The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1566–1580

ASYMPTOTIC APPROXIMATIONS FOR THE PROBABILITY THAT A SUM OF LATTICE RANDOM VECTORS LIES IN A CONVEX SET[1]

BY JAMES K. YARNOLD

McDonnell-Douglas Astronautics Company

0. Summary. Ranga Rao [10] developed a version of the Edgeworth asymptotic expansion for $\Pr(X_n \in B)$, where $X_n = n^{-\frac{1}{2}} \sum_{i=1}^n Z_i$, $[Z_n]$ is a sequence of independent random vectors in R_k having a common lattice distribution with mean vector zero and nonsingular covariance matrix Σ, and B is a Borel set. Use of this expansion is very difficult, except for the distribution function of X_n.

In this paper, Ranga Rao's expansion is used to obtain a different expansion, when B is convex. This new expansion is much simpler to evaluate. In the special case when $B = [x \mid x^T\Sigma^{-1}x < c]$, the new expansion assumes its simplest form.[2] The first partial sum is the usual multivariate normal approximation, and Esséen ([6] pages 110–111) determined the order of magnitude of its error, i.e.,

$$\Pr(X_n \in B) = K_k(c) + O(n^{-k/(k+1)})$$

where $K_k(c)$ is the chi-square distribution function with k degrees of freedom. Note that the order of magnitude of the error is $n^{-\frac{1}{2}}$ for $k = 1$ and approaches n^{-1} as k increases. The second partial sum is

$$\Pr(X_n \in B) = K_k(c) + (N(nc) - V(nc)) \frac{\exp(-c/2)}{(2\pi n)^{k/2}|\Sigma|^{\frac{1}{2}}} + O(n^{-1})$$

where $N(nc)$ is the number of integer vectors m in the ellipsoid $(m + na)^T\Sigma^{-1}(m + na) < nc$ having center at $-na$, and $V(nc)$ is the volume of this ellipsoid. This provides a new expansion for the distribution function of the quadratic form $X_n^T\Sigma^{-1}X_n$.

When Z_i has a multinomial distribution with parameters $N = 1, p_1, \cdots, p_m$, $\sum_{i=1}^m p_i = 1$, $X_n^T\Sigma^{-1}X_n$ is the chi-square goodness-of-fit statistic, and the new expansion (with $k = m - 1$) provides very accurate approximations for its distribution function. The accuracy of the first several partial sums, and of the Edgeworth approximation under the (inappropriate) assumption that Z_i has a continuous distribution, is examined numerically for a number of multinomial distributions. It is concluded that the Edgeworth approximation assuming a

---

Received October 9, 1969; revised January 12, 1972.

[1] This article is based upon a dissertation [11] for the Department of Statistics, University of Chicago, which was partially sponsored by the Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government. Additional sponsors were: General Analysis Corp.; CEIR; INC; URS Systems Corp.; International Telephone and Telegraph Corp.; McDonnell Douglas Corp.; U.S. Army Computer Command, CS3 and TACFIRE Projects; U.S. Army Test Command.

[2] In this article, x^T is the transpose of the vector x.

continuous distribution should never be used when Z_i has a lattice distribution, and that the second partial sum of the new expansion is much more accurate than the normal approximation for all multinomial distributions examined.

1. Introduction. Let $[Z_n]$ be a sequence of independent, identically distributed random vectors in R_k with mean vector zero and nonsingular covariance matrix Σ, and let $X_n = n^{-\frac{1}{2}}(Z_1 + \cdots + Z_n)$. In this paper we study approximations to $\Pr(X_n \in B)$ based on the normal approximation

$$\Pr(X_n \in B) \doteq \int_B d\Phi(x) \tag{1.1}$$

where $\Phi(x)$ is the normal distribution function having the same mean vector and covariance matrix as Z_i. The central limit theorem asserts that the error in the normal approximation goes to zero as $n \to \infty$ for all Borel sets B such that the probability of the boundary of B is zero under the normal approximation.

The normal approximation is not always satisfactory and several asymptotic expansions have been developed. Studies of the errors in the normal approximation and its improvements are important. The multidimensional case has been studied by Bergstrom [1], Bhattacharya [2], Esséen [6], and Ranga Rao [10].

When the normal approximation does not provide sufficient accuracy, the multivariate Edgeworth approximation is available:

$$\Pr(X_n \in B) \doteq \int_B d[\Phi(x) + \sum_{v=1}^{s-3} n^{-v/2} P_v(-\Phi(x))]\,. \tag{1.2}$$

The definition of $P_v(-\Phi(x))$ is given in Ranga Rao [10]. It is proven by the author [13] that $P_v(-\Phi(x))$ is equal to

$$\sum_{h=1}^{v} \frac{(-1)^{v+2h}}{h!} \sum_{\substack{v_1=1 \\ v_1+\cdots\cdots+v_h=v}}^{v} \cdots \sum_{v_h=1}^{v} \left[\prod_{r=1}^{h} \frac{(\lambda_1 D_1 + \cdots + \lambda_k D_k)^{v_r+2}}{(v_r+2)!}\right]\Phi(x)\,. \tag{1.3}$$

In the latter expression the differential operator D_i represents partial differentiation with respect to the ith component of the vector x. The quantity $(\lambda_1 D_1)^{s_1} \cdots (\lambda_k D_k)^{s_k}\Phi(x)$ is interpreted as a symbolic product which is equal to

$$\lambda_{s_1 s_2 \cdots s_k} \frac{\partial^{s_1+\cdots+s_k}}{(\partial x_1)^{s_1} \cdots (\partial x_k)^{s_k}} \Phi(x)\,,$$

where $\lambda_{s_1 s_2 \cdots s_k}$ is the multivariate cumulant of Z_i of orders $s_1, s_2, \cdots, s_k$. Let β_{js} be the sth absolute moment of the jth component of Z_i, and let $\beta_s = \sum_{j=1}^{k} \beta_{js}$. Ranga Rao [10] has shown that if the characteristic function $h(t)$ of Z_i satisfies Cramér's condition C

$$\lim_{|t|\to\infty} \sup |h(t)| < 1 \tag{1.4}$$

and if $\beta_{s+1} < \infty$, then the error in (1.2) is $O(n^{-(s-2)/2})$, uniformly in all convex sets B. This is an extension of Cramér's [4] result for the case $k = 1$. Condition C is satisfied if the distribution of Z_i has an absolutely continuous part.

When Z_i has a lattice distribution, Cramér's condition C is not satisfied and the preceding expansion is not valid. In this paper it is assumed that the possible values of Z_i are restricted to the set of lattice points $U = [a + m \mid m$ is an integer vector in R_k], where $a = (a_1, \cdots, a_k)$ is an arbitrary vector in R_k. Consequently X_n has a lattice distribution with possible values restricted to the lattice

$$(1.5) \qquad L = [(na + m)/n^{\frac{1}{2}} \mid m \text{ is an integer vector}] .$$

In this case, there is a local Edgeworth approximation for the probability at each point x in the lattice L:

$$(1.6) \qquad \Pr(X_n = x) \doteq n^{-k/2}[\phi(x) + \textstyle\sum_{v=1}^{s-3} n^{-v/2} P_v(-\phi(x))]$$

where $\phi(x)$ is the normal density function corresponding to $\Phi(x)$ and $P_v(-\phi(x))$ is defined by (1.3), upon replacing Φ by ϕ. Ranga Rao [10] has shown that if $\beta_s < \infty$, then the error in (1.6) is $O(n^{-(s-2+k)/2})$ uniformly in x.

Esséen [6] showed for $k = 1$ and Ranga Rao [10] generalized to any k that the local expansion (1.6) can be summed over any set B and expressed in the form

$$(1.7) \qquad \Pr(X_n \in B) \doteq \textstyle\int_B d\pi_n^{(s)}(x) .$$

If $\beta_{s+1} < \infty$, the approximation (1.7) is valid to $O(n^{-(s-2)/2})$ uniformly in Borel sets B. The definition of $\pi_n^{(s)}(x)$ is

$$(1.8) \qquad \pi_n^{(s)}(x) = [\textstyle\prod_{j=1}^{k} T_j^{(s)}]\psi_n^{(s)}(x) ,$$

where $\psi_n^{(s)}(x) = \sum_{v=0}^{s-3} n^{-v/2} P_v(-\Phi(x))$ and $T_j^{(s)}$ is the operator defined for $s \geqq 3$ as follows:

$$T_j^{(s)} = \textstyle\sum_{i=0}^{s-3} (-1)^i n^{-i/2} S_i(x_j n^{\frac{1}{2}} - na_j)(D_j)^i .$$

The function $S_j(t)$ is periodic in t with period one for $j = 0, 1, \cdots$, and is absolutely continuous for $j \geqq 2$. The definition of $S_j(t)$ is as follows:

$$S_0(t) = 1 ; \qquad S_1(t) = t - [t] - \tfrac{1}{2} , \qquad \text{where } [t] \text{ is the greatest integer} \leqq t ;$$

for $m = 1, 2, \cdots$

$$S_{2m}(t) = (-1)^{m-1} \sum_{n=1}^{\infty} \frac{2\cos(2n\pi t)}{(2n\pi)^{2m}} \qquad \text{and}$$

$$S_{2m+1}(t) = (-1)^{m-1} \sum_{n=1}^{\infty} \frac{2\sin(2n\pi t)}{(2n\pi)^{2m+1}} .$$

Further properties of these functions are given by Knopp ([9], page 522) in his discussion of the Euler—Maclaurin sum formula for the case $k = 1$.

2. Evaluation of a Lebesgue-Stieltjes integral. Ranga Rao's re-expression (1.7) of the lattice sum as a Stieltjes integral is of value only for sets with sufficient regularity. For rectangles aligned parallel to the coordinate axes, the result is a sum of 2^k terms requiring no integrations. For an arbitrary set B, on the other hand, the natural evaluation of the Stieltjes integral leads back to the sum of the local approximation over all lattice points in B.

In this paper, a useful evaluation is obtained for convex sets, or more generally for *extended convex* sets B whose sections parallel to each coordinate axis are all intervals. By carrying out the first integration over an interval, the probability is expressed as a sum of 2^{k-1} sum-integrals, and orders of magnitude of various terms can be studied.

An extended convex set B has representation (2.1) for every $r \in [1, \cdots, k]$:

(2.1) $$B = [x \mid w_r(x') < x_r < \theta_r(x') \text{ and } x' = (x_1, \cdots, x_{r-1}, x_{r+1}, \cdots, x_k) \in B_r]$$

where $B_r \subset R_{k-1}$ and w_r, θ_r are continuous functions on R_{k-1} into R_1 and $x_1, \cdots, x_k$ are the components of x.

Let F be a function on R_k into R_1 which does not depend upon n and let $T_j^{(s)}$ be the operator defined in (1.8). In Theorem 1, the integral $\int_B d(\prod_{j=1}^k T_j^{(s)})F(x)$ is evaluated for an extended convex set B. This is accomplished by integrating first with respect to x_r (for various r) over the interval $(w_r(x'), \theta_r(x'))$ for fixed $x' \in B_r$.

In Section 3, Theorem 1 will be used to obtain a new asymptotic expansion for $\Pr(X_n \in B)$ when B is an extended convex set. This is accomplished by applying Theorem 1 to each F_i in $\psi_n^{(s)}(x) = \sum_i c_i(n)F_i(x)$ in order to evaluate $\int_B d[\prod_{j=1}^k T_j^{(s)}]\psi_n^{(s)}(x)$ (see (1.7) and (1.8)).

The following notation is used in the rest of the paper: $\chi_B(x)$ is the characteristic function of the set B, and $\chi_{B_r}(x')$ is the characteristic function of the set B_r.

$$(h(x))_{w_r(x')}^{\theta_r(x')} = h(x_1, \cdots, x_{r-1}, \theta_r(x'), x_{r+1}, \cdots, x_k) - h(x_1, \cdots, x_{r-1}, w_r(x'), x_{r+1}, \cdots, x_k) .$$

$L_j = [(na_j + m)/n^{\frac{1}{2}} \mid m$ is an integer$]$ is a lattice in R_1. The Cartesian product $L_1 \times \cdots \times L_k$ is the lattice L in R_k defined in (1.5).

$$\Delta_j F(x) = F(x_1, \cdots, x_{j-1}, x_j + 0, x_{j+1}, \cdots, x_k) - F(x_1, \cdots, x_{j-1}, x_j - 0, x_{j+1}, \cdots, x_k) .$$

W is the set of all subsets of $\Lambda = [1, \cdots, k]$.

For $I \in W$, $C(I) = [x \mid x_j \in L_j$ if $j \in I$; $x_j \in R_1$ if $j \in I^c]$, where I^c is the complement of I in Λ.

E is the set of all subsets of W which do not contain r.

For $I \in E$, $C_r(I) = [x' \mid x_j \in L_j$ if $j \in I$; $x_j \in R_1$ if $j \in I^*]$, where I^* is the complement of I in $[1, \cdots, r-1, r+1, \cdots, k]$.

We define abbreviated notation for sum-integrals of special types over R_k and R_{k-1}.

If $I = [i_1, \cdots, i_m]$, then

$$\sum_{C(I) \cap B}\int g(x) = \sum_{x_{i_1} \in L_{i_1}} \cdots \sum_{x_{i_m} \in L_{i_m}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \chi_B(x) g(x) \prod_{j \in I^c} dx_j$$

and provided $I \in E$,

$$\sum_{C_r(I) \cap B_r}\int h(x') = \sum_{x_{i_1} \in L_{i_1}} \cdots \sum_{x_{i_m} \in L_{i_m}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \chi_{B_r}(x') h(x') \prod_{j \in I^*} dx_j .$$

The set $I_0 = [r+1, \cdots, k]$ occurs frequently, and we define $C_r^* = C_r(I_0)$.

THEOREM 1. *Let F be a function on R_k into R_1 which has partial derivatives of* *all orders in* R_k *and which does not depend upon* n. *Let* $f(x) = (\prod_{i=1}^k D_i)F(x)$. *If* B *is a bounded extended convex set, then*

$$\int_B d(\prod_{j=1}^k T_j^{(s)})F(x) \tag{2.2}$$
$$= \int_B dF(x) + \sum_{i=1}^{s-3}(-1)^i n^{-i/2}[\sum_{r=1}^k n^{-(k-r)/2} \sum_{C_r^* \cap B_r} \int (S_i(n^{\frac{1}{2}}x_r - na_r)D_r^{i-1}f(x))_{w_r(x')}^{\theta_r(x')}] + O(n^{-(s-2)/2}) .$$

Furthermore, for each i, *the quantity within brackets is* $O(1)$.

The proof of Theorem 1 is given in the Appendix.

3. Asymptotic expansion for $\Pr(X_n \in B)$, **when** B **is an extended convex set and** Z_i **has a lattice distribution.** The following theorem provides a new asymptotic expansion for $\Pr(X_n \in B)$.

THEOREM 2. *If* B *is a bounded extended convex set, if* Z_i *has possible values in the lattice* U, *and if* $\beta_{s+1} < \infty$, *then*

$$\Pr(X_n \in B) = \int_B \phi(x)\, dx + \sum_{j=1}^{s-3} n^{-j/2} \int_B P_j(-\phi(x))\, dx + \sum_{j=1}^{s-3} n^{-j/2} \sum_{i=1}^{j} (-1)^i \sum_{r=1}^k n^{-(k-r)/2} \sum_{C_r^* \cap B_r} \int (S_i(n^{\frac{1}{2}}x_r - na_r)D_r^{i-1}P_{j-i}(-\phi(x)))_{w_r(x')}^{\theta_r(x')} + O(n^{-(s-2)/2}) .$$

Furthermore, the jth *term is* $O(n^{-j/2})$.

PROOF. It is known from (1.7) that if $\beta_{s+1} < \infty$, then $\Pr(X_n \in B) = \int_B d[\prod_{j=1}^k T_j^{(s)}]\phi_n^{(s)}(x) + O(n^{-(s-2)/2})$. By definition, $\phi_n^{(s)}(x) = \sum_{v=0}^{s-3} n^{-v/2} \times P_v(-\Phi(x))$ so that $\int_B d[\prod_{j=1}^k T_j^{(s)}]\phi_n^{(s)}(x) = \sum_{v=0}^{s-3} n^{-v/2} \int_B d[\prod_{j=1}^k T_j^{(s)}]P_v(-\Phi(x))$. Because $P_v(-\Phi(x))$ is a function on R_k into R_1 which has partial derivatives of all orders and which is independent of n, Theorem 1 may be applied to each integral in the latter sum, resulting in:

$$\Pr(X_n \in B) = \sum_{j=0}^{s-3} n^{-j/2} \int_B P_j(-\phi(x))\, dx + \sum_{v=0}^{s-3} \sum_{i=1}^{s-3} (-1)^i n^{-(i+v)/2} \sum_{r=1}^k n^{-(k-r)/2} \times \sum_{C_r^* \cap B_r} \int (S_i(y_r)D_r^{i-1}P_v(-\phi(x)))_{w_r(x')}^{\theta_r(x')} + O(n^{-(s-2)/2}) \tag{3.1}$$

where $y_r = n^{\frac{1}{2}}x_r - na_r$ and we have used $(\prod_{i=1}^k D_i)P_v(-\Phi(x)) = P_v(-\phi(x))$.

We next consider the second term (call it A) on the right-hand side of (3.1). Note that the i, v term in A is $O(n^{-(i+v)/2})$, using the second conclusion of Theorem 1. Upon putting all terms in A such that $i + v \geqq s - 2$ into the remainder and substituting $j = i + v$, it follows that

$$A = \sum_{j=1}^{s-3} n^{-j/2} \sum_{i=1}^{j} (-1)^i \sum_{r=1}^k n^{-(k-r)/2} \sum_{C_r^* \cap B_r} \int (S_i(y_r)D_r^{i-1}P_{j-i}(-\phi(x)))_{w_r(x')}^{\theta_r(x')} + O(n^{-(s-2)/2}) . \tag{3.2}$$

The proof of Theorem 2 is completed by substituting (3.2) into (3.1).

The simplest version of this new expansion, beyond the normal approximation term $\int_B \phi(x)\, dx$, is given in Theorem 3.

THEOREM 3. *If B is a bounded extended convex set, if Z_i has possible values in the lattice U, and if $\beta_5 < \infty$, then*

$$\Pr(X_n \in B) = \int_B \phi(x)\,dx - n^{-\frac{1}{2}} \sum_{\substack{n_1=0 \\ n_1+\cdots+n_k=3}}^{3} \cdots \sum_{n_k=0}^{3} \frac{\lambda_{n_i \cdots n_k}}{(n_1)! \cdots (n_k)!} \int_B \left(\prod_{i=1}^k D_i^{n_i}\right)\phi(x)\,dx$$

$$- \sum_{r=1}^k n^{-(k-r+1)/2} \sum_{C_r^* \cap B_r} \int (S_1(n^{\frac{1}{2}}x_r - na_r)\phi(x))_{w_r(x')}^{\theta_r(x')} + O(n^{-1}) .$$

Furthermore, the second and third terms on the right-hand side of the equality are $O(n^{-\frac{1}{2}})$.

4. Asymptotic expansion for the distribution function of the quadratic form $Q_n = X_n^T\Sigma^{-1}X_n$ when Z_i has a lattice distribution. When B is an ellipsoid of the form $B = [x \mid x^T\Sigma^{-1}x < c]$, it is known [5] that the normal approximation for $\Pr(X_n \in B)$ is $\int_B \phi(x)\,dx = K_k(c)$, the χ^2 distribution function with k degrees of freedom. Esséen ([6] pages 110–111) has shown that if Q_n is the quadratic form $X_n^T\Sigma^{-1}X_n$ and if $\beta_4 < \infty$, then

$$\Pr(Q_n < c) = K_k(c) + O(n^{-k/(k+1)}) . \tag{4.1}$$

Note that the order of magnitude of the error committed by the χ^2 approximation is $n^{-\frac{1}{2}}$ for $k = 1$ and approaches n^{-1} as k increases.

The simplest version of the new asymptotic expansion for $\Pr(Q_n < c)$—beyond the χ^2 approximation term—is the following second partial sum:

THEOREM 4. *If Z_i has possible values in the lattice U, $\beta_5 < \infty$, and $c < \infty$, then*

$$\Pr(Q_n < c) = K_k(c) + [N(nc) - V(nc)] \frac{\exp(-c/2)}{(2\pi n)^{\frac{1}{2}k}|\Sigma|^{\frac{1}{2}}} + O(n^{-1}) \tag{4.2}$$

and

$$[N(nc) - V(nc)] \frac{\exp(-c/2)}{(2\pi n)^{\frac{1}{2}k}|\Sigma|^{\frac{1}{2}}} = O(n^{-k/(k+1)}) \tag{4.3}$$

where $N(nc)$ is the number of integer vectors m in the ellipsoid $(m + na)^T\Sigma^{-1}(m + na) < nc$ having center at $-na$, and $V(nc) = (\pi nc)^{k/2}|\Sigma|^{\frac{1}{2}}/\Gamma(\frac{1}{2}k + 1)$ is the volume of this ellipsoid.

PROOF. The first step is to show that the second term on the right-hand side of the equality in Theorem 3 vanishes when $B = [x \mid x^T\Sigma^{-1}x < c]$. For every $(n_1, \cdots, n_k)$, $(\prod_{i=1}^k D_i^{n_i})\phi(x)$ is an odd function, and the integral of an odd function over the ellipsoid B vanishes.

We next evaluate the third term on the right-hand side of the equality in Theorem 3 (call it J). When $B = [x \mid x^T\Sigma^{-1}x < c]$, $\theta_r(x')$ and $w_r(x')$ are the values of x_r such that $x^T\Sigma^{-1}x = c$. Thus

$$\phi(x) = \frac{\exp(-c/2)}{(2\pi)^{k/2}|\Sigma|^{\frac{1}{2}}} = d$$

when $x_r = \theta_r(x')$ or $x_r = w_r(x')$, and

$$J = -d \sum_{r=1}^k n^{-(k-r+1)/2} \sum_{x_{r+1} \in L_{r+1}} \cdots \sum_{x_k \in L_k} \int_{B_r} \cdots \int [S_1(y_r)]_{w_r(x')}^{\theta_r(x')} \prod_{i=1}^{r-1} dx_i ,$$

where $y_r = n^{\frac{1}{2}}x_r - na_r$. We apply (8.10) and the proof which follows it to $G_r(x_r) = S_1(y_r)$, which is differentiable except on a discrete set L_r; to $\mathscr{F}_{x'}(x_r) = 1$, which is differentiable on the real line; and to $\bar{\alpha}(x_r) = G_r\mathscr{F}_{x'}(x_r) = S_1(y_r)$ to prove that

$$J = -d \sum_{r=1}^{k} n^{-(k-r+1)/2} \sum_{x_{r+1}\in L_{r+1}} \cdots \sum_{x_k \in L_k} \int \cdots \int_{B_r} \Big(\int_{w_r(x')}^{\theta_r(x')} D_r S_1(y_r)\, dx_r + \sum_{w_r(x');x_r\in L_r}^{\theta_r(x')} \Delta_r S_1(y_r)\Big) \prod_{i=1}^{r-1} dx_i .$$

By definition $S_1(t) = t - [t] - \frac{1}{2}$, so that $D_r S_1(y_r) = n^{\frac{1}{2}}$, $\Delta_r S_1(y_r) = -1$, and

$$\begin{aligned} J &= -d \sum_{r=1}^{k} \Big[n^{-(k-r)/2} \sum_{x_{r+1}\in L_{r+1}} \cdots \sum_{x_k\in L_k} \int \cdots \int_B \prod_{i=1}^{r} dx_i \\ &\quad - n^{-(k-r+1)/2} \sum_{x_r\in L_r} \cdots \sum_{x_k\in L_k} \int \cdots \int_B \prod_{i=1}^{r-1} dx_i\Big] \\ &= d\Big(n^{-k/2} \sum_{x_1\in L_1} \cdots_B \sum_{x_k\in L_k} 1 - \int \cdots \int_B \prod_{i=1}^{k} dx_i\Big) . \end{aligned}$$

$\sum_{x_1\in L_1} \cdots_B \sum_{x_k\in L_k} 1$ is the number of points of the lattice L which are in $B = [x \mid x^T\Sigma^{-1}x < c]$. Because $x \in L$ if and only if $x = (m + na)/n^{\frac{1}{2}}$ and m is an integer vector, the latter number is equal to $N(nc)$, the number of integer vectors m in the ellipsoid $(m + na)^T\Sigma^{-1}(m + na) < nc$ having center at $-na$. It is known ([5] page 120) that

$$n^{k/2} \int \cdots \int_B \prod_{i=1}^{k} dx_i = \frac{(\pi nc)^{k/2}|\Sigma|^{\frac{1}{2}}}{\Gamma(\frac{1}{2}k + 1)} \equiv V(nc)$$

is the volume of the latter ellipsoid. Thus $J = dn^{-k/2}(N(nc) - V(nc))$, which is the desired second term in (4.2).

The last step in the proof is the verification of (4.3), i.e. $J = O(n^{-k/(k+1)})$. From Esséen ([6] page 117) it is known that for $k \geqq 2$, $V(nc) - N(nc) = O(n^{k/2-k/(k+1)})$. Thus if $k \geqq 2$, then $J = dn^{-k/2}O(n^{k/2-k/(k+1)}) = O(n^{-k/(k+1)})$. In the case $k = 1$, Esséen ([6] Chapter IV) has shown that $J = O(n^{-\frac{1}{2}})$.

The third partial sum of this asymptotic expansion for $\Pr(Q_n < c)$ is given in Theorem 5. In the rest of this article, Kaplan's notation for multivariate cumulants is used, e.g. $\lambda(i_1, i_2, i_3)$ is the general expression for the third order cumulant of components i_1, i_2, i_3 of Z_i. If $i_1 \neq i_2 \neq i_3$, $\lambda(i_1, i_2, i_3) = \lambda_{111}$ for components i_1, i_2, i_3. If $i_1 = i_2 = i_3$, $\lambda(i_1, i_2, i_3) = \lambda_3$ for component i_1. If $i_1 = i_2 \neq i_3$, $\lambda(i_1, i_2, i_3) = \lambda_{21}$ for components i_1, i_3. σ^{ij} is an element of Σ^{-1}.

THEOREM 5. *If Z_i has possible values in the lattice U, $c < \infty$, and $\beta_6 < \infty$, then*

$$\begin{aligned} &\Pr(Q_n < c) \\ &\quad = K_k(c) + (N(nc) - V(nc))\frac{\exp(-c/2)}{(2\pi n)^{k/2}|\Sigma|^{\frac{1}{2}}} \\ &\qquad - \Big[\frac{\delta_1}{n}\Big(\sum_{t=0}^{2}(-1)^{2-t}\tbinom{2}{t}K_{k+2t}(c)\Big) + \frac{\delta_2}{n}\Big(\sum_{t=0}^{3}(-1)^{3-t}\tbinom{3}{t}K_{k+2t}(c)\Big)\Big] \\ &\qquad - \frac{1}{n}\Big[\sum_{r=1}^{k} n^{-(k-r)/2} \sum_{C_r^*\cap B_r}\int (S_1(n^{\frac{1}{2}}x_r - na_r)P_1(-\phi(x)))_{w_r(x')}^{\theta_r(x')}\Big] \\ &\qquad + \frac{1}{n}\Big[\sum_{r=1}^{k} n^{-(k-r)/2} \sum_{C_r^*\cap B_r}\int (S_2(n^{\frac{1}{2}}x_r - na_r)D_r\phi(x))_{w_r(x')}^{\theta_r(x')}\Big] + O(n^{-\frac{3}{2}}) , \end{aligned} \tag{4.4}$$

where

$$\delta_1 = \tfrac{1}{8} \sum_{i_1=1}^k \sum_{i_2=1}^k \sum_{i_3=1}^k \sum_{i_4=1}^k \sigma^{i_1 i_2}\sigma^{i_3 i_4}\lambda(i_1, i_2, i_3, i_4)$$

$$\delta_2 = \sum_{i_1=1}^k \cdots \sum_{i_6=1}^k \left[\frac{\sigma^{i_1 i_2}\sigma^{i_3 i_4}\sigma^{i_5 i_6}}{8} + \frac{\sigma^{i_1 i_4}\sigma^{i_2 i_5}\sigma^{i_3 i_6}}{12}\right] \lambda(i_1, i_2, i_3)\lambda(i_4, i_5, i_6)$$

and B is the ellipsoid $[x \mid x^T \Sigma^{-1} x < c]$.

Theorem 5 is the case $s = 5$ of Theorem 2. The only point which requires verification is that the first bracketed expression is equal to $-n^{-1} \int_B P_2(-\phi(x))\, dx$. This is done in [13], where it is also shown that:

$$\frac{\delta_1}{n} = \tfrac{1}{8}(\sigma^2(Q_n) - 2k) \tag{4.5}$$

$$\frac{\delta_2}{n} = -24\left(\frac{\sigma^2(Q_n)}{4} - \frac{\bar{\mu}_3(Q_n)}{48} - \frac{k}{3}\right)$$

where $\sigma^2(Q_n)$ is the variance of Q_n and $\bar{\mu}_3(Q_n)$ is the third central moment of Q_n (to order n^{-1}). The second method of evaluating δ_1 and δ_2, as defined by (4.5), is convenient when the second and third moments of Q_n are known. The first method of evaluating δ_1 and δ_2, as defined in (4.4), requires the cumulants of Z_i and σ^{ij}.

The last two terms in (4.4) are at most of order n^{-1}. Whether they are of higher order is not known.

5. Asymptotic expansion for the distribution function of the chi-square goodness-of-fit statistic. Given a multinomial distribution with parameters $N = 1, p_1, \cdots, p_m$, where $\sum_{j=1}^m p_j = 1$, let Z_i be the $k = m - 1$ dimensional random vector obtained by deleting the mth component of the multinomial random vector and centering at the mean vector, $(p_1, \cdots, p_{m-1})$. Given a sample of size n from this multinomial, it is desired to test the null hypothesis that $p_1, \cdots, p_m$ have specified values. The classical chi-square goodness-of-fit statistic, X^2, is a quadratic form of the type $Q_n = X_n^T \Sigma^{-1} X_n$ to which the asymptotic expansions in Theorems 2, 4, and 5 apply, with $|\Sigma| = p_1 \cdots p_m$ and $a = -(p_1, \cdots, p_{m-1})$.

Hoel [8] applied the multivariate Edgeworth approximation for a continuous distribution in (1.2) to $\Pr(X^2 < c)$, with $s = 5$, obtaining:

$$\Pr(X^2 < c) \doteq K_{m-1}(c) - \frac{\delta_1}{n}\left(\textstyle\sum_{t=0}^2 (-1)^{2-t}\binom{2}{t} K_{m-1+2t}(c)\right) - \frac{\delta_2}{n}\left(\textstyle\sum_{t=0}^3 (-1)^{3-t}\binom{3}{t} K_{m-1+2t}(c)\right). \tag{5.1}$$

This is not a valid expansion, as Hoel stated, because Z_i has a lattice distribution. The error in (5.1) is $O(n^{-(m-1)/m})$, which is the same as for the χ^2 approximation. It is clear from the valid expansion given in (4.4) and from (4.3) that the *continuous Edgeworth* expansion (5.1) neglects the discontinuous term which is $O(n^{-(m-1)/m})$.

However, δ_1 and δ_2 are required to use the third partial sum in (4.4), and Hoel's evaluation of δ_1 and δ_2 for the multinomial (using a different and much more complex method) provides a check upon the new first and second methods of evaluating δ_1 and δ_2 given in (4.4) and (4.5). The second method (4.5) requires only the second and third central moments of X^2, which were obtained by Haldane [7]:

$$\sigma^2(X^2) = 2(m-1) + \frac{1}{n}\left(\sum_{i=1}^m \frac{1}{p_i} - m^2 - 2m + 2\right)$$

$$\mu_3(X^2) = 8(m-1) - \frac{1}{n}\left(18m^2 + 36m - 32 - 22\sum_{i=1}^m \frac{1}{p_i}\right) + O(n^{-2}) .$$

Thus, using (4.5)

$$\frac{\delta_1}{n} = \tfrac{1}{8}(\sigma^2(X^2) - 2(m-1)) = \frac{1}{8n}\left(\sum_{i=1}^m \frac{1}{p_i} - m^2 - 2m + 2\right)$$

$$\frac{\delta_2}{n} = -24\left(\frac{\sigma^2(X^2)}{4} - \frac{\mu_3(X^2)}{48} - \frac{m-1}{3}\right) = \frac{1}{n}\left(5\sum_{i=1}^m \frac{1}{p_i} - 3m^2 - 6m + 4\right).$$

6. Numerical investigation of the accuracy of approximations for $\Pr(X^2 < c)$. In this section the accuracy of the following three approximations for $\Pr(X^2 < c)$ is compared with that of the χ^2 approximation for a number of different multinomial distributions:

Approximation A is the second partial sum given in (4.2).

Continuous Edgeworth is the multivariate Edgeworth approximation for a continuous distribution given in (5.1).

Approximation B is the third partial sum in (4.4), without the last two bracketed terms.

For a multinomial distribution with expectations $np_1, np_2, \cdots, np_m$, Table 1 gives for each approximation the maximum absolute error in distribution functions (true-approximate), at possible values of X^2 (with the sign of the largest error attached).

A more extensive investigation of the accuracy of these approximations, including the two-parameter Γ distribution, the χ^2 approximation with continuity correction [3], and the $C(\underline{m})$ approximation is given in [12], using maximum error over the whole range and over the upper ten percent.

7. Conclusions on the accuracy of approximations for $\Pr(X^2 < c)$. The following conclusions on the accuracy of these approximations are based upon the order of magnitude of their remainders and inspection of Table 1:

A. Never use the continuous Edgeworth approximation. The χ^2 approximation is almost as accurate or better.

B. Approximation A is much more accurate than the χ^2 approximation for every multinomial distribution examined.

C. Approximations A and B are very much more accurate than the χ^2 approximation when the χ^2 approximation is only fair (or better). The accuracy of A and B increases as $Q = \sum_{i=1}^{m} 1/(np_i)$ decreases, and B is more accurate than A when Q is sufficiently small.

D. Approximations A, B, and χ^2 are all very inaccurate when there are too many small expectations. Under the assumption that, as $n \to \infty$, some expectations remain finite while the rest increase without limit, the limiting distribution of X^2 is the $C(\underline{m})$ distribution. This $C(\underline{m})$ approximation is accurate when there are any number of small expectations. The basic idea for the $C(\underline{m})$ approximation is due to Cochran [3], and a proof for the general case is given in [12].

E. It is recommended that the two-parameter Γ approximation and the X^2 approximation with continuity correction not be used (see [12]).

8. Appendix.

PROOF OF THEOREM 1. Let

$$T_j = T_j^{(s)} \tag{8.1}$$

$$T_j^* = T_j^{(s)} - 1 \,. \tag{8.2}$$

Then $\prod_{j=1}^k T_j = 1 + \sum_{r=1}^k T_r^* \prod_{j=r+1}^k T_j$ and

$$\int_B d(\textstyle\prod_{j=1}^k T_j)F(x) = \int_B dF(x) + \sum_{r=1}^k \int_B d(T_r^* \prod_{j=r+1}^k T_j)F(x) \,. \tag{8.3}$$

We now evaluate the last integral in (8.3). The quantity $(T_r^* \prod_{j=r+1}^k T_j)F$ is a sum of terms of the form

$$(\textstyle\prod_{j=r}^k g_j(x_j) \prod_{j=r}^k D_j^{m_j})F \,,$$

where $g_j(x_j) = S_{m_j}(n^{\frac{1}{2}}x_j - na_j)(-1)^{m_j}n^{-m_j/2}$ is differentiable except when $x_j \in L_j$, and each term may be rewritten as

$$(\textstyle\prod_{j=r}^k g_j(x_j)) \cdot (\prod_{j=r}^k D_j^{m_j}F) \,.$$

Lemma 1, which appears at the end of this section, gives a reduction to sums and Lebesgue integrals for a Stieltjes integral of the type $\int_B d(\prod_{j=1}^k G_j)\mathscr{F}(x)$. It applies immediately to the integrals required above with $\mathscr{F} = \prod_{j=r}^k D_j^{m_j}F$, $G_j = g_j$ for $j \geqq r$, and $G_j = 1$ for $j < r$. The result of recombining all the terms is equivalent to applying Lemma 1 directly to

$$\int_B d(T_r^* \textstyle\prod_{j=r+1}^k T_j)F(x)$$

$$\text{with} \quad G_j = \begin{bmatrix} 1 & \text{for } j = 1, \cdots, r-1 \\ T_r^* & \text{for } j = r \\ T_j & \text{for } j = r+1, \cdots, k \end{bmatrix} \qquad \text{and} \qquad \mathscr{F} = F \,,$$

even though Lemma 1 is not directly applicable to operators T_j; we do not give the details of this step. The latter direct application of Lemma 1 gives the following result:

$$\int_B d(T_r^* \textstyle\prod_{j=r+1}^k T_j)F(x) = \sum_{I \in E} Q(I) \tag{8.4}$$

where

$$Q(I) = \sum_{C_r(I) \cap B_r} \int [(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)F(x)]_{w_r(x')}^{\theta_r(x')} .$$

It follows that

$$(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)F(x) = (\prod_{\substack{I \\ j \neq j_0}} \Delta_j G_j)\Delta_{j_0} G_{j_0}(G_r)(\prod_{I^*} D_j G_j)F(x)$$

because G_j is a function only of x_j and because the differential and difference operators can be applied in any order. If $\exists j \in I \ni j < r$, say $j = j_0$, then $G_{j_0} = 1$ so $G_{j_0}(G_r)(\prod_{I^*} D_j G_j)F(x)$ is differentiable with respect to x_{j_0} and $\Delta_{j_0} G_{j_0}(G_r) \times (\prod_{I^*} D_j G_j)F(x) = 0$. Thus the E on the right-hand side of (8.4) may be replaced by E_0, the set of all subsets of $[r+1, \cdots, k]$.

If $I \in E_0$, we define $I^{**} = I^* \cap [r+1, \cdots, k]$ and the definition of the G_j's implies

$$Q(I) = \sum_{C_r(I) \cap B_r} \int [(\prod_I \Delta_j T_j)(T_r^*)(\prod_{j=1}^{r-1} D_j \prod_{I^{**}} D_j T_j)F(x)]_{w_r(x')}^{\theta_r(x')} .$$

Using the expressions for $\Delta_j T_j$ and $D_j T_j$ given in Lemmas 2 and 3 (which appear at the end of this section), it follows that if $I \in E_0$ has m elements $[i_1, \cdots, i_m]$, then

$$(8.5) \qquad Q(I) = n^{-m/2} \sum_{C_r(I) \cap B_r} \int [(\sum_{i=1}^{s-3} (-1)^i n^{-i/2} S_i(y_r) D_r^{i-1})(-n^{-\frac{1}{2}})^{(s-3)(k-r-m)} \times (\prod_{I^{**}} S_{s-3}(y_j) D_j^{s-3}) f(x)]_{w_r(x')}^{\theta_r(x')}$$

where $y_j = n^{\frac{1}{2}} x_j - na_j$.

If $m < k - r$, then the power of $n^{-\frac{1}{2}}$ in the integrand of $Q(I)$ is at least $1 + (s-3) = s - 2$. Also $S_i(y_j)$ is bounded by a constant independent of n, and the same is true of every partial derivative of $f(x)$. It follows from the latter and the fact that the region of integration in $Q(I)$ is independent of n that the integral in $Q(I)$ is $O(n^{-(s-2)/2})$. To estimate the multiple sum over $(x_{i_1}, \cdots, x_{i_m}) \in L_{i_1} \times \cdots \times L_{i_m}$ in $Q(I)$, note that from the definition of the lattice L it is known the spacing between values of $x_j \in L_j$ is $n^{-\frac{1}{2}}$. If d is the diameter of B_r, an immediate estimate of the multiple sum in $Q(I)$ is $O((dn^{\frac{1}{2}})^m)$. Thus

$$Q(I) = n^{-m/2} \cdot O((dn^{\frac{1}{2}})^m) \cdot O(n^{-(s-2)/2}) = O(n^{-(s-2)/2})$$

if $m < k - r$. There is only one $I \in E_0$ such that $m = k - r$, namely $I_0 = [r+1, \cdots, k]$. Consequently for $r = 1, \cdots, k$

$$(8.6) \qquad \int_B d(T_r^* \prod_{j=r+1}^{k} T_j)F(x) = Q(I_0) + O(n^{-(s-2)/2}) .$$

Because $I_0^* = [1, \cdots, r-1]$ and I_0^{**} is the null set, it follows from (8.5) that

$$(8.7) \qquad Q(I_0) = \sum_{i=1}^{s-3} (-1)^i n^{-i/2} [n^{-(k-r)/2} \sum_{C_r^* \cap B_r} \int [S_i(n^{\frac{1}{2}} x_r - na_r) D_r^{i-1} f(x)]_{w_r(x')}^{\theta_r(x')}] .$$

For each i and r, the expression in outer brackets in (8.7) is $O(1)$, by the argument just used to determine the order of magnitude of $Q(I)$.

The conclusions of Theorem 1 follow from (8.3), (8.6), and (8.7).

LEMMA 1. *Let G_j be a function of x_j on R_1 into R_1 which is differentiable except on a discrete set*[3] *L_j. Let $\mathscr{F}$ and $\alpha = (\prod_{j=1}^k G_j)\mathscr{F}$ be functions on R_k into R_1, where $\mathscr{F}$ is differentiable everywhere and α is of bounded variation on compact sets. If B has the representation* (2.1), *then*

$$\int_B d\alpha(x) = \sum_{I\in E}\Big(\sum_{C_r(I)\cap B_r}\int_I [(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)]_{w_r(x')}^{\theta_r(x')}\Big)\,. \tag{8.8}$$

PROOF. From the definition of the Lebesgue-Stieltjes integral, it follows that

$$\int_B d\alpha(x) = \sum_{I\in W}\Big[\sum_{C(I)\cap B}\int_I (\prod_I \Delta_j G_j)(\prod_{I^c} D_j G_j)\mathscr{F}(x)\Big]\,, \tag{8.9}$$

where

$$(\prod_I \Delta_j G_j)(\prod_{I^c} D_j G_j)\mathscr{F}(x) = \prod_I \Delta_j \prod_{I^c} D_j \prod_{j=1}^k G_j\mathscr{F}(x)\,.$$

Note that the range of every differential or difference operator is understood to be everything to its right.

For every $I \in E$ we define $I_r = I \cup [r]$, which provides a one-to-one correspondence of E onto E^c: $I \Leftrightarrow I_r$. (Note that $(I_r)^c = I^*$.) Consequently

$$\int_B d\alpha(x) = \sum_{I\in E}\Big[\sum_{C(I)\cap B}\int_I (\prod_I \Delta_j G_j)(\prod_{I^c} D_j G_j)\mathscr{F}(x) + \sum_{C(I_r)\cap B}\int_{I_r} (\prod_I \Delta_j G_j)(\prod_{I^*} D_j G_j)\mathscr{F}(x)\Big]\,.$$

By Fubini's theorem,

$$\sum_{C(I)\cap B}\int_I (\prod_I \Delta_j G_j)(\prod_{I^c} D_j G_j)\mathscr{F}(x) = \sum_{C_r(I)\cap B_r}\int_I \int_{w_r(x')}^{\theta_r(x')} (\prod_I \Delta_j G_j)(D_r G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)dx_r\,.$$

$$\sum_{C(I_r)\cap B}\int_{I_r} (\prod_I \Delta_j G_j)(\prod_{I^*} D_j G_j)\mathscr{F}(x) = \sum_{C_r(I)\cap B_r}\int_I \sum_{\substack{w_r(x')\\ x_r\in L_r}}^{\theta_r(x')} (\prod_I \Delta_j G_j)(\Delta_r G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)\,.$$

$$\begin{aligned} &\int_{w_r(x')}^{\theta_r(x')} (\prod_I \Delta_j G_j)(D_r G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)\,dx_r \\ &\quad + \sum_{\substack{w_r(x')\\ x_r\in L_r}}^{\theta_r(x')} (\prod_I \Delta_j G_j)(\Delta_r G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x) \\ &\qquad = \int_{w_r(x')}^{\theta_r(x')} d_{x_r}[(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)]\mathscr{F}(x) \\ &\qquad = [(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)]_{w_r(x')}^{\theta_r(x')}\,. \end{aligned} \tag{8.10}$$

The proof of the first equality in (8.10) is as follows: The hypotheses on G_j and $\mathscr{F}(x)$ imply the x' section $\bar{\mathscr{F}}_{x'}(x_r)$ of $\bar{\mathscr{F}}(x) = (\prod_I \Delta_j G_j)(\prod_{I^*} D_j G_j)\mathscr{F}(x)$ is differentiable with respect to x_r on the real line. G_r, $\bar{\mathscr{F}}_{x'}(x_r)$, $\bar{\alpha}(x_r) = G_r\bar{\mathscr{F}}_{x'}(x_r)$ satisfy the hypotheses of Lemma 1 in the one-dimensional case with $x = x_r$, $B = (w_r(x'), \theta_r(x'))$, and the first equality in (8.10) is given by (8.9) (upon replacing $\alpha(x)$ by $\bar{\alpha}(x_r)$, $\mathscr{F}(x)$ by $\bar{\mathscr{F}}_{x'}(x_r)$, and Λ by $\bar{\Lambda} = [r]$.)
Consequently

$$\int_B d\alpha(x) = \sum_{I\in E}\Big[\sum_{C_r(I)\cap B_r}\int_I [(\prod_I \Delta_j G_j)(G_r)(\prod_{I^*} D_j G_j)\mathscr{F}(x)]_{w_r(x')}^{\theta_r(x')}\Big]\,,$$

which completes the proof of Lemma 1.

[3] In Lemma 1 L_j is defined more generally as a discrete set, instead of a lattice. The definitions of $C(I)$ and $C_r(I)$ are the same with this more general definition of L_j.

TABLE 1
Maximum error of approximations for Pr$(X^2 < c)$

Expectations	χ^2	A	B	Continuous Edgeworth
2,2,2	−.2701	−.0504	−.0344	−.2465
3,3,3	−.1982	−.0315	−.0201	−.1841
5,5,5	−.1348	−.0173	−.0142	−.1254
10,10,10	−.0977	−.0102	−.0057	−.0932
20,20,20	−.0659	−.0044	−.0021	−.0638
50,50,50	−.0315	−.0016	−.0009	−.0310
100,100,100	−.0201	−.0012	−.0009	−.0198
.1,10.45,10.45	−.0849	−.0638	−.1829	−.2144
.5,10.25,10.25	−.2260	−.0544	−.0319	−.2038
1,10,10	−.1454	−.0385	−.0284	−.1353
2,10,10	−.1183	−.0134	−.0077	−.1101
3,10,10	−.1021	−.0137	−.0089	−.0969
5,10,10	−.0777	−.0073	−.0057	−.0721
.1,.1,10.8	−.1667	−.1261	+.1225	−.0764
.5,.5,10	−.3935	−.1055	−.0784	−.3617
1,1,10	−.2751	−.0583	−.0407	−.2574
2,2,10	−.1728	−.0203	−.0198	−.1598
3,3,10	−.1316	−.0210	−.0148	−.1253
5,5,10	−.0964	−.0128	−.0062	−.0917
1,1,1,1	−.3339	−.1039	−.0619	−.2919
2,2,2,2	−.1603	−.0418	−.0291	−.1472
3,3,3,3	−.1330	−.0271	−.0131	−.1190
5,5,5,5	−.0852	−.0134	−.0075	−.0770
10,10,10,10	−.0515	−.0062	−.0026	−.0472
.1,10.3,10.3,10.3	.1080	−.0546	−.1605	.1573
.5,10.17,10.17,10.17	−.0971	.0314	−.0174	−.0863
1,10,10,10	−.0766	.0213	.0135	−.0683
2,10,10,10	−.0568	.0100	−.0060	−.0542
3,10,10,10	−.0474	−.0080	−.0023	−.0418
5,10,10,10	−.0329	−.0067	−.0016	−.0278
.1,.1,10.4,10.4	.2091	−.1460	−.3065	−.3021
.5,.5,10,10	−.2103	.0700	−.0487	−.1866
1,1,10,10	−.1184	.0498	.0280	−.1045
2,2,10,10	−.0638	.0166	−.0131	−.0612
3,3,10,10	−.0638	−.0121	−.0052	−.0565
5,5,10,10	−.0427	−.0079	−.0032	−.0380
.1,.1,.1,10.7	−.2229	−.1555	−.1302	−.1976
.5,.5,.5,10.5	−.3233	−.0897	−.0660	−.2994
1,1,1,10	−.1663	.0679	.0380	−.1456
2,2,2,10	−.0975	.0182	−.0133	−.0924
3,3,3,10	−.0926	−.0154	−.0063	−.0833
5,5,5,10	−.0571	−.0096	−.0051	−.0526

LEMMA 2. *$D_j T_j^{(s)} = (-1)^{s-3} n^{-(s-3)/2} S_{s-3}(y_j) D_j^{s-2}$, where $y_j = n^{\frac{1}{2}} x_j - na_j$.*

PROOF.

$$(8.11) \quad \begin{aligned} D_j T_j^{(s)} &= D_j\left(\sum_{i=0}^{s-3} (-1)^i n^{-i/2} S_i(y_j) D_j^{\,i}\right) \\ &= \sum_{i=0}^{s-3} (-1)^i n^{-i/2} S_i(y_j) D_j^{\,i+1} + \sum_{i=0}^{s-3} (-1)^i n^{-i/2} [D_j S_i(y_j)] D_j^{\,i} . \end{aligned}$$

A property ([9] page 522) of the function $S_i(t)$ is that $(d/dt)S_i(t) = S_{i-1}(t)$. Thus

$$(8.12) \quad \begin{aligned} &\sum_{i=0}^{s-3} (-1)^i n^{-i/2} [D_j S_i(y_j)] D_j^{\,i} \\ &= \sum_{i=0}^{s-3} (-1)^i n^{-i/2} n^{\frac{1}{2}} S_{i-1}(y_j) D_j^{\,i} \qquad \text{where} \quad S_{-1}(t) \equiv 0 \\ &= -\sum_{i=0}^{s-4} (-1)^i n^{-i/2} S_i(y_j) D_j^{\,i+1} . \end{aligned}$$

From (8.11) and (8.12), Lemma 2 follows.

LEMMA 3. *If $x_j \in L_j$, then $\Delta_j T_j^{(s)} = n^{-\frac{1}{2}} D_j$.*

PROOF. $x_j \in L_j$ implies that $y_j = n^{\frac{1}{2}} x_j - na_j$ is integral; also

$$\begin{aligned} \Delta_j T_j^{(s)} &= T_j^{(s)}(y_j + 0) - T_j^{(s)}(y_j - 0) \\ &= \sum_{i=0}^{s-3} (-1)^i n^{-i/2} S_i(y_j + 0) D_j^{\,i} - \sum_{i=0}^{s-3} (-1)^i n^{-i/2} S_i(y_j - 0) D_j^{\,i} . \end{aligned}$$

For $i \neq 1$, $S_i(y_j + 0) = S_i(y_j - 0)$ because $S_i(t)$ is absolutely continuous. Thus the latter expression is equal to

$$-n^{-\frac{1}{2}}[S_1(y_j + 0) D_j - S_1(y_j - 0) D_j] = n^{-\frac{1}{2}} D_j .$$

Acknowledgments. The author is indebted to the faculty of the Department of Statistics, in particular to my advisor David L. Wallace, and to William H. Kruskal—who suggested the present investigation.

I am also indebted to Thomas Caldwell and David Walker for some of the computer programming.

The following provided useful comments on this article: Alex Mood, George Brown, Howard Tucker, Thomas S. Ferguson, Paul Meier, R. R. Bahadur, W. G. Cochran, Steve Webb.

REFERENCES

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

[2] BHATTACHARYA, R. N. (1967). Berry-Esséen bounds for the multidimensional central limit theorem. Ph. D. dissertation, Univ. of Chicago.

[3] COCHRAN, W. G. (1942). The χ^2 correction for continuity. *Iowa State College J. Sci.* **16** 421-436.

[4] CRAMÉR, H. (1928). On the composition of elementary errors. *Skand. Aktuarietidskr.* **11** 13-74 and 141-180.

[5] CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[6] ESSÉEN, C. G. (1945). Fourier analysis of distribution functions. *Acta. Math.* **77** 1-125.

[7] HALDANE, J. B. S. (1937). The exact value of the moments of the distribution of χ^2, used as a test of goodness of fit, when expectations are small. *Biometrika* **29** 133-143.

[8] HOEL, P. G. (1938). On the chi-square distribution for small samples. *Ann. Math. Statist.* **9** 158-165.

[9] KNOPP, K. (1951). *Theory and Application of Infinite Series*, (2d English ed.). Blackie and Son, London.

[10] RANGA RAO, R. (1961). Some problems in probability theory. Ph. D. dissertation, Calcutta University. Abstract in *Bull. Amer. Math. Soc.* **67** 359–361.

[11] YARNOLD, J. K. (1968). The accuracy of seven approximations for the null distribution of the chi-square goodness of fit statistic. Ph. D. dissertation, Univ. of Chicago.

[12] YARNOLD, J. K. (1970). The minimum expectation in chi-square goodness of fit tests, and the accuracy of approximations for the null distribution. *J. Amer. Statist. Assoc.* **65** 864–886.

[13] YARNOLD, J. K. (1972). Edgeworth approximations for the distribution of the quadratic form $X_n{}^T\sum^{-1}X_n$, when X_n is a sum of independent random vectors with a common continuous distribution.

19592 TRIDENT LANE
HUNTINGTON BEACH, CALIF. 92646

Added in proof. In the proof of Lemma 1, see Hildebrandt, T. H. (1963) *Introduction to the Theory of Integration*. Academič Press, New York. Theorem 13.8, page 60, for $k = 1$ and problem 6, page 139 for $k = 2$.

The second equality in (8.10) is given as Property I of the Stieltjes integral, page 60, Widder, D. V. (1947) *Advanced Calculus*. Prentice-Hall, New York.

FORTRAN IV computer programs for all of the computations in these articles are available from the author. One program calculates the exact upper tail probability of the χ^2 distribution and the seven approximations, at each possible value of χ^2, for arbitrary $n, m, p_1, \cdots, p_m$. The second program does the same (more efficiently) when all p_i are equal. The third program calculates the upper tail probability of the $C(\underline{m})$ distribution. The first program also computes $N(nc)$, $V(nc)$, and the lattice remainder $N(nc) - V(nc)$, which also arise in the lattice point problem in the theory of numbers.