result of Ingham (see [8], Corollary 3.1). For $\varepsilon > 0$ define the probability measure K_ε by

$$(1.43)\qquad K_\varepsilon(A) = K(\varepsilon^{-1}A) \quad (A \text{ Borel set; } \varepsilon^{-1}A = \{\varepsilon^{-1}x : x \in A\}) .$$

The effect of smoothing by convolution with K_ε is provided by the following lemma (see [8], Corollary 2.1).

LEMMA 1.6. *Let G be a finite measure and H a finite signed measure such that $G(R^k) = H(R^k)$, and let K be a probability measure on R^k. If*

$$(1.44)\qquad K(B(0:1)) = 1 \qquad (B(0:1) = \{||x|| < 1\}) ,$$

then for every $\varepsilon > 0$ and every real-valued, bounded, Borel measurable function f on R^k one has

$$(1.45)\qquad |\textstyle\int_{R^k} f d(G - H)| \leqq \omega_f(R^k)||(G - H) * K_\varepsilon|| + \bar{\omega}_f(2\varepsilon : |H|)$$

where $|H|$ is the total variation of H.

In this lemma let $G = Q_n$, $H = \Phi + \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)$ and K as specified earlier. The optimum ε is of the order e^{-dn} where d is a positive constant satisfying

$$(1.46)\qquad 0 < d < -\frac{1}{k} \log \theta , \qquad \theta = \sup\{|\hat{Q}_1(t)| : ||t|| > (16\rho_3)^{-1}\} .$$

Cramér's condition (1.36) ensures that $\theta < 1$ and that, consequently, such a choice of d is possible. Since $(\hat{G} - \hat{H}) \cdot \hat{K}_\varepsilon$ is integrable, one may use Fourier inversion and Theorem 1.1 to estimate the variation norm $||(G - H) * K_\varepsilon||$. Integrability of $\hat{K}_\varepsilon$ and the fact that $\sup\{|\hat{Q}_n(t)| : ||t|| > n^{\frac{1}{2}}/(16\rho_3)\} = \theta^n$ makes an adequate estimation of the tail integral of $(\hat{G} - \hat{H}) \cdot \hat{K}_\varepsilon$ possible.

To apply Theorem 1.5 note that the right side in (1.41) is $o(n^{-(s-2)/2})$ if (0.8) holds. For example, consider the class of Borel sets

(1.47) $\mathscr{A}_\alpha(a:\Phi) = \{A: A \text{ Borel set, } \Phi((\partial A)^\varepsilon) \leqq a\varepsilon^\alpha \text{ for } \varepsilon > 0\}$,

where α and a are specified positive numbers. Then one has

(1.48) $\sup_{A \in \mathscr{A}_\alpha(a:\Phi)} |Q_n(A) - \Phi(A) - \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)(A)| = o(n^{-(s-2)/2})$.

It was first shown by Ranga Rao [36] and later by von Bahr [2] that for the *class $\mathscr{C}$ of all Borel measurable convex subsets* of R^k one has (a complete proof may be found in [11], Corollary 3.2)

(1.49) $\sup_{C \in \mathscr{C}} \Phi((\partial C)^\varepsilon) \leqq a(k)\varepsilon \qquad \varepsilon > 0$.

It follows from (1.48), (1.49) that

(1.50) $\sup_{C \in \mathscr{C}} |Q_n(C) - \Phi(C) - \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)(C)| = o(n^{-(s-2)/2})$.

We make two more observations on Theorem 1.5. First, suppose $\mathscr{P}$ is a relatively *norm compact* class of probability measures Q_1 satisfying, in addition to the hypothesis of Theorem 1.5, the condition

(1.51) $\sup_{Q_1 \in \mathscr{P}} \int_{R^k} ||x||^{s+1} Q_1(dx) < \infty$.

It is then simple to show, using norm compactness, that on $\mathscr{P}$ the quantity θ defined in (1.46) is bounded away from one. Hence (1.41) and, therefore, (1.48), (1.50) hold uniformly over such a class $\mathscr{P}$. The second remark concerns the extension of (1.41) to unbounded f. Such an extension is possible if

(1.52) $M_s(f) \equiv \sup_{x \in R^k} (1 + ||x||^s)^{-1} |f(x)| < \infty$.

Indeed, if $M_r(f) < \infty$ for some integer r, $0 \leqq r \leqq s$, then (1.41) holds (see [11], Theorem 20.1) with $\omega_f(R^k)$ replaced by

(1.53) $M_r^*(f) \equiv 2 \inf_{c \in R^1} M_r(f - c)$.

Observe that if $M_s(f) = \infty$, then $\int f\, dQ_n$ may not exist.

Theorem 1.5 still leaves out the entire class of discrete probability measures as well as many nonatomic singular distributions. If Q_1 is of the lattice type, then $|\hat{Q}_1|$ is periodic and, consequently, the lim sup of $|\hat{Q}_1(t)|$ as $||t|| \to \infty$ is one. For an arbitrary discrete Q_1, the ch.f. $\hat{Q}_1$ is a uniform limit of trigonometric polynomials and is, therefore, *almost periodic* in the sense of Bohr; hence $\limsup |\hat{Q}_1(t)| = |\hat{Q}_1(0)| = 1$. Now it is possible to show ([11], Theorem 17.5), no matter what the type of the distribution Q_1 is, that an affine subspace of dimension m $(0 \leqq m < k)$ has Q_n measure at most $O(n^{-(k-m)/2})$ provided $\rho_3 < \infty$. If Q_1 is of the lattice type, then this bound is actually attained, and it follows that the distribution function F_n of Q_n has jumps of order $O(n^{-\frac{1}{2}})$. But $\Phi + \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)$ is absolutely continuous. Thus Theorem 1.5 can not be true in the lattice case. However, because of the lattice structure a different expansion of $Q_n(A)$ for special rectangles A may be given. This is discussed in Section 2. If no assumption is made on the type of Q_1 one may still estimate $\int f\, d(Q_n - \Phi)$.

THEOREM 1.7. *If $\rho_3 < \infty$, then for every real-valued, bounded, Borel measurable function f on R^k one has*

$$(1.54) \qquad |\int_{R^k} f\, d(Q_n - \Phi)| \leqq c_4(k)\omega_f(R^k)\rho_3 n^{-\frac{1}{2}} + c_5(k)\omega_f^*(\varepsilon_n : \Phi) ,$$

where $\varepsilon_n = c_5'(k)\rho_3 n^{-\frac{1}{2}}$.

To prove this one chooses, as in the proof of Theorem 1.5, a smoothing kernel K. However, this time the probability measure K is chosen so that $\hat{K}$ vanishes outside a compact set. This rules out the possibility of K having a compact support. Instead one requires

$$(1.55) \qquad \gamma \equiv K(\{||x|| < 1\}) > \tfrac{1}{2} , \qquad \int_{R^k} ||x||^{k+1} K(dx) < \infty .$$

Define K_ε by (1.43). Then one has, instead of Lemma 1.6 (see [8], Corollary 2.2),

LEMMA 1.8. *If G is a finite measure and H is a finite signed measure such that $G(R^k) = H(R^k)$, and if K_ε is as above, then*

$$(1.56) \quad |\int_{R^k} f\, d(G - H)| \leqq (2\gamma - 1)^{-1}[\tfrac{1}{2}\omega_f(R^k)||(G - H) * K_\varepsilon|| + \omega_f^*(2\varepsilon : |H|)] ,$$

for every bounded measurable f.

One also needs the analytical result (see [11], Lemma 11.6)

LEMMA 1.9. *There exists a positive constant $c_6(k)$ such that if g satisfies*

$$(1.57) \qquad \int_{R^k} (1 + ||x||^{k+1})|g(x)|\, dx < \infty ,$$

then

$$(1.58) \qquad ||g||_1 \equiv \int_{R^k} |g(x)|\, dx \leqq c_6(k) \max_{|\nu|=0,k+1} ||D^\nu \hat{g}||_1 .$$

Write $G = Q_n$, $H = \Phi$, $\varepsilon = c_7(k)\rho_3 n^{-\frac{1}{2}}$, and let g be the density of $(G - H) * K_\varepsilon$. Assume $\rho_{k+1} < \infty$, so that (1.58) may apply. Note that $D^\nu[(\hat{G} - \hat{H}) \cdot \hat{K}_\varepsilon](t)$ vanishes outside a sphere of radius $O(n^{\frac{1}{2}})$. Thus Theorem 1.1 is adequate in showing that the right side in (1.58) is $O(n^{-\frac{1}{2}})$. Now use Lemma 1.8 to complete the proof of Theorem 1.7 in the case $\rho_{k+1} < \infty$. Finiteness of ρ_{k+1} is assured by the hypothesis of the theorem if $k = 1$ or 2. For $k > 2$ one uses truncation (see [9] or [11]).

Letting f in (1.54) be the indicator function of a Borel set A one gets

$$(1.59) \qquad |Q_n(A) - \Phi(A)| \leqq c_8(k)\rho_3 n^{-\frac{1}{2}} + c_8'(k) \sup_{y \in R^k} \Phi((\partial A)^{\varepsilon_n} + y) .$$

In view of (1.49) and the fact that $\mathscr{C}$ is translation invariant, it follows that

$$(1.60) \qquad \sup_{C \in \mathscr{C}} |Q_n(C) - \Phi(C)| \leqq c_9(k)\rho_3 n^{-\frac{1}{2}} .$$

The inequality (1.60) is an improvement of an earlier result of Ranga Rao [37]. Inequalities (1.59) and (1.60) were proved independently by von Bahr [2] and the present author [7] under slightly more stringent moment conditions (e.g., in [7] it is assumed that $\rho_{3+\delta} < \infty$ for some $\delta > 0$). Later the present form of (1.60) was obtained by Sazonov [38]. Theorem 1.7 is due to the author [9]. An extension to unbounded f and applications to nonuniform rates of convergence and mean central limit theorems may also be found in [9].

Although Theorem 1.7 seems adequate for most applications in which no assumption is made on the type of distribution Q_1, it is still important to know if ω_f^* may be replaced by $\bar{\omega}_f$ in (1.54). Very recently, Sweeting [41] has settled this important issue by proving that this is indeed possible. That this is possible if ε_n is also replaced by $\varepsilon_n' = \varepsilon_n \log n$ was shown earlier in [7], [8].

The next theorem of this section provides a limited expansion under a relaxation of Cramér's condition (1.36). To state this we define a *strongly nonlattice* probability measure Q_1 to be one for which

$$(1.61) \qquad |\hat{Q}_1(t)| < 1 \qquad \text{for all} \quad t \neq 0 .$$

It is easy to show that in one dimension the terms *nonlattice* and *strongly nonlattice* are equivalent. This is not the case in higher dimensions. Indeed, given a real c and a nonzero vector t_0 one may easily construct a nonlattice (or even nondiscrete) probability measure Q_1 which concentrates all its mass on the countable set of hyperplanes $\{x : \langle t_0, x\rangle = c + 2n\pi\}$, $n = 0, \pm 1, \pm 2, \cdots$. Clearly $|\hat{Q}_1(t_0)| = 1$, so that $\hat{Q}_1$ is *not* strongly nonlattice.

THEOREM 1.10. *If Q_1 is strongly nonlattice and $\rho_3 < \infty$, then the relation*

$$(1.62) \qquad \int_{R^k} f\, d[Q_n - \Phi - n^{-\frac{1}{2}} P_1(-\Phi)] = o(n^{-\frac{1}{2}})$$

holds uniformly for every class $\mathscr{F}$ of functions satisfying

$$(1.63) \qquad \sup_{f \in \mathscr{F}} \omega_f(R^k) < \infty , \qquad \sup_{f \in \mathscr{F}} \omega_f^*(\varepsilon : \Phi) = O(\varepsilon) \quad \textit{as} \quad \varepsilon \downarrow 0 .$$

The proof of Theorem 1.10 is analogous to that of Theorem 1.7. Let η be any small number, and let $\varepsilon = n^{-\frac{1}{2}}\eta$. In Lemma 1.8 take $G = Q_n$, $H = \Phi + n^{-\frac{1}{2}}P_1(-\Phi)$, and let K_ε be as in the proof of Theorem 1.7. To estimate $||(G - H) * K_\varepsilon||$ use Lemma 1.8 with the density of $(G - H) * K_\varepsilon$ as g, and then apply Theorem 1.1 to estimate the integral of $|D^\nu \hat{g}| = |D^\nu[\hat{Q}_n - \hat{\Phi} - n^{-\frac{1}{2}}\hat{P}_1(-\Phi)\hat{K}_\varepsilon]|$ over a ball of radius $c_1' n^{\frac{1}{2}}$, say. This estimate is $\omega_f(R^k) \cdot o(n^{-\frac{1}{2}})$. Since $\hat{K}_\varepsilon(t) = 0$ for $||t|| > n^{\frac{1}{2}}/\eta$, one needs to estimate the integral also over the set $B_n = \{c_1' n^{\frac{1}{2}} < ||t|| \leqq n^{\frac{1}{2}}/\eta\}$. Since Q_1 is strongly nonlattice, one has

$$(1.64) \qquad \delta(u) \equiv \sup_{c' < ||t|| < u} |\hat{Q}_1(t)| < 1 \qquad u > c_1' ,$$

and $|\hat{Q}_n(t)| = |\hat{Q}_1(t/n^{\frac{1}{2}})|^n \leqq (\delta(\eta^{-1}))^n$ on B_n. But $(\delta(\eta^{-1}))^n$ goes to zero exponentially fast as $n \to \infty$, and the estimation is complete. It is also clear that a detailed knowledge of the asymptotic behavior of $\delta(\cdot)$ at infinity would enable one to refine (1.62). For example, if $\rho_4 < \infty$ and $\delta(u) = O(1 - u^{-1})$ as $u \to \infty$, then by taking one more term in the asymptotic expansion one may replace the remainder $o(n^{-\frac{1}{2}})$ by $O(n^{-1})$ in (1.62). The relation (1.62) is also uniform over every relatively norm compact class of probability measures Q_1 (strongly nonlattice and normalized) whose fourth moments are bounded away from infinity.

In many applications one needs to estimate the probability $Q_n(\{||x|| > a_n\})$ where $a_n \to \infty$ as $n \to \infty$. The estimation (1.60) is usually not adequate for this purpose. In case the *Laplace–Stieltjes transform* $\lambda \to \int \exp\{\langle\lambda, x\rangle\} Q_1(dx)$ is finite in a neighborhood of the origin, precise estimates may be given, provided $a_n = o(n^{\frac{1}{2}})$. In one dimension this was done by Khinchin [26] in a special case and Cramér [19] in the general case. For a multidimensional extension we refer to von Bahr [3]. We shall not discuss this *large deviations* theory here. For applications discussed in this article the following result (due to von Bahr [1]) is adequate.

THEOREM 1.11. *If $\rho_s < \infty$ for some integer $s \geqq 3$, then for each $\delta > 0$ one has*

$$\sup_{a \geqq (((s-2)+\delta)\log n)^{\frac{1}{2}}} a^s Q_n(\{||x|| \geqq a\}) = \theta_n n^{-(s-2)/2} \tag{1.65}$$

where θ_n goes to zero as $n \to \infty$.

To prove this we need to apply Lemma 1.6 to $G = Q_n$, $H = \Phi + \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi)$, K_ε as used in the proof of Theorem 1.5, and the function

$$\begin{aligned} f(x) &= 0 \qquad \text{if} \quad ||x|| < a \\ &= a^s \qquad \text{if} \quad ||x|| \geqq a\,. \end{aligned}$$

Also the quantity $\omega_f(R^k)$ in the bound (1.45) has to be replaced by $M_s(f)$ defined by (1.52). In this case $M_s(f) = a^s/(1 + a^s) < 1$, and $\bar{\omega}_f(2\varepsilon : |H|) = a^s|H|(\{a - 2\varepsilon < ||x|| < a + 2\varepsilon\})$. Since a is large, the average modulus of oscillation is small (namely, $o(n^{-(s-2)/2})$ if $\varepsilon = n^{-\frac{1}{2}} \log n$). The variation norm $||(G - H) * K_\varepsilon||$ is estimated as usual by appealing to Theorem 1.1 and doing a separate estimation of the integral of $D^\nu[(\hat{G} - \hat{H}) \cdot \hat{K}_\varepsilon](t)$ over the region $\{||t|| > n^{\frac{1}{2}}\}$. This last integration is facilitated by our choice of the kernel K (whose Fourier transform goes to zero fast at infinity).

2. Asymptotic expansion in the lattice case. A discrete subgroup L of R^k is a *lattice* if it is of rank k, i.e., if L has a representation

$$L = \mathbb{Z} \cdot \xi_1 + \cdots + \mathbb{Z} \cdot \xi_k = \{\textstyle\sum_{l=1}^k m_l \xi_l : m_1, \cdots, m_k \in \mathbb{Z}\}\,. \tag{2.1}$$

Here $\xi_1, \cdots, \xi_k$ are k linearly independent vectors of R^k which are said to form a *basis of L*, and $\mathbb{Z}$ is the set of all integers. A probability measure Q on R^k is *of the lattice type* (or, simply, *lattice*) if there exist a lattice L and a vector x_0 such that

$$Q(\{x_0 + L\}) = 1\,. \tag{2.2}$$

A *lattice random vector* is one whose distribution is of the lattice type. If Q is lattice and *nondegenerate* (i.e., no hyperplane carries the entire mass of Q), then there exists a smallest lattice L_0, called the *minimal lattice of Q*, such that (2.2) holds with $L = L_0$ and some x_0 (see [11], Lemma 21.4). It is obvious that if Q has finite second moments then it is nondegenerate if and only if its covariance matrix is nonsingular. If the *standard Euclidean basis* $\{e_1, \cdots, e_k\}$ is a basis of a lattice L, then $L = \mathbb{Z}^k$. For the sake of simplicity we assume below that Q_1 has a nondegenerate lattice distribution with minimal lattice $\mathbb{Z}^k$. Suppose X_1 has a nondegenerate lattice distribution whose minimal lattice has a basis $\{\xi_1, \cdots, \xi_k\}$.

Note that if T is the linear transformation mapping $\xi_1, \cdots, \xi_k$ into $e_1, \cdots, e_k$, then the random vector $Y_1 = TX_1$ has minimal lattice $\mathbb{Z}^k$. Since such a transformation changes the covariance matrix, we must now deal with an arbitrary covariance matrix V instead of the identity I. In addition, it would be convenient to take $x_0 = 0$ in (2.2). However, in order that one does not lose generality, one should then deal with an arbitrary mean vector. Throughout this section, therefore, we require that the lattice random vector Y_1 has a distribution Q with minimal lattice $\mathbb{Z}^k$ and that

$$EY_1 = \mu\,, \qquad \operatorname{Cov} Y_1 = V\,, \tag{2.3}$$

where V is nonsingular. Let then $\{Y_n\colon n \geqq 1\}$ be a sequence of i.i.d. lattice random vectors with Y_1 as specified. Let Q_n denote the distribution of $(Y_1 + \cdots + Y_n - n\mu)/n^{\frac{1}{2}}$. Write

$$\begin{aligned} y_{\alpha,n} &= \frac{\alpha - n\mu}{n^{\frac{1}{2}}}\,, \\ p_n(\alpha) &= \Pr(Y_1 + \cdots + Y_n = \alpha) = Q_n(\{y_{\alpha,n}\}) \qquad \alpha \in \mathbb{Z}^k\,, \\ q_{n,s}(x) &= n^{-k/2}[\phi_V(x) + \textstyle\sum_{r=1}^{s-1} n^{-r/2}P_r(-\phi_V)(x)]\,. \end{aligned} \tag{2.4}$$

Here ϕ_V is the *normal density* on R^k *having zero mean and covariance matrix* V, and $P_r(-\phi_V)$ is obtained by replacing ϕ by ϕ_V in (1.24). The polynomials $\tilde{P}_r$ are the same as before with the understanding that the cumulants χ_ν are now those of $Y_1 - \mu$.

THEOREM 2.1. *If* $\rho_s \equiv E||Y_1 - \mu||^s < \infty$ *for some integer* $s \geqq 2$, *then*

$$\begin{aligned} &\sup_{\alpha\in\mathbb{Z}^k} (1 + ||y_{\alpha,n}||^s)|p_n(\alpha) - q_{n,s}(y_{\alpha,n})| = o(n^{-(k+s-2)/2})\,, \\ &\textstyle\sum_{\alpha\in\mathbb{Z}^k} |p_n(\alpha) - q_{n,s}(y_{\alpha,n})| = o(n^{-(s-2)/2}) \qquad n \to \infty\,. \end{aligned} \tag{2.5}$$

In order to prove (2.5) first note that the ch.f. $\hat{Q}^n$ of $Y_1 + \cdots + Y_n$ is the multiple Fourier series

$$\hat{Q}^n(t) = \textstyle\sum_{\alpha\in\mathbb{Z}^k} \exp\{i\langle t, \alpha\rangle\}p_n(\alpha) \qquad t \in R^k\,, \tag{2.6}$$

so that

$$\begin{aligned} p_n(\alpha) &= (2\pi)^{-k} \textstyle\int_{(-\pi,\pi]^k} \hat{Q}^n(t) \exp\{-i\langle t, \alpha\rangle\}\, dt \\ &= (2\pi)^{-k}n^{-k/2} \textstyle\int_{(-n^{\frac{1}{2}}\pi,n^{\frac{1}{2}}\pi]^k} \exp\{-i\langle \tau, y_{\alpha,n}\rangle\}\hat{Q}_n(\tau)\, d\tau\,, \end{aligned} \tag{2.7}$$

changing variables $t \to \tau = n^{\frac{1}{2}}t$ in the second step. Now approximate $\hat{Q}_n$ in (2.7) by its asymptotic expansion (a change of variables will convert Theorem 1.1 into the needed expansion corresponding to an arbitrary covariance matrix V) and compare the resulting expression with

$$\begin{aligned} q_{n,s}(y_{\alpha,n}) = (2\pi)^{-k}n^{-k/2} &\textstyle\int_{R^k} \exp\{-i\langle \tau, y_{\alpha,n}\rangle\} \\ &\times [1 + \textstyle\sum_{r=1}^{s-2} n^{-r/2}\tilde{P}_r(i\tau)] \exp\{-\frac{1}{2}\langle \tau, V\tau\rangle\}\, d\tau\,. \end{aligned} \tag{2.8}$$

Similarly $y_{\alpha,n}^\nu p_n(\alpha)$ and $y_{\alpha,n}^\nu q_{n,s}(\alpha)$ are compared by inverting derivatives of $\hat{Q}_n$ and those of its expansion. The first relation in (2.5) is obtained in this way; the second follows from the first on summing over α.

The local expansion (2.5) is precise. The next problem is to find a method for summing up these approximations of point masses over sets. In one dimension Esseen [21] adapted the classical Euler–Maclaurin sum formula for this purpose. Ranga Rao [36], [37] proved a generalization of this summation formula in multidimension and used it to obtain expansions of Q_n. To explain this we introduce a sequence of functions S_j ($j = 0, 1, 2, \cdots$) on R^1 which are periodic with period one, differentiable at all nonintegral points, and satisfy

$$(2.9) \qquad S_0 \equiv 1\,, \qquad \frac{d}{dx} S_{j+1}(x) = S_j(x) \quad \text{for all} \quad x \quad \text{if} \quad j \geqq 1\,,$$
$$\frac{d}{dx} S_1(x) = 1 \quad \text{for nonintegral} \quad x\,.$$

Assume also that S_1 is right continuous. These conditions completely specify the sequence. For example,

$$(2.10) \qquad \begin{aligned} &S_1(x) = x - \tfrac{1}{2}\,, \qquad S_2(x) = \tfrac{1}{2}(x^2 - x + \tfrac{1}{6})\,, \\ &S_3(x) = \tfrac{1}{6}(x^3 - \tfrac{3}{2}x^2 + \tfrac{1}{2}x)\,, \qquad 0 \leqq x < 1\,. \end{aligned}$$

For $j \geqq 2$ the functions S_j are absolutely continuous on R^1, while S_1 has jumps -1 at all integral points. Let now f be an arbitrary real-valued function on R^1 having continuous and integrable derivatives $D^j f$, $0 \leqq j \leqq r$. Write

$$(2.11) \qquad \begin{aligned} F(x) &= \textstyle\int_{-\infty}^{x} f(t)\,dt\,, \\ F_r(x) &= \textstyle\sum_{j=0}^{r} (-1)^j S_j(x) D^j F(x) + (-1)^{r+1} \int_{-\infty}^{x} S_r(t) D^{r+1} F(t)\,dt\,. \end{aligned}$$

Then F_r is the distribution function of a finite signed measure and an integration by parts yields

$$(2.12) \qquad \textstyle\sum_{m \leqq x} f(m) = F_r(x) \qquad x \in R^1\,.$$

The summation on the left is over integers m. To extend (2.12) to multidimension consider a function f on R^k having continuous and integrable derivatives $D^\nu f$, $0 \leqq |\nu| \leqq r$. Define

$$(2.13) \qquad F(x) = \textstyle\int_{-\infty}^{x^{(1)}} \cdots \int_{-\infty}^{x^{(k)}} f(y)\,dy \qquad x = (x^{(1)}, \cdots, x^{(k)}) \in R^k\,.$$

Define operators $I_{r,j}$, $T_{r,j}$ acting on such functions F by

$$(2.14) \qquad \begin{aligned} I_{r,j}(F)(x) &= \textstyle\int_{-\infty}^{x^{(j)}} S_r(t)(D_j^{r+1}F)(x^{(1)}, \cdots, x^{(j-1)}, t, x^{(j+1)}, \cdots, x^{(k)})\,dt \\ &= \textstyle\int_{-\infty}^{x^{(1)}} \cdots \int_{-\infty}^{x^{(k)}} S_r(y^{(j)}) D_j{}^r f(y)\,dy \qquad x = (x^{(1)}, \cdots, x^{(k)})\,, \\ T_{r,j}(F) &= (1 - S_1(x^{(j)})D_j + \cdots + (-1)^r S_r(x^{(j)}) D_j{}^r \\ &\quad + (-1)^{r+1} I_{r,j})(F)\,. \end{aligned}$$

Since the operators $T_{r,j}$ are associative and commutative one may define

$$(2.15) \qquad \begin{aligned} F_r(x) &= (\textstyle\prod_{j=1}^{k} T_{r,j})(F)(x) \\ &= \textstyle\prod_{j=1}^{k} \{1 - S_1(x^{(j)})D_j + \cdots + (-1)^r S_r(x^{(j)}) D_j{}^r \\ &\quad + (-1)^{r+1} I_{r,j}\}(F)(x) \qquad x = (x^{(1)}, \cdots, x^{(k)})\,. \end{aligned}$$

Again one may show that F_r is of bounded variation; and an induction on k using (2.12) yields

$$\sum_{\{\alpha^{(1)}\leqq x^{(1)},\cdots,\alpha^{(k)}\leqq x^{(k)}\}} f(\alpha) = F_r(x) \qquad x = (x^{(1)}, \cdots, x^{(k)})\,. \tag{2.16}$$

The summation on the left is over integral vectors $\alpha = (\alpha^{(1)}, \cdots, \alpha^{(k)})$. To apply this result to our specific situation define

$$f(x) = q_{n,s}\left(\frac{x - n\mu}{n^{\frac{1}{2}}}\right) \qquad x \in R^k\,, \tag{2.17}$$

and obtain (taking $r = s - 1$ in (2.16))

$$\begin{aligned}\sum_{\{\alpha:\alpha\leqq n^{\frac{1}{2}}x+n\mu\}} q_{n,s}\left(\frac{\alpha - n\mu}{n^{\frac{1}{2}}}\right) = \prod_{j=1}^{k}\{1 &- S_1(n\mu^{(j)} + n^{\frac{1}{2}}x^{(j)})D_j + \cdots \\ &+ (-1)^{s-1}S_{s-1}(n\mu^{(j)} + n^{\frac{1}{2}}x^{(j)})D_j^{s-1} \\ &+ (-1)^s I_{s-1,j}\}(F)(n^{\frac{1}{2}}x + n\mu)\,.\end{aligned} \tag{2.18}$$

By expanding the product in (2.18) and omitting terms of order $O(n^{-j/2})$, $j \geqq s - 1$, one has

$$\begin{aligned}&\sum_{\{\alpha\leqq n^{\frac{1}{2}}x+n\mu\}} q_{n,s}\left(\frac{\alpha - n\mu}{n^{\frac{1}{2}}}\right) \\ &\quad = \sum_{|\nu|\leqq s-2} n^{-|\nu|/2}(-1)^{|\nu|}S_\nu(n\mu + n^{\frac{1}{2}}x)D^\nu\Phi_V(x) \\ &\qquad + n^{-\frac{1}{2}}\sum_{|\nu|\leqq s-3} n^{-|\nu|/2}(-1)^{|\nu|}S_\nu(n\mu + n^{\frac{1}{2}}x)D^\nu P_1(-\Phi_V)(x) + \cdots \\ &\qquad + n^{-(s-2)/2}P_{s-2}(-\Phi_V)(x) + o(n^{-(s-2)/2})\end{aligned} \tag{2.19}$$

uniformly for $x \in R^k$. Here for each multiindex $\nu = (\nu^{(1)}, \cdots, \nu^{(k)})$

$$S_\nu(x) = S_{\nu^{(1)}}(x^{(1)}) \cdots S_{\nu^{(k)}}(x^{(k)}) \qquad x = (x^{(1)}, \cdots, x^{(k)})\,. \tag{2.20}$$

Combining (2.5) and (2.19) one has

THEOREM 2.2. *If $\rho_s < \infty$ for some integer $s \geqq 3$ and F_n denotes the distribution function of Q_n, then*

$$\begin{aligned}\sup_{x\in R^k} |F_n(x) &- \sum_{|\nu|\leqq s-2} n^{-|\nu|/2}(-1)^{|\nu|}S_\nu(n\mu + n^{\frac{1}{2}}x)D^\nu\Phi_V(x) \\ &- n^{-\frac{1}{2}}\sum_{|\nu|\leqq s-3} n^{-|\nu|/2}(-1)^{|\nu|}S_\nu(n\mu + n^{\frac{1}{2}}x)D^\nu P_1(-\Phi_V)(x) - \cdots \\ &- n^{-(s-2)/2}P_{s-2}(-\Phi_V)(x)| = o(n^{-(s-2)/2})\,.\end{aligned} \tag{2.21}$$

Note that if μ_n denotes the signed measure whose distribution function appears on the right side of (2.18), then (by virtue of (2.18) and (2.5))

$$|Q_n(A) - \mu_n(A)| = o(n^{-(s-2)/2}) \tag{2.22}$$

uniformly over all Borel sets A. Unfortunately, it has not been possible so far to obtain computable expressions of $\mu_n(A)$ for sets A other than rectangles whose sides are parallel to the hyperplanes $\{x^{(j)} = 0\}$, $1 \leqq j \leqq k$. In case Y_1 has a minimal lattice with basis $\{\xi_1, \cdots, \xi_k\}$, Theorem 2.2 is easily modified to apply to rectangles whose sides are parallel to the hyperplanes $\{x \equiv \sum_1^k y^{(j)}\xi_j : y^{(l)} = 0\}$, $1 \leqq l \leqq k$. The difficulty is caused by the presence of terms involving S_1. One of the most outstanding problems in the subject is to obtain "good" estimations of $Q_n(A)$ in the lattice case for sets A other than rectangles properly aligned with the lattice. Perhaps the nature of the problem is best appreciated by linking it with the lattice point problem of analytic number theory. We do this in the following section.

3. The lattice point problem. Confining ourselves to the standard lattice $\mathbb{Z}^k$, we define a *lattice-point* as a point in $\mathbb{Z}^k$. Let V be a positive definite symmetric matrix and consider the ellipsoids

$$E(c: V) = \{x \in R^k : \langle x, V^{-1}x\rangle \leqq c\} \qquad c > 0 . \tag{3.1}$$

Let $N(c: V)$ denote the number of lattice points in $E(c: V)$. An important problem in analytic number theory is to obtain asymptotic estimates of $N(c: V)$ as $c \to \infty$. If no further specification is made on k and V, then the best known result is that of Landau, namely,

$$|N(c: V) - \text{volume of } E(c: V)| = O(c^{k/2 - k/k+1}) \qquad c \to \infty . \tag{3.2}$$

Esseen [21] showed that (3.2) is essentially equivalent to the following theorem specialized to lattice random vectors.

THEOREM 3.1. *If* $\{Y_n : n \geqq 1\}$ *is a sequence of* i.i.d. *random vectors each with mean* μ, *covariance matrix* V, *and a finite fourth absolute moment, then*

$$\sup_{a \geqq 0} |Q_n(E(a: V)) - \Phi_V(E(a: V))| = O(n^{-k/(k+1)}) , \tag{3.3}$$

where Q_n *is the distribution of* $n^{-\frac{1}{2}} \sum_{j=1}^{n} (Y_j - \mu)$.

The proof of Theorem 3.1 is rather long and is given in [21]. We shall only give a sketch of Esseen's argument linking (3.2) and (3.3). Note that the right side in (3.3) goes to zero faster than $n^{-\frac{1}{2}}$ if $k > 1$. If the distribution Q_1 satisfies Cramér's condition (1.36), then a faster rate of convergence (with an error $O(n^{-1})$) may be obtained from (1.50) with $s = 4$. Here one uses the fact that $P_1(-\phi_V)$ is an odd function and, therefore,

$$P_1(-\Phi_V)(E(a: V)) = P_1(-\Phi)(E(a: I)) = 0 . \tag{3.4}$$

The strength of Esseen's result lies, however, in its generality. For example, suppose Y_1 in the theorem is lattice having $\mathbb{Z}^k$ as its minimal lattice. Without loss of generality assume $\Pr(Y_1 \in \mathbb{Z}^k) = 1$. The local expansion (2.5) with $s = 4$ yields (we take $\mu = 0$ for simplicity)

$$\begin{aligned} \sup_{a \geqq 0} \Big| Q_n(E(a: V) &- (\det V)^{-\frac{1}{2}}(2\pi n)^{-k/2} \\ &\times \sum_{\{\alpha \in \mathbb{Z}^k : \langle \alpha - n\mu, V^{-1}(\alpha - n\mu)\rangle \leqq an\}} \exp\left\{-\frac{1}{2n}\langle \alpha - n\mu, V^{-1}(\alpha - n\mu)\rangle\right\}\Big| \\ &= O(n^{-1}) , \end{aligned} \tag{3.5}$$

again because $P_1(-\phi_V)$ is an odd function and the set of vectors $y_{\alpha,n} = (\alpha - n\mu)/n^{\frac{1}{2}}$ over which $P_1(-\phi_V)(y_{\alpha,n})$ is to be summed is symmetric. Combining (3.3) and (3.5) one gets

$$\sup_{a\geq 0}\Big|(\det V)^{-\frac{1}{2}}(2\pi n)^{-k/2} \times \sum_{\{\langle\alpha-n\mu, V^{-1}(\alpha-n\mu)\rangle\leq an\}} \exp\left\{-\frac{1}{2n}\langle\alpha-n\mu, V^{-1}(\alpha-n\mu)\rangle\right\} - (\det V)^{-\frac{1}{2}}(2\pi)^{-k/2}\int_{\{\langle x, V^{-1}x\rangle\leq a\}}\exp\{-\tfrac{1}{2}\langle x, V^{-1}x\rangle\}\,dx\Big| = O(n^{-k/(k+1)})\,. \tag{3.6}$$

Now write $N(u)$ for the number of lattice points in $E(u: V) + n\mu = \{x: \langle x - n\mu, V^{-1}(x - n\mu)\rangle \leq u\}$, $u > 0$. Also write $B(u)$ for the volume of $E(u: V)$, and let

$$R(u) = N(u) - B(u)\,. \tag{3.7}$$

Then (3.6) reduces to

$$\sup_{a\geq 0}\left|(\det V)^{-\frac{1}{2}}(2\pi n)^{-k/2}\int_{[0,an]}\exp\left\{-\frac{u}{2n}\right\}dR(u)\right| = O(n^{-k/(k+1)})\,. \tag{3.8}$$

An integration by parts immediately gives

$$\sup_{a\geq 0}\left|(\det V)^{-\frac{1}{2}}(2\pi n)^{-k/2}\left[e^{-a/2n}R(an) + \frac{1}{2n}\int_{[0,an]}e^{-u/2n}R(u)\,du\right]\right| = O(n^{-k/(k+1)})\,. \tag{3.9}$$

Since $(an)^{-1}\int_{[0,an]}|R(u)|\,du$ is of order not larger than that of $|R(an)|$, (3.9) leads to

$$R(an) = O(n^{k/2-k/(k+1)}) \tag{3.10}$$

for all $a > 0$. Landau's result (3.2) follows from (3.10). Conversely, on retracing the steps one can deduce (3.3) for lattice random vectors from (3.2). More precisely, it has been shown by Yarnold [43] using the expansion in Section 2 that in the lattice case one has

$$\sup_{a\geq 0}|Q_n(E(a: V)) - \Phi_V(E(a: V)) - R(an)e^{-a/2}(2\pi n)^{-k/2}(\det V)^{-\frac{1}{2}}| = O(n^{-1})\,. \tag{3.11}$$

It is a simple consequence of (3.2) that the number of lattice points on the surface $\{x: \langle x, V^{-1}x\rangle = c\}$ is of the order $O(c^{k/2-k/(k+1)})$ as $c \to \infty$. Also observe that we can derive a weaker estimate $O(c^{(k-1)/2})$ for this as well as for the remainder in (3.2) more simply from the inequality (1.60) without appealing to Theorem 3.1 or the material in Section 2.

The foregoing discussion virtually rules out the possibility of obtaining computable expansions of $Q_n(A)$ in the lattice case except for sets A properly aligned with the lattice. Under the circumstances perhaps the best one can hope for is an extension of Theorem 3.1. Of course, for sets A which are not symmetric the analogue of (3.3) is

$$|Q_n(A) - \Phi_V(A) - n^{-\frac{1}{2}}P_1(-\Phi_V)(A)| = O(n^{-k/(k+1)})\,. \tag{3.12}$$

Recently, Matthes [30] proved this for a class of convex bodies A having sufficiently smooth surfaces whose Gaussian curvatures are bounded away from zero and infinity. The result is delicate. Note that it does not hold for rectangles. It would appear that (3.12) will not hold if ∂A contains too many points of the lattice $n^{-k/2}\mathbb{Z}^k$. Technically, the proof by Matthes uses an estimate of Esseen [21] on the value distribution of $|\hat{Q}_1|$ and a result of Herz [23] asserting

$$(3.13) \qquad |\hat{I}_A(t)| = O(||t||^{-(k+1)/2}) \qquad ||t|| \to \infty$$

for the convex sets A discussed by Matthes. Since convexity is a bothersome restriction, it would be useful to extend the result of Matthes by proving (3.13) for other smooth sets A.

The lattice point problem is intimately related to the asymptotic distribution of eigenvalues of self adjoint elliptic operators in the theory of partial differential equations. For a delightful discussion of this we refer to Courant and Hilbert [18] (pages 429–445).

4. Asymptotic distributions of a class of statistics. In this section we briefly sketch the derivation of normal approximations and asymptotic expansions of a class of statistics commonly used for purposes of statistical inference. Simplest examples of such statistics are functions of sample moments. Detailed proofs will appear elsewhere. Theorem 4.1 is based on joint work with J. K. Ghosh. Theorem 4.2 and the expansion (4.15) were obtained by Chibisov [15] under more restrictive assumptions on the functions f_i $(1 \leqq i \leqq k)$ below using different methods. While we merely require differentiability, Chibisov [15] assumes analyticity of these functions, but obtains estimates of the variation norm.

Let $\{Y_n \equiv (Y_n^{(1)}, \cdots, Y_n^{(m)}) : n \geqq 1\}$ be a sequence of i.i.d. random vectors with values in R^m $(m \geqq 1)$. Let G denote their common distribution. We introduce real-valued, Borel measurable functions $f_1, \cdots, f_k$ on R^m and assume

A_1: $E|f_i(Y_1)|^s < \infty$, $1 \leqq i \leqq k$. *Here s is a positive integer, $s \geqq 3$.*

A_2: *H is a real-valued Borel measurable function defined on a neighborhood N of*

$$(4.1) \qquad \mu \equiv (Ef_1(Y_1), \cdots, Ef_k(Y_1)) .$$

H has bounded and continuous derivatives of order p_0 or less in some neighborhood $M(\subset N)$ of μ. Here $p_0 \geqq 2$. Also,

$$(4.2) \qquad (\text{grad } H)(\mu) \equiv (D_1 H, \cdots, D_k H)(\mu) \neq 0 .$$

The derivatives of H at μ are denoted by

$$(4.3) \qquad \begin{aligned} &l_j = (D_j H)(\mu) \quad (1 \leqq j \leqq k), \quad l = (l_1, \cdots, l_k) ; \\ &l_{i_1 \cdots i_p} = (D_{i_1} \cdots D_{i_p} H)(\mu) \quad (1 \leqq i_1, \cdots, i_p \leqq k; p \leqq p_0) . \end{aligned}$$

Also write

$$(4.4) \qquad \begin{aligned} &Z_n = (f_1(Y_n), \cdots, f_k(Y_n)) , \qquad \bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i , \\ &W_n = n^{\frac{1}{2}}[H(\bar{Z}) - H(\mu)] , \end{aligned}$$

and introduce the function

$$g_n(z) = n^{\frac{1}{2}}\left[H\left(\mu + \frac{z}{n^{\frac{1}{2}}}\right) - H(\mu)\right]. \tag{4.5}$$

By extending H, if necessary, arbitrarily (but measurably) over all of R^k we make W_n and g_n well defined. Note that $\{Z_n : n \geqq 1\}$ is an i.i.d. sequence, $EZ_1 = \mu$. Let

$$V = \operatorname{Cov} Z_1 . \tag{4.6}$$

It is easy to see that V is singular if and only if the functions $1, f_1, \cdots, f_k$ are linearly dependent on the support of G, i.e., if and only if there exist $\theta^{(1)}, \cdots, \theta^{(k)}$, $c \in R^1$, not all zero, such that

$$G(\{y : \textstyle\sum_{i=1}^k \theta^{(i)} f^{(i)}(y) = c\}) = 1 . \tag{4.7}$$

Let λ, Λ *denote the smallest and largest eigenvalues of* V, respectively. Recall that ϕ_V is the normal density on R^k with mean zero and covariance V. The following result holds.

THEOREM 4.1. *If* A_1 *holds with* $s = 3$, A_2 *holds with* $p_0 = 2$, *and if* V *is nonsingular, then*

$$\sup_{u \in R^1} |\Pr(W_n \leqq u) - \textstyle\int_{\{g_n(z) \leqq u\}} \phi_V(z)\, dz| \leqq dn^{-\frac{1}{2}} , \tag{4.8}$$

where d *depends only on the moments of* Z_1 *of orders three or less and on the first order derivatives of* H *on* M.

In order to prove this theorem one cannot appeal to (1.59) directly, since, in general, V is not the identity matrix. But transforming the random vectors $Z_1, \cdots, Z_n$, by a nonsingular linear transformation one easily obtains (from (1.59))

$$\begin{aligned} |Q_n(A) - \Phi_V(A)| &\leqq c_8(k)\lambda^{-\frac{3}{2}} E\|Z_1 - \mu\|^3 n^{-\frac{1}{2}} \\ &\quad + c_8'(k) \sup_{y \in R^k} \Phi_V((\partial A)^\eta + y) , \\ \eta &= c_6(k)\lambda^{-\frac{3}{2}} E\|Z_1 - \mu\|^3 n^{-\frac{1}{2}} , \end{aligned} \tag{4.9}$$

where Q_n is the distribution of $n^{-\frac{1}{2}}(Z_1 + \cdots + Z_n - n\mu)$. To apply (4.9) one may take $A = \{z \in R^k : g_n(z) \leqq u\}$ or, in view of Theorem 1.11, its restriction to the set $\{\|z\| \leqq ((s-1)\Lambda \log n)^{\frac{1}{2}}\}$. A fairly straightforward computation yields

$$\sup_{y \in R^k} \Phi_V((\partial A)^\varepsilon + y) \leqq d'\varepsilon \qquad \varepsilon > 0 , \tag{4.10}$$

uniformly in u, and (4.8) follows from (4.9) and (4.10).

To obtain asymptotic expansions going beyond (4.8) we assume

A_3: *The distribution* G *of* Y_1 *(or* G^{*r} *for some positive integer* r*) has a nonzero absolutely continuous component* H. *Further, there exists a nonempty open set* B *of* R^m *on which the density of* H *is positive and the functions* $1, f_1, \cdots, f_k$ *are continuously differentiable and linearly independent.*

Note that if A_3 holds, then by Lemma 1.4 the distribution of $Z_1 - \mu$ satisfies Cramér's condition (1.36). Theorem 1.5 then implies (by a linear transformation of Z_j's)

THEOREM 4.2. *Assume* A_1, A_2, A_3 *hold with* $p_0 = s \geqq 3$. *Then*

$$\sup_{u\in R^1} |\Pr(W_n \leqq u) - \int_{\{g_n(z)\leqq u\}} [\phi_V(z) + \sum_{r=1}^{s-2} n^{-r/2} P_r(-\phi_V)(z)]\,dz| = o(n^{-(s-2)/2}) \qquad n \to \infty\,. \tag{4.11}$$

Since the domain of integration $\{g_n(z) \leqq u\}$ is not simple to deal with we now provide a more computable expression for the expansion. For this we first introduce the function

$$h_{s-1}(z) = \sum_{j=1}^{k} l_j z^{(j)} + \frac{1}{2n^{\frac{1}{2}}} \sum_{1\leqq i,j\leqq k} l_{ij} z^{(i)} z^{(j)} + \cdots + \frac{1}{(s-1)!\, n^{(s-2)/2}} \sum_{1\leqq i_1,\cdots,i_{s-1}\leqq k} l_{i_1\cdots i_{s-1}} z^{(i_1)} \cdots z^{(i_{s-1})} \tag{4.12}$$

and note that h_{s-1} is a Taylor expansion of g_n and, therefore, for all constants $c > 0$,

$$\sup_{\{||z||<c \log^{\frac{1}{2}} n\}} |g_n(z) - h_{s-1}(z)| = O(n^{-(s-1)/2} \log^{s/2} n)\,. \tag{4.13}$$

In view of (4.13) and Theorem 1.11 one may replace g_n by h_{s-1} and the random variable W_n by

$$W_n' \equiv h_{s-1}(n^{-\frac{1}{2}}(Z_1 + \cdots + Z_n - n\mu))\,. \tag{4.14}$$

The next task is to derive the expansion

$$\int_{\{h_{s-1}(z)\leqq u\}} \psi(z)\,dz = \int_{\{\sum_1^k l_j z^{(j)}\leqq u\}} [\psi(z) + \sum_{r=1}^{s-2} n^{-r/2}\psi_r(z)]\,dz + O(n^{-(s-1)/2})\,, \tag{4.15}$$

where $\psi, \psi_1, \cdots, \psi_r$ are polynomial multiples of ϕ_V whose coefficients do not depend on n. This may be done by an appropriate change of variables; but we omit the details. For example, one may easily show (assuming $l_k \neq 0$)

$$\int_{\{h_2(z)\leqq u\}} \psi(z)\,dz = \int_{\{\sum l_j z^{(j)}\leqq u\}} \left[\psi(z)\left(\left(1 - \frac{\sum_j l_{kj} z^{(j)}}{l_k n^{\frac{1}{2}}}\right) - (D_k\psi)(z)\frac{\sum_{i,j} l_{ij} z^{(i)} z^{(j)}}{2 l_k n^{\frac{1}{2}}}\right)\right] dz + O(n^{-1})\,. \tag{4.16}$$

Applying (4.13) and (4.16) in Theorem 4.1 one obtains

$$\sup_{u\in R^1} |\Pr(W_n \leqq u) - \int_{-\infty}^{u} \phi_{\tilde\sigma^2}(v)\,dv| \leqq d' n^{-\frac{1}{2}}\,, \tag{4.17}$$

where d' is a positive constant and

$$\tilde\sigma^2 = \langle l, Vl\rangle\,, \qquad \phi_{\tilde\sigma^2}(v) = \frac{1}{(2\pi)^{\frac{1}{2}}\tilde\sigma} \exp\left\{-\frac{v^2}{2\tilde\sigma^2}\right\}. \tag{4.18}$$

Similarly from Theorem 4.2 one gets

$$\sup_{u\in R^1}\left|\Pr(W_n \leqq u) - \int_{\{\langle l,z\rangle\leqq u\}}\left[\phi_V(z)\left(1 - \frac{1}{l_k n^{\frac{1}{2}}}\sum_j l_{kj} z^{(j)}\right) - \frac{1}{2 l_k n^{\frac{1}{2}}}(\sum_{i,j} l_{ij} z^{(i)} z^{(j)})(D_k\phi_V)(z) + \frac{1}{n^{\frac{1}{2}}} P_1(-\phi_V)(z)\right] dz\right| = O(n^{-1})\,, \tag{4.19}$$

if A_1, A_2, A_3 hold with $p_0 = s = 4$.

By a linear transformation $z \to x$, with $x^{(1)}(z) = \langle l, z\rangle$, the right side of (4.15) may be reduced by integration to yield

$$(4.20) \qquad \int_{\{h_{s-1}(z)\leq u\}} \{\phi_V(z) + \sum_{r=1}^{s-1} n^{-r/2}P_r(-\phi_V)(z)\}\, dz \\ - \int_{-\infty}^{u} [1 + \sum_{r=1}^{s-2} n^{-r/2}q_r(v)]\phi_{\tilde\sigma^2}(v)\, dv = O(n^{-(s-1)/2}),$$

where $q_1, \cdots, q_{s-2}$ are polynomials whose coefficients do not depend on n. To identify the polynomials q_j's we describe another formal procedure for expanding the distribution function of W_n. Since W_n may not have finite moments of orders up to s, a formal method for computing "approximate cumulants" of W_n is used. This is the so-called *delta method.* Assume, for the sake of simplicity, that Z_1 has finite moments of all orders. Since h_{s-1} is a polynomial of degree $s-1$, the moments and cumulants of W_n' can be computed in terms of those of $Z_1 - \mu$ either directly (algebraically), or using Theorem 1.5 with $f = h_{s-1}$, or using Theorem 1.1. One may show

$$(4.21) \qquad j\text{th cumulant of } \quad W_n' = K_{j,n} + o(n^{-(s-2)/2}),$$

where

$$(4.22) \qquad \begin{aligned} K_{1,n} &= \sum_{i=1}^{s-2} n^{-i/2}b_{1i}, \\ K_{j,n} &= n^{-(j-2)/2}K_j + \sum_{j-1}^{s-2} n^{-i/2}b_{ji} \qquad j \geqq 2, \\ K_j &= j\text{th cumulant of } \quad \langle l, Z_1 - \mu\rangle, \end{aligned}$$

and b_{ji}'s depend only on cumulants of $Z_1 - \mu$ of orders s^2 and less. Also note that $K_2 = \tilde\sigma^2$. Now write

$$(4.23) \qquad \exp\left\{itK_{1,n} + \frac{(it)^2}{2}(K_{2,n} - \tilde\sigma^2) + \sum_{r=3}^{s} \frac{(it)^r}{r!} K_{r,n}\right\} \\ = 1 + \sum_{r=1}^{s-2} n^{-r/2}\Pi_r(it) + o(n^{-(s-2)/2}) \qquad t \in R^1,$$

where Π_r's are polynomials whose coefficients depend only on the cumulants of $Z_1 - \mu$ of orders s and less. One would then expect (as in the case of the Edgeworth expansion in Section 1)

$$(4.24) \qquad \Pr(W_n \leqq u) = \int_{-\infty}^{u} [1 + \sum_{r=1}^{s-2} n^{-r/2}\Pi_r(-D)]\phi_{\tilde\sigma^2}(v)\, dv + o(n^{-(s-2)/2}),$$

where $\Pi_r(-D)$ is the differential operator obtained by formally substituting $(-1)^j D^j$ for $(it)^j$ in the polynomial $\Pi_r(it)$, $j \geqq 0$. The integrands on the right sides of (4.20) and (4.24) are identical, i.e.,

$$(4.25) \qquad q_r(v) = \phi_{\tilde\sigma^2}^{-1}(v) \cdot \Pi_r(-D)\phi_{\tilde\sigma^2}(v) \qquad r \geqq 1.$$

One proves this by showing that the two densities under the integral signs in (4.20) and (4.24) have the same moments of all orders.

We refer to Wallace [42] for a description of the original conjecture about the validity of an expansion analogous to (4.24) using $n^{-(j-2)/2}K_j$ instead of $K_{j,n}$. In [12] Bickel modified this conjecture essentially in its present form. We do emphasize, however, that the moments and cumulants of W_n are not quite relevant for the above expansion; for the asymptotic distribution of W_n depends only on the local behavior of H at μ. Bickel [12] also discusses the possibility of applying the above expansion to other types of statistics. The main problem, of course, is to prove that there exists an expansion.

The results of this section extend in a fairly straightforward way to vector-valued functions H, and to probabilities of other sets of interest (not merely intervals or rectangles).

5. Other applications, extensions. A number of applications other than those discussed in Sections 3, 4 are listed below.

(a) *U-statistics.* Suppose $\{X_n : n \geqq 1\}$ is a sequence of i.i.d. observations with values in some space S. Let ψ be a real or vector-valued function on $S \times S$ such that $\psi(x, y) = \psi(y, x)$. The function $U_n = \binom{n}{2}^{-1} \sum_{1 \leqq i < j \leqq n} \psi(X_i, X_j)$ is a *U-statistic with kernel* ψ. Subtracting the expectation if necessary, we assume that $E\psi(X_1, X_2) = 0$. Also suppose $E||\psi(X_1, X_2)||^2 < \infty$. Let $\psi_1(x) = E\psi(X_1, x)$; then $\{\psi_1(X_n) : n \geqq 1\}$ is an i.i.d. sequence. By comparing U_n with $S_n = n^{-1} \sum_1^n \psi_1(X_i)$, Höeffding (see [35], page 58) showed that as $n \to \infty$ the statistic $n^{\frac{1}{2}} U_n$ converges in distribution to Φ_V, where $V = \operatorname{Cov} \psi_1(X_1)$. By an attractive argument Bickel [12] has recently shown that if ψ is real-valued and bounded, then

$$\sup_{u \in R^1} \left| \Pr(n^{\frac{1}{2}} U_n \leqq u) - \frac{1}{(2\pi)^{\frac{1}{2}} \sigma} \int_{-\infty}^{u} e^{-v^2/2\sigma^2} \, dv \right| = O(n^{-\frac{1}{2}}) , \tag{5.1}$$

where $\sigma^2 = E\psi_1^2(X_1)$. It would be useful to relex Bickel's assumption of boundedness of ψ, to extend (5.1) to vector-valued ψ, and, more importantly, to obtain an asymptotic expansion under appropriate assumptions. There are similar important problems concerning the so-called *rank statistics* (see [12]).

(b) *Maximum likelihood estimators.* Let $[X_n : n \geqq 1\}$ be a sequence of i.i.d. observations from a distribution with a strictly positive density $f(x; \theta)$ (relative to some σ-finite measure), where the *parameter* θ lies in an open subset of R^k (or, more generally, in a k-dimensional manifold). Assume that f is twice differentiable in θ and that the *likelihood equations* (in θ)

$$\sum_{i=1}^{n} \frac{\partial \log f(X_i; \theta)}{\partial \theta^{(j)}} = 0 \qquad 1 \leqq j \leqq k \tag{5.2}$$

have a unique solution $\hat{\theta}_n$, the *maximum likelihood estimator of* θ. If the *information matrix* $I(\theta) = -((E_\theta D_i D_j \log f(X_1, \theta)))$ is nonsingular, then under regularity assumptions one shows that $T_n = n^{\frac{1}{2}}(\hat{\theta}_n - \theta)$ is asymptotically normal Φ_V, where $V = I^{-1}(\theta)$. For the case $k = 1$ Berry–Esseen bounds and asymptotic expansions of the distribution function of T_n have been obtained by Linnik and Mitrofanova [29] and Pfanzagl [34]. A complete derivation for multidimensional parameters is still not available.

An entirely analogous problem arises in mathematical economics [10]. Here the summands in (5.2) are *excess demands* of individuals, θ is the (normalized) price vector. The solution $\hat{\theta}_n$ is the *equilibrium price*. One is interested in the asymptotic behavior of $\hat{\theta}_n$ when n, the number of agents in the economy, is large.

(c) *Law of the iterated logarithm.* The classical law of the iterated logarithm (LIL) is essentially tied up with the central limit theorem. Indeed, a very useful method of proving LIL's is by using the classical Berry–Esseen theorem. This method is originally due to Chung [16] (also see [17] pages 231–237) and was later rediscovered by Petrov [33]. It can be used to derive classical as well as Strassen type LIL's for independent as well as dependent random variables with the help of such Berry–Esseen type bounds as obtained by Statulevicius [39] and Stein [40]. This method is comparable in effectiveness with that using the Skorokhod representation (of successive partial sums of a sequence of random variables as values of the Brownian path at appropriately defined successive stopping times).

(d) *Statistical mechanics.* In his important works on the mathematical foundations of classical statistical mechanics Khinchin [27] used (his own) results on refinements of the central limit theorem to provide an analytical derivation of the Gibbs canonical ensemble and the laws of classical thermodynamics. Khinchin's book [27] is still one of the most penetrating studies on the foundations of equilibrium statistical mechanics of ideal gasses.

We conclude this article with a few additional remarks. First, the main theorems of Sections 1 and 2 have appropriate analogs in the non-i.i.d. case; these analogs may be found in the cited references. Secondly, note that if $\{X_n : n \geqq 1\}$ is an i.i.d. sequence of k-dimensional random vectors such that X_1 has independent coordinates, then obtaining rates of convergence and asymptotic expansions of the distribution Q_n (of the normalized partial sum) reduces to a one-dimensional problem. This is obvious if one is approximating the distribution function of Q_n; but even for more general sets (e.g., the class $\mathscr{C}$ of Borel measurable convex sets) one only needs to use the classical Berry–Esseen theorem and the Fubini theorem. Thus one can easily show (see [8], Theorem 4.7)

$$\sup_{C \in \mathscr{C}} |Q_n(C) - \Phi(C)| \leqq 2c_0 \left(\sum_{i=1}^{k} E|X_1^{(i)}|^3\right) n^{-\frac{1}{2}}, \tag{5.3}$$

where the universal constant c_0 is the one appearing in the Berry–Esseen bound (see Van Beek [4] for an estimation $c_0 = .7975$). Thus, in our context, the complexity of higher dimensionality arises only through the dependence among coordinate variables.

As a third remark it may be mentioned that errors of normal approximation have also been estimated by methods different from the Fourier analytic method used here (e.g., see [5], [31], [38], [41]). Because these methods are somewhat more direct it is possible that they will yield better estimates of constants involved in the bounds. However, none of these other methods have been successful in providing asymptotic expansions. Our final remark concerns the moment condition "$\rho_3 < \infty$" in Theorem 1.7. Rates of convergence can be obtained when $\rho_{2+\delta}$ is assumed finite for some δ, $0 \leqq \delta < 1$ (see, e.g., Section 18 in [11]). For distribution functions in one dimension definitive results have been obtained in this case by Heyde [24] and Ibragimov [25].

Acknowledgment. The author is indebted to Professor J. K. Ghosh for discussions on Section 4. He wishes to thank Professor V. Paulauskas for some helpful comments and for drawing the author's attention to the articles [15], [44]. Finally, thanks are due to Professors Ron Pyke and Patrick Billingsley for their kind suggestions.

REFERENCES

[1] BAHR, B. VON (1967). On the central limit theorem in R_k. *Ark. Mat.* **7** 61–69.

[2] BAHR, B. VON (1967). Multi-dimensional integral limit theorems. *Ark. Mat.* **7** 71–88.

[3] BAHR, B. VON (1967). Multi-dimensional integral limit theorems for large deviations. *Ark. Mat.* **7** 89–99.

[4] BEEK, P. VAN (1972). An application of the Fourier method to the problem of sharpening the Berry-Esseen inequality. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **23** 187–197.

[5] BERGSTRÖM, H. (1945). On the central limit theorem in the space R_k, $k > 1$. *Skad. Aktuarietidskr.* **28** 106–127.

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

[7] BHATTACHARYA, R. N. (1968). Berry-Esseen bounds for the multi-dimensional central limit theorem. *Bull. Amer. Math. Soc.* **74** 285–287.

[8] BHATTACHARYA, R. N. (1972). Recent results on refinements of the central limit theorem. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 453–484. Univ. of California Press.

[9] BHATTACHARYA, R. N. (1975). On errors of normal approximation. *Ann. Probability* **3** 815–828.

[10] BHATTACHARYA, R. N. and MAJUMDAR, M. (1973). Random exchange economies. *J. Econ. Theory* **6** 37–67.

[11] BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

[12] BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

[13] BIKJALIS, A. (1968). Asymptotic expansions of distribution functions and the density functions of sums of independent and identically distributed random vectors. *Litovsk. Mat. Sb.* **6** 405–422 (in Russian).

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

[15] CHIBISOV, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion. *Theor. Probability Appl.* **17** 620–630.

[16] CHUNG, K. L. (1950). *Notes on Limit Theorems.* Columbia Univ. Graduate Mathematical Statistical Society (mimeo).

[17] CHUNG, K. L. (1974). *A Course in Probability Theory* (2nd ed.). Academic Press, New York.

[18] COURANT, R. and HILBERT, D. (1953). *Methods of Mathematical Physics* I. Wiley, New York.

[19] CRAMÉR, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilitiés. *Actualités Sci. Indust.* **736**.

[20] CRAMÉR, H. (1962). *Random Variables and Probability Distributions* (2nd ed.). Cambridge Univ. Press.

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

[22] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions of Sums of Independent Random Variables.* English translation by K. L. Chung. Addison-Wesley, Reading, Mass.

[23] Herz, C. S. (1962). Fourier transforms related to convex sets. *Ann. of Math.* **75** 81–92.

[24] Heyde, C. C. (1967). On the influence of moments on the rate of convergence to the normal distribution. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 12–18.

[25] Ibragimov, I. A. (1966). On the accuracy of Gaussian approximation to the distribution functions of sums of independent variables. *Theor. Probability Appl.* **11** 559–579.

[26] Khinchin, A. (1929). Über einen neuen Grenzwertsatz der Wahrscheinlichkeitsrechnung. *Math. Ann.* **101** 745–752.

[27] Khinchin, A. (1949). *Mathematical Foundations of Statistical Mechanics.* English translation by G. Gamow. Dover, New York.

[28] Liapounov, A. M. (1901). Nouvelle forme du théorème sur la limite de probabilité. *Mem. Acad. Sci. St. Petersbuug* **12.**

[29] Linnik, Yu. V. and Mitrofanova, N. (1965). Some asymptotic expansions of maximum likelihood estimates. *Sankhyā Ser. A* **27** 73–82.

[30] Matthes, T. K. (1975). The multivariate central limit theorem for regular convex sets. *Ann. Probability* **3** 503–515.

[31] Paulauskas, V. (1969). One estimate of the remainder term in the multidimensional central limit theorem, I, II. *Litovsk. Mat. Sb.* **9** 329–342, 791–815 (in Russian).

[32] Petrov, V. V. (1964). On local limit theorems for sums of independent random variables. *Theor. Probability Appl.* **9** 312–320.

[33] Petrov, V. V. (1966). On a relation between an estimate of the remainder in the central limit theorem and the law of iterated logarithm. *Theor. Probability Appl.* **11** 454–458.

[34] Pfanzagl, J. (1973). Asymptotic expansions related to minimum contrast estimators. *Ann. Statist.* **1** 993–1026.

[35] Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

[36] Rao, Ranga R. (1960). Some problems in probability theory. Ph. D. thesis, Calcutta Univ.

[37] Rao, Ranga R. (1961). On the central limit theorem in R_k. *Bull. Amer. Math. Soc.* **67** 359–361.

[38] Sazanov, V. V. (1968). On the multi-dimensional central limit theorem. *Sankhyā Ser. A* **30** 181–204.

[39] Statulevicius, V. (1973). Asymptotic analysis of the distribution of sums of dependent random variables and the regularity conditions. *Internat. Conf. Prob. Theory Math. Statist.* Vilnius, U.S.S.R. Abstracts.

[40] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 583–602. Univ. of California Press.

[41] Sweeting, T. J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Probability* **5** 28–41.

[42] Wallace, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29** 635–654.

[43] Yarnold, J. K. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. *Ann. Math. Statist.* **43** 1566–1580.

[44] Yurunskii, V. V. (1972). Bounds for characteristic functions of certain degenerate multidimensional distributions. *Theor. Probability Appl.* **17** 101–113.

Department of Mathematics
University of Arizona
Tucson, Arizona 85721