AD-A225 640

(4)

EVALUATION OF NORMAL PROBABILITIES OF SYMMETRIC REGIONS

BY

SATISH IYENGAR

TECHNICAL REPORT NO. 433
JULY 26, 1990

Prepared Under Contract
N00014-89-J-1627 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

1. INTRODUCTION

Let $\mathbf{X} = (X_1,\ldots,X_n)'$ be a nonsingular multivariate normal random vector which is standardized so that $E(\mathbf{X}) = 0$ and $\mathrm{var}(\mathbf{X}) = R$, where $R = (\rho_{ij})$ with $\rho_{ii} = 1$ for all i, and let $A \in \mathbb{R}^n$ be a permutation-symmetric region. In this paper, we provide an approximation to $P(\mathbf{X} \in A)$ which is both easy to compute and accurate over a wide range of parameter values. In typical applications, A is an orthant $\{\underset{\sim}{x}: x_i \geq c;\ i=1,\ldots,n\}$, a union of orthants $\{\underset{\sim}{x}: \sum_k I(x_k \geq c) = p\}$, a cube $\{\underset{\sim}{x}: -c \leq x_i \leq c;\ i=1,\ldots,n\}$, or a sphere $\{\underset{\sim}{x}: \sum_k x_k^2 \leq c\}$.

The problem of evaluating $P(\mathbf{X} \in A)$ arises in many contexts. In genetic models of the transmission of traits, this probability is related to an individual's risk of contracting a disease: see Rice, et al.[24], Henery[8], and Curnow[2]. In geostatistics, it is related to the probability of correct classification in indicator kriging: see Journel[14], and Solow[28]. It is useful in the construction of simultaneous confidence intervals: see Miller[19], and Uusipaikka[32]. It also appears in the middle of some time series or regression problems: see especially Keenan[15], who shows that orthant probabilities are prediction error probabilities and transition probabilities for certain binary time series, and Keener and Waldman[16], who need to evaluate orthant probabilities (for dimensions as large as 35) for computing the likelihood function for rank-censored data. Our restriction on the set A may seem quite severe; however, this formulation is already useful in the applications just mentioned. In fact, the probability content of the positive orthant has been the subject of many papers: see Martynov[17], and Moran[21]. Also, the probability content of the sphere gives the distribution of a weighted sum of independent χ^2 variates: see Solomon and Stephens[27], and Imhof[10], for other approaches.

If a multivariate normal probability were required only infrequently, then ordinary simulation or numerical quadrature would often be adequate. However, in many applications (see especially [15] and [16]) the computation of such probabilities is just a small component of a much larger problem. This larger problem typically involves the estimation of parameters, and uses an iterative algorithm; thus a subroutine which evaluates multivariate probabilities is repeatedly called, and so speed and accuracy are important considerations.

This problem is an old one, and various approaches have been proposed; we now briefly review them. Numerical integration and simulation in high dimensions are known to be quite slow. Recently, however, Schervish[26] wrote a program for computing rectangular probabilities, and tested it for dimensions $n \leq 5$, and $\rho_{ij} \equiv \rho$. Simulation is, of course, always available in its simplest form, but standard variance reduction techniques are hard to obtain; a notable exception is the recent work of Moran[21], in which he gave a clever method for estimating the probability of the positive orthant, and provided a control variate for the estimator. Plackett[23] gave a reduction formula, but it seems practical only for $n \leq 5$. The study of probability inequalities (see Tong[31], and Eaton[5]) has provided many interesting insights and useful results, but the inequalities themselves often yield poor approximations; see Section 5 below for further discussion. The evaluation of multivariate normal probabilities is closely related to the determination of volumes of sets on the surface of the unit ball in $\mathbb{R}^n$, but this relationship has only occasionally been used: see for instance the work of Abrahamson[1], and Ruben[25]; there is a recent revival of geometrical methods in various contexts, as seen in the work of Diaconis and Efron[4], Johansen and Johnstone[12], and Naiman[22]. Next, expansions such as the tetrachoric series have been tried in several cases, but difficulties here include their slow convergence or even divergence over much of the parameter space: see Harris and Soms[7], and Moran[20]. Finally, sometimes a direct approach is rewarding. For instance, if $\rho_{ij} \equiv \rho \geq 0$, or if $\rho_{ij} = a_i a_j$ where $-1 < a_i < 1$, then many relevant computations reduce to the evaluation of one-dimensional integrals, which are easy to do: see Steck[29], and Curnow and Dunnett[3].

In this paper, we will exploit the symmetry of A to provide an approximation to $P(\mathbf{X} \in A)$ that is expressible in terms of one-dimensional integrals; such integrals are easily evaluated by computer. Some of the methods mentioned above come into play: for example, we will use the case $\rho_{ij} \equiv \rho$, about which we will expand a Taylor series which turns out to be an (integrated) Gram-Charlier series, and we will use Schervish's program for comparison. After establishing our notation in Section 2, we introduce a simple approximation in Section 3. We compute and study correction terms in Section 4, do two examples in some detail in Section 5, and summarize our numerical work there. We then conclude with a brief discussion in Section 6.

2. NOTATION AND PRELIMINARIES

Let $\mathbf{X} = (X_1,\ldots,X_n)'$ be a standardized nonsingular multivariate normal random vector whose distribution is denoted $N_n(0,R)$; thus $E(\mathbf{X}) = 0$, and $\mathrm{var}(\mathbf{X}) = R = (\rho_{ij})$ with $\rho_{ii}=1$ for all i. The density of $\mathbf{X}$ is denoted $\phi_n(\underset{\sim}{x};R)$; the univariate normal distribution is denoted $\Phi(x) = P(X_1 \leq x)$ and its density is $\phi(x)$. Let $d = n(n-1)/2$, and string out the correlations into a vector $\underset{\sim}{\rho} = (\rho_{12},\rho_{13},\ldots,\rho_{23},\ldots,\rho_{2n},\ldots)' \in \mathbb{R}^d$; also, let $\bar{\rho}$ be the average of the components of $\underset{\sim}{\rho}$, and let $\bar{\underset{\sim}{\rho}} = (\bar{\rho},...,\bar{\rho})' \in \mathbb{R}^d$. E_α is an n-by-n equicorrelation matrix with parameter α, where $-(n-1)^{-1} < \alpha < 1$. P_n is the set of n-by-n permutation matrices, an element of which is denoted π; Π is a random matrix which is independent of $\mathbf{X}$ and is uniformly distributed over P_n. The permutation-symmetric set $A \in \mathbb{R}^n$ satisfies $A = \pi A = \{\pi \underset{\sim}{a}: \underset{\sim}{a} \in A\}$ for all $\pi \in P_n$. Our main example below involves the positive orthant $Q_n = \{\underset{\sim}{x}: x_k \geq 0; k=1,...,n\} = \{\underset{\sim}{x}: \underset{\sim}{x} \geq 0\}$. Unless stated otherwise, all summations are over the entire range of the index variable: thus, for instance, $\sum_\pi$ is a summation over all of P_n. Finally, let

(1) $\quad h_n(\underset{\sim}{\rho}) = P(\mathbf{X} \in A).$

We suppress the dependence of h_n on A.

To derive his reduction formula, Plackett[23] proved and used the following identity for the multivariate normal:

(2) $\quad (\partial/\partial \rho_{ij})\ \phi_n(x;R) = (\partial^2/\partial x_i \partial x_j)\ \phi_n(x;R).$

This identity has been used to establish probability inequalities and monotonicity results for multivariate normal probabilities (see Tong[31]). We use it repeatedly below: it simplifies many computations, and is a natural way of deriving the multivariate Gram-Charlier series in Section 4.

The equicorrelated case is important in our discussion below, so we describe it now. If $Z = (Z_1,...,Z_n)'$ is a $N_n(0,I)$ vector, Z_o is a standard normal independent of Z, $\underset{\sim}{e} = (1,...,1)' \in \mathbb{R}^n$, and $\alpha \geq 0$, then $V = (1-\alpha)^{1/2}Z + \alpha^{1/2} Z_o \underset{\sim}{e}$ is a $N_n(0,E_\alpha)$ variate. Upon conditioning on Z_o, we get the single integral

(3) $$P(V \in A) = \int_{-\infty}^{\infty} P\left(Z \in \left(\frac{A - t\alpha^{1/2}\underset{\sim}{e}}{(1-\alpha)^{1/2}}\right)\right) \phi(t)\, dt.$$

When A is an orthant or a cube, the integrand in (3) is just a product of one-dimensional normal marginal probabilities, and if A is a sphere, it is a noncentral χ^2 probability. Thus, for many cases of interest, the right side of (3) is easily evaluated. The analogous formula for $\alpha<0$ is more involved, but still tractable: see Steck[29]. Also, moments of the form

(4) $\quad E[V_1^{a_1} \cdots V_n^{a_n}\, I(V \in A)]$

are needed below, and they can be expressed as a single integral just like (3). The same argument applies to the case $\rho_{ij}=a_i a_j$, with $-1 < a_i < 1$ for all i, which is generated by $V_i = (1-a_i^2)^{1/2}Z_i + a_i Z_o$, $i=1,\dots,n$.

3. A SYMMETRY ARGUMENT

Since $A = \pi A$ for all $\pi \epsilon P_n$, $P(\mathbf{X} \epsilon A) = P(\pi \mathbf{X} \epsilon A)$ for all $\pi \epsilon P_n$, and $P(\mathbf{X} \epsilon A) = P(\Pi \mathbf{X} \epsilon A)$ also. Consider $\Pi \mathbf{X}$; it has exchangeable components and it is a scale mixture of normals with density

$$(5) \qquad \psi_n(\underset{\sim}{x};R) = (n!)^{-1}\sum_{\pi} \phi_n(\underset{\sim}{x};\pi R \pi').$$

Its first two moments are $E(\Pi \mathbf{X}) = 0$ and $\mathrm{var}(\Pi \mathbf{X}) = (n!)^{-1}\sum_{\pi} \pi R \pi' = E_{\bar{\rho}}$. We get our simplest approximation by fitting to ψ_n the normal density $\phi_n(\underset{\sim}{x};E_{\bar{\rho}})$, which shares the same first two moments: thus,

$$(6) \qquad h_n(\underset{\sim}{\rho}) \cong h_n(\underset{\sim}{\bar{\rho}}).$$

This approximation has several appealing properties. First, it is easily evaluated for many cases by the argument leading to (3). Second, there is the following heuristic argument. Let P_d be the set of d-by-d permutation matrices, and let P_d' be the subset of it induced by the correlation vectors of $\{\pi \mathbf{X}: \pi \epsilon P_n\}$. The points $\{\pi \underset{\sim}{\rho}: \pi \epsilon P_d'\}$ are the vertices of a regular polygon centered at $\underset{\sim}{\bar{\rho}}$. Since h_n has the same value at each vertex of this polygon, and since it is a smooth function of $\underset{\sim}{\rho}$, its value at the center of the polygon should not be far from its value at a vertex. Third, it comes from the least squares fit of the equicorrelation matrices to R; that is, it minimizes $\sum_{i,j}(\rho_{ij}-a)^2$ over a. And fourth, a Taylor expansion of h_n about $\underset{\sim}{a}=(a,\dots,a)'$ gives

$$(7) \qquad h_n(\underset{\sim}{\rho}) = h_n(\underset{\sim}{a}) + Dh_n(\underset{\sim}{a})'(\underset{\sim}{\rho}-\underset{\sim}{a}) + \text{remainder},$$

where D denotes the gradient. Each component of $Dh_n(\underset{\sim}{a})$ is, by Plackett's identity (2),

$$(8) \qquad (\partial/\partial\rho_{ij}) \int_A \phi_n(\underset{\sim}{x};E_a)\, d\underset{\sim}{x} = \int_A (\partial/\partial\rho_{ij})\, \phi_n(\underset{\sim}{x};E_a)\, d\underset{\sim}{x}$$

$$= \int_A (\partial^2/\partial x_i \partial x_j)\, \phi_n(\underset{\sim}{x};E_a)\, d\underset{\sim}{x} = \int_A (\partial^2/\partial x_1 \partial x_2)\, \phi_n(\underset{\sim}{x};E_a)\, d\underset{\sim}{x}.$$

The interchange of the order of differentiation and integration in (8) is easily justified, and the last equation there comes from the symmetry of A. The linear term in (7) is thus zero if $a=\bar{\rho}$, so that (6) is also a good first order approximation.

Approximation (6) is based on only two moments, so it is not expected to be accurate for extreme tail probabilities, for example for $P(X_i \geq c;$ all $i)$ for large c; see Steck[30] and Iyengar[11] for a treatment of such cases. Also, it is clear from the heuristic argument above that the approximation should improve as $\sum_{i,j}(\rho_{ij}-\bar{\rho})^2$ decreases. Thus, we need higher moments of $\Pi\mathbf{X}$ to provide correction terms, so we now turn to the higher order terms of the Taylor expansion (7).

4. CORRECTION TERMS

To write the Taylor expansion of $h_n(\underset{\sim}{\rho})$ about $h_n(\bar{\underset{\sim}{\rho}})$, we need additional notation: let $\underset{\sim}{k} = (k_{1,2},\dots,k_{n-1,n})$ consist of nonnegative integers $k_{i,j}$ for $i<j$, and let $|\underset{\sim}{k}| = k_{1,2} + \dots + k_{n-1,n}$; let $\underset{\sim}{m} = (m_1,\dots,m_n)$, where $m_i = \sum_{i<j} k_{i,j} + \sum_{i>j} k_{j,i}$, and note that $|\underset{\sim}{m}| = 2|\underset{\sim}{k}|$. Next, let

(9) $\quad \underset{\sim}{\rho}^k = \rho_{1,2}^{\,k_{1,2}} \cdots \rho_{n-1,n}^{\,k_{n-1,n}};$

(10) $\quad D^k h_n(\underset{\sim}{\rho}) = (\partial^p/\partial\rho_{1,2}^{\,k_{1,2}} \cdots \partial\rho_{n-1,n}^{\,k_{n-1,n}})\, h_n(\underset{\sim}{\rho}),$ for $|\underset{\sim}{k}| = p;$

(11) $\quad D^m \phi_n(\underset{\sim}{x};R) = (\partial^{2p}/\partial x_1^{\,m_1} \cdots \partial x_n^{\,m_n})\, \phi_n(\underset{\sim}{x};R),$ for $|\underset{\sim}{m}| = 2p.$

(12) $\quad C(p,\underset{\sim}{k}) = p!\, /\, (k_{1,2}! \cdots k_{n-1,n}!),$ for $|\underset{\sim}{k}| = p.$

We then have

(13) $\quad h_n(\underset{\sim}{\rho}) = \sum_{p=0}^{N} (p!)^{-1} \sum_{|k|=p} C(p,\underset{\sim}{k})\, (\underset{\sim}{\rho} - \bar{\underset{\sim}{\rho}})^k\, D^k h_n(\underset{\sim}{\rho})$ + remainder.

Consider the integrands in (13): on the left side, we have $\phi_n(\underset{\sim}{x};R)$; by changing the order of differentiation and integration on the right side and by repeatedly using Plackett's identity (2) there, we get

(14) $\quad \phi_n(\underset{\sim}{x};R) = \sum_{p=0}^{N} (p!)^{-1} \sum_{|k|=p} C(p,\underset{\sim}{k})\, (\underset{\sim}{\rho} - \bar{\underset{\sim}{\rho}})^k\, D^m \phi_n(\underset{\sim}{x};E_{\bar{\rho}})$ + remainder

$$= \Sigma_{p=0}^{N} (p!)^{-1} \sum_{|k|=p} C(p,k) \; (\underset{\sim}{\rho} - \bar{\underset{\sim}{\rho}})^k \; H_{\underset{\sim}{m}}(\underset{\sim}{x}; E_{\bar{\rho}}) \; \phi_n(\underset{\sim}{x}; E_{\bar{\rho}}) + \text{remainder.}$$

where $H_{\underset{\sim}{m}}(x;E_{\bar{\rho}})$ is the m^{th} order Hermite polynomial. Thus, the Taylor expansion (13) yields the Gram-Charlier series expansion of $\psi_n(\underset{\sim}{x};R)$ about $\phi_n(\underset{\sim}{x};E_{\bar{\rho}})$; see Johnson and Kotz[13] and Mihaila[18] for a discussion of the multivariate Gram-Charlier series and Hermite polynomials. Of course, we could also interpret the integrands of (13) as the Gram-Charlier expansion of the mixture of normals $\phi_n(\underset{\sim}{x};R)$ about $\phi_n(\underset{\sim}{x};E_{\bar{\rho}})$; we get a different expansion, but either interpretation gives the same correction terms for the approximation (6).

For our purposes, the important feature of (13) is that each $D^k h_n(\bar{\underset{\sim}{\rho}})$ can be expressed as a one-dimensional integral; this is because of the Hermite polynomials of (14) and the argument leading to (3) and (4). It is easy to see that the p^{th} term of (13) requires (on the order of) p^d additions, where $d = n(n-1)/2$. Note that no matrix inversion is necessary to evaluate the correction terms. In the next section, we give details of one special case to illustrate our use of (13).

The convergence of (13) is a difficult issue in general. One special case that is tractable is the trivariate positive orthant, O_3, for which $h_3(\underset{\sim}{\rho}) = 1/2 - (1/4\pi)(\cos^{-1}(\rho_{12}) + \cos^{-1}(\rho_{13}) + \cos^{-1}(\rho_{23}))$. Here, (13) comes from the expansion of $\cos^{-1}$ about $\bar{\rho}$. For any fixed $\bar{\rho} \geq 0$ or $\bar{\rho} \leq 3\text{-}2(3^{1/2})$, (13) is convergent for all choices of $\underset{\sim}{\rho}$ leading to that $\bar{\rho}$ and yielding a positive definite matrix R. However, when $\bar{\rho} \in (3\text{-}2(3^{1/2}),0)$, (13) is not always convergent: take for instance $\underset{\sim}{\rho} = (0.45,-0.60,-0.75)'$; the radius of convergence for $\cos^{-1}$ about $\bar{\rho} = -0.30$ is 0.70, and $\rho_{12} = 0.45$ lies beyond that. We will rigorously study the convergence of (13) and (14) elsewhere; here, we depend upon numerical examples to assess the performance our approximation.

5. EXAMPLES

Our first example deals with the positive orthant $O_n = \{\underset{\sim}{x}: \underset{\sim}{x} \geq 0\}$. We give explicit expressions for the first two terms of the Taylor expansion (13), and describe their numerical evaluation. The algebra here is straightforward but rather tedious, so we suggest the use of MACSYMA for other applications. We only do the case $\bar{\rho} \geq 0$ here.

The first term of the Taylor expansion is

$$(15) \qquad h_n(\bar{\underset{\sim}{\rho}}) = \int_{-\infty}^{\infty} \Phi(t\tau^{1/2})^n \, \phi(t) \, dt,$$

where $\tau = \bar{\rho}/(1-\bar{\rho})$. This expression can be accurately evaluated using a Gaussian quadrature formula and Hill's [9] algorithm for computing Φ. For $n \leq 5$, quadrature is not necessary; this is because (15) reduces to 1/2, $1/2 - (1/2\pi)\cos^{-1}(\bar{\rho})$, and $1/2 - (3/4\pi)\cos^{-1}(\bar{\rho})$ for n = 1, 2, and 3, respectively. For n=4, a simple formula is not available, but Steck[29] provides an accurate approximation; and for n=5, we have

$$(16) \qquad h_5(\bar{\rho}) = \frac{-3}{4} + \frac{5}{4\pi}\cos^{-1}(\bar{\rho}) + \frac{5}{2}h_4(\bar{\rho}).$$

We need (15) for $n \leq 0$ below, so define $h_{-n}(\bar{\rho}) = 0$ for n=1,2,..., and $h_0(\bar{\rho}) = 1$. We also need the following restricted moments:

$$(17) \qquad f_n(\alpha; a_1, \ldots, a_t) = E[V_1^{a_1} \ldots V_t^{a_t} \, I(V \in Q_n)],$$

where V is a $N_n(0, E_\alpha)$ variate, $a_1 \geq \ldots \geq a_t \geq 1$, and $1 \leq t \leq n$. Using the argument leading to (3), we can rewrite the integral in (17) as a one-dimensional integral: for instance,

$$(18) \qquad f_n(\alpha;1) = (1-\alpha)^{1/2} \int_{-\infty}^{\infty} [\phi(t\delta) + t\delta\Phi(t\delta)] \, \Phi(t\delta)^{n-1} \, \phi(t) \, dt,$$

where $\delta = (\alpha/(1-\alpha))^{1/2}$. For small values of n, this integral can be evaluated explicitly; else, numerical quadrature is required.

Next, the quadratic form in (13) has an expansion whose terms are proportional to

$$(19) \qquad (\rho_{ij}-\bar{\rho})(\rho_{kl}-\bar{\rho}) \int_{Q_n} (\partial^4/\partial x_i \partial x_j \partial x_k \partial x_l) \, \phi_n(x; E_{\bar{\rho}}) \, dx,$$

with $i<j$ and $k<l$. Such a set $\{i,j,k,l\}$ has two, three, or four distinct elements. If there are four, then the integral in (19) is

$$(20) \qquad \phi_4(0; E_{\bar{\rho}}) \int_{Q_{n-4}} \phi_{n-4}(x; E_\alpha) \, dx,$$

where $\alpha = \bar{\rho}/(1+4\bar{\rho})$; the integral here is evaluated in the same way as (15). When $\{i,j,k,l\}$ has three distinct elements, the integral in (19) is

$$(21) \qquad (n-3) \, \phi_3(0; E_{\bar{\rho}}) \, A(\bar{\rho}) \, f_{n-3}(\beta;1),$$

where

$$(22)\qquad A(\bar{\rho}) = \bar{\rho}\,(1+3\bar{\rho})^{1/2}\ /\ (1+(n-1)\bar{\rho})\ ((1-\bar{\rho})(1+2\bar{\rho}))^{1/2},$$

and $\beta=\bar{\rho}/(1+3\bar{\rho})$. For $n\leq 6$, the expectation f_{n-3} in (21) can be evaluated explicitly in terms of elementary functions; we omit those details. And when $\{i,j,k,l\}$ has two distinct elements, the integral in (19) is

$$(23)\qquad \phi_2(0;E_{\bar{\rho}})\ [B(\bar{\rho})\ (f_{n-2}(\gamma;2)+(n-3)f_{n-2}(\gamma;1,1)) + C(\bar{\rho})\int_{Q_{n-2}} \phi_{n-2}(\underline{x};E_\gamma)\ d\underline{x}],$$

where

$$(24)\qquad B(\bar{\rho}) = (n-2)(1+2\bar{\rho})\ /\ (1-\bar{\rho}^2)\ (1+(n-1)\bar{\rho})^2,$$

$$(25)\qquad C(\bar{\rho}) = \bar{\rho}\ /\ (1-\bar{\rho})\ (1+(n-1)\bar{\rho}),$$

and $\gamma=\bar{\rho}/(1+2\bar{\rho})$; the expectations f_{n-2} here can be explicitly evaluated without numerical quadrature if $n\leq 5$. Thus, we can use the expressions (20),(21), and (23) to get the first correction term. We omit the next correction term which involves the third derivatives of $h_n(\underline{\rho})$.

We now turn to our numerical work. For dimensions n = 3, 4, and 5, we generated correlation matrices R with $\rho_{ij} \geq 0$ for all i,j. For each case, we computed the true probability of the positive orthant using the program of Schervish. We also computed approximation (6) and corrections from the second and third derivatives of (13). We then plotted the relative error of the three approximations against the "variance in ρ": $(n)^{-1} \sum_{i,j} (\rho_{ij} - \bar{\rho})^2$; these plots are shown in Figs. 1-9.

For n=3, the relative error decreases considerably as we add higher order terms of (13). Two parts of the scatterplot are clearly separated in Figs. 1 and 3 (and somewhat less clearly in Fig.2): the upper part corresponds to "extreme" $\underline{\rho}$ of the form $(0.0,\rho_{13},0.9)$, and the lower part corresponds to less extreme $\underline{\rho}$. For n = 4, the relative error of the third-order approximation is often smaller than that of the second-order one, although the difference between the two is not great; thus, the range of relative errors shown in Figs. 5 and 6 are the same. For n = 5, this improvement is somewhat greater. A conclusion from our work is that (for correlation matrices of this kind) three terms of the series (13) seem to be enough to give a relative error of less than about seven percent.

Next, we turn to the speed of our algorithm. In Table 1, we compare the time required (T_1) to compute the first three terms of our approximating series with the time required (T_2) to compute the exact probability (demanding three-place accuracy) using Schervish's program. All times are given in seconds; they are the average of five runs on a VAX 750. For all cases cited in Table 1, the relative error of our approximation was less than 0.03. Notice that Schervish's program is much faster when $\underset{\sim}{\rho}$ is close to $\bar{\underset{\sim}{\rho}}$ than when it is not. The reduction in the time needed is especially evident for n = 5 when the correlation matrix is not too close to the equicorrelated case.

Our second example involves exceedance probabilities. For a fixed constant $c \geq 0$, let $S_n = \sum_k I(X_k \geq c)$, and let $p_{n,k}(\underset{\sim}{\rho}) = P(S_n = k)$. Then

$$(26) \qquad p_n(\underset{\sim}{\rho}) = (p_{n,0}(\underset{\sim}{\rho}), \ldots, p_{n,n}(\underset{\sim}{\rho}))$$

is the exceedance distribution which we approximate by $p_n(\bar{\underset{\sim}{\rho}})$. Note that the mean of the exceedance distribution and that of its approximation are the same: $E(S_n;\underset{\sim}{\rho}) = E(S_n;\bar{\underset{\sim}{\rho}}) = n\Phi(-c)$. We now study the second moment of the exceedance distribution and that of its approximation; this study will yield new results about our approximation of orthant probabilities. The variance is given by $\mathrm{var}(S_n;\underset{\sim}{\rho}) = n\Phi(-c) - n^2\Phi(-c)^2 + \sum_{ij} P(X_i \geq c, X_j \geq c)$, which depends only on bivariate quadrant probabilities. We need the following lemma, the proof of which is an easy application of Plackett's identity.

LEMMA 1: Let V be a $N_2(0,R)$ variate with $\rho_{12} = r$. Then $F(r) = P(V_1 \geq c, V_2 \geq c)$ is convex for r in (0,1); for $c \geq 2^{1/2} - 1$, F is convex in (-1,1); and for c=0, F is concave in (-1,0).

This lemma yields the following

PROPOSITION 1: If (i) $\rho_{ij} \geq 0$ for all i and j, or (ii) if $c \geq 2^{1/2} - 1$, then $\mathrm{var}(S_n;\underset{\sim}{\rho}) \geq \mathrm{var}(S_n;\bar{\underset{\sim}{\rho}})$; if (iii) c=0 and $\rho_{ij} \leq 0$ for all i and j, then $\mathrm{var}(S_n;\underset{\sim}{\rho}) \leq \mathrm{var}(S_n;\bar{\underset{\sim}{\rho}})$.

Proof: The function $g(\underset{\sim}{\rho}) = \mathrm{var}(S_n;\underset{\sim}{\rho})$ is a symmetric function of the components of $\underset{\sim}{\rho}$. Under (i) or (ii), g is Schur convex function of $\underset{\sim}{\rho}$ (see Tong[31,p.106]); and under (iii), g is a Schur concave function of $\underset{\sim}{\rho}$. Since $\underset{\sim}{\rho}$ majorizes $\bar{\underset{\sim}{\rho}}$, the result follows.

Now consider the case of the positive orthant (c=0, and S_n=n) and suppose that $\rho_{ij} \geq 0$ for all i and j; let $\underset{\sim}{\alpha} = (\alpha,\ldots,\alpha)$, where $\alpha = \min\{\rho_{ij}\}$ and $\underset{\sim}{\beta} = (\beta,\ldots,\beta)$, where $\beta =$

max $\{\rho_{ij}\}$. From the variance inequality in Proposition 1, it is tempting to conjecture that

$$(27) \qquad p_{n,n}(\underline{\rho}) \geq p_{n,n}(\bar{\underline{\rho}}).$$

This is indeed true when n=3 (because $-\cos^{-1}$ is convex in (0,1)), and it represents an improvement over the bounds derived from the result of Slepian: $p_{n,n}(\underline{\alpha}) \leq p_{n,n}(\underline{\rho}) \leq p_{n,n}(\underline{\beta})$ (see Tong[31,p.10 and p.169]). Examples show that inequality (27) is not true for larger n. However, Figures 10-12 show that the right side of (27) is a much better approximation to the positive orthant probability than are the Slepian bounds: for the same correlation matrices studied above, these Figures compare the relative error of the Slepian bounds with the relative error of our approximation (6); our approximation was better in every case (except, of course, when the the matrix was already equicorrelated, in which case all three coincided with the true probability.)

The same method gives similar results for the cube: $T_n = \sum_k I(|X_k| \leq c)$, $q_{n,k} = P(T_n = k)$. The analogs of Lemma 1 and Proposition 1 come from the behavior of G(r) $= P(|V_1| \leq c, |V_2| \leq c)$.

6. DISCUSSION

The problem of evaluating multivariate normal probabilities is a difficult one, and it is likely that there is no panacea; that is, an approximation which is tailored to work well for one set of parameter values will probably be inadequate for another. Thus, many methods have been proposed in the literature. Our contribution to this literature is to show (from our theoretical and numerical work above) that (13) provides a good approximation to probabilities of permutation-symmetric regions, and that it is easily evaluated. One drawback of our proposal is that useful error bounds for our approximation are not yet available; here, we depend upon numerical work to assess the error. Our methods can also be applied to the problem of evaluation $Ef(\mathbf{X})$, where $f(x) = f(\pi x)$ for all $\pi \in P_n$, and also to random variables with other elliptically contoured distributions.

Acknowledgement: I thank Fred Huffer for his very generous help in this investigation, and Herbert Solomon for his encouragement. This work was supported by ONR contract N00014-76-C-0475 at Stanford University and 5 +32 MH 15758-07 at Carnegie-Mellon University. All computations were done on a VAX 750 at the Statistics Department at Carnegie-Mellon.

7. REFERENCES

1. Abrahamson,I. (1964) Orthant Probabilities for the Quadrivariate Normal Distribution. *Ann.Math.Stat.* 35,1685-1703.

2. Curnow,R. (1984) Progeny Testing for all-or-none Traits when a Multifactorial Model Applies. *Biometrics* 40,375-382.

3. Curnow,R.,et al. (1962) The Numerical Evaluation of Certain Multivariate Normal Integrals. *Ann.Stat.* 33,571-579.

4. Diaconis,P.,et al. (1985) Testing for Independence in a Two-Way Table: New Interpretations for the Chi-Square Statistic. *Ann.Stat.* 13,845-909.

5. Eaton,M.L. (1982) A Review of Selected Topics in Multivariate Probability Inequalities

6. Gupta,S. (1963) Bibliography on the Multivariate Normal Integrals and Related Topics. *Ann.Math.Stat.* 34,829-838.

7. Harris,B.,et al. (1980) The Use of the Tetrachoric Series for Evaluating Multivariate Normal Probabilities. *J.Mult.Analysis* 10,252-267.

8. Henery,R. (1981) An approximation to Certain Multivariate Normal Probabilities. *JRSS-B.* 43,81-85.

9. Hill,I.D. (1973) Algorithm AS66. The Normal Integral. *Applied Statistics.* 22,424-427.

10. Imhof,J.P. (1961) Computing the Distribution of Quadratic Forms in Normal Variables. *Biometrika.* 48,417-426.

11. Iyengar,S. (1986) On a Lower Bound for the Multivariate Normal Mills' Ratio. To appear in *Ann.Prob.*

12. Johansen,S.,et al. (1985) On Some Uses of Spherical Geometry in Simultaneous inference and Data Analysis. Stanford U. Tech. Rep.

13. Johnson,N.,et al. (1970) *Distributions in Statistics. Vol.1-3* Houghton-Mifflin.

14. Journel,A. (1982) Nonparametric Estimation of Spatial Distributions. *Math.Geol.* 15,445-468.

15. Keenan,D. (1982) A Time Series Analysis of Binary Data. *JASA.* 77,816-821.

16. Keener,R.,et al. (1985) Maximum Likelihood Regression of Rank-Censored Data. *JASA.* 80,385-392.

17. Martynov,G. (1981) Evaluation of the Normal Distribution Function. *J. of Soviet Math.* 17,1857-1875.

18. Mihaila,I. (1968) Development of the Trivariate Frequency Function in Gram-Charlier Series. *Rev. Roumaine de Math. Pures et Appliq.* 13,803-813.

19. Miller,R. (1981) *Simultaneous Statistical Inference.* Springer.

20. Moran,P. (1983) A New Expansion for the Multivariate Normal Integral. *Australian J. Stat.* 25,339-344.

21. Moran,P. (1984) The Monte Carlo Evaluation of Orthant Probabilities for Multivariate Normal Distributions. *Aust.Jour.Stat.* 26,39-44.

22. Naiman,D. (1983) Comparing Scheffe-type to Constant-width Confidence Bounds in Regression. *J.Amer.Stat.Assoc.* 78,906-912.

23. Plackett,R. (1954) A Reduction Formula for Multivariate Normal Integrals. *Biometrika.* 41,351-360.

24. Rice,J.,et al. (1978) Multifactorial Inheritance with Cultural Transmission and Assortative Mating. *Amer.J.Hum.Gen.* 30,618-643.

25. Ruben,H. (1959) On the Numerical Evaluation of a Class of Multivariate Normal Integrals. *Proc.Soc.Edin.* A65,272-281.

26. Schervish,M. (1984) Multivariate Normal Probabilities with Error Bound. *Applied Statistics.* 33,81-94.

27. Solomon,H.,et al. (1978) Approximations to Density Functions Using Pearson Curves. *JASA.* 73,153-160.

28. Solow,A. (1985) Mapping by Simple Indicator Kriging. Stanford U. Tech. Report.

29. Steck,G. (1962) Orthant Probabilities for the Equicorrelated Normal Distribution. *Biometrika.* 49,433-445.

30. Steck,G. (1979) Lower Bounds for the Multivariate Normal Mills' Ratio. *Ann.Prob.* 7,547-551.

31. Tong,Y. (1980) *Probability Inequalities in Multivariate Distributions.* Academic Press.

32. Uusipaikka,E. (1985) Exact Simultaneous Confidence Intervals for Multiple Comparisons Among Three or Four Mean Values. *J.Amer.Stat.Assoc.* 80,196-201.

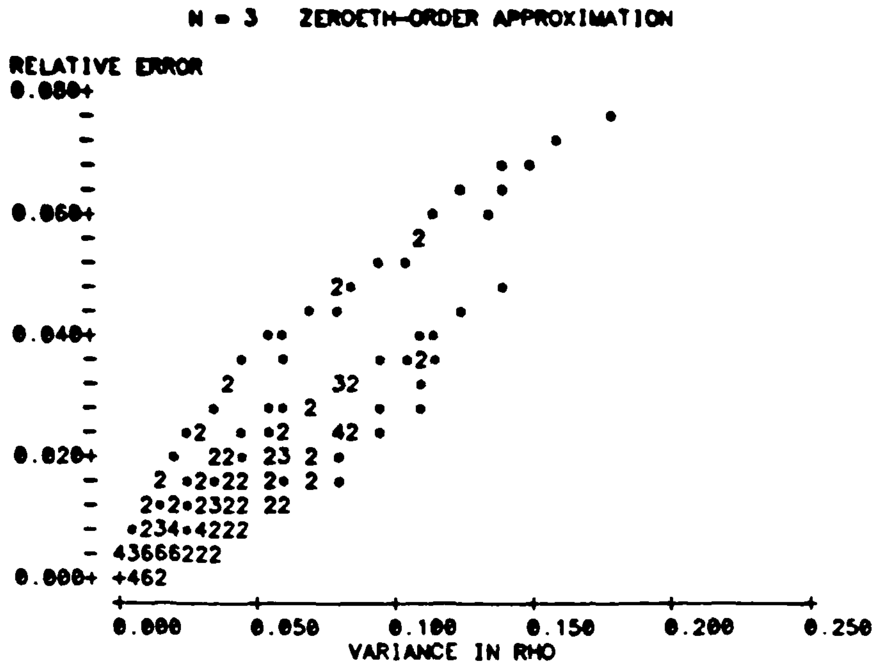
N = 3 ZEROETH-ORDER APPROXIMATION
RELATIVE ERROR
0.080
0.060
0.040
0.020
0.000
0.000
0.050
0.100
0.150
0.200
0.250
VARIANCE IN RHO

Figure 1.

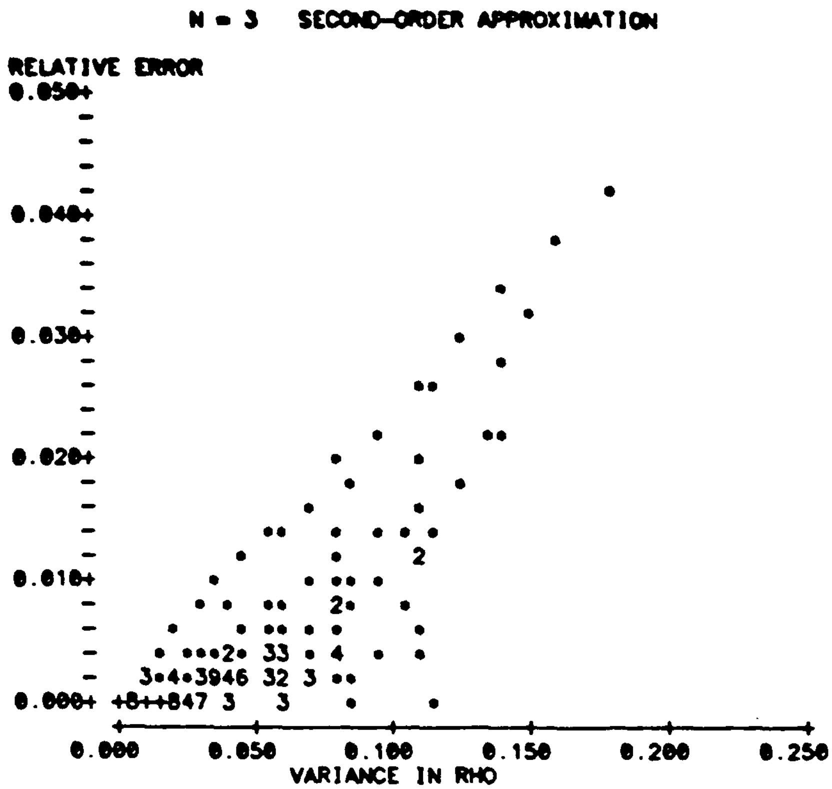
N = 3 SECOND-ORDER APPROXIMATION
RELATIVE ERROR
0.050
0.040
0.030
0.020
0.010
0.000
0.000
0.050
0.100
0.150
0.200
0.250
VARIANCE IN RHO

Figure 2.

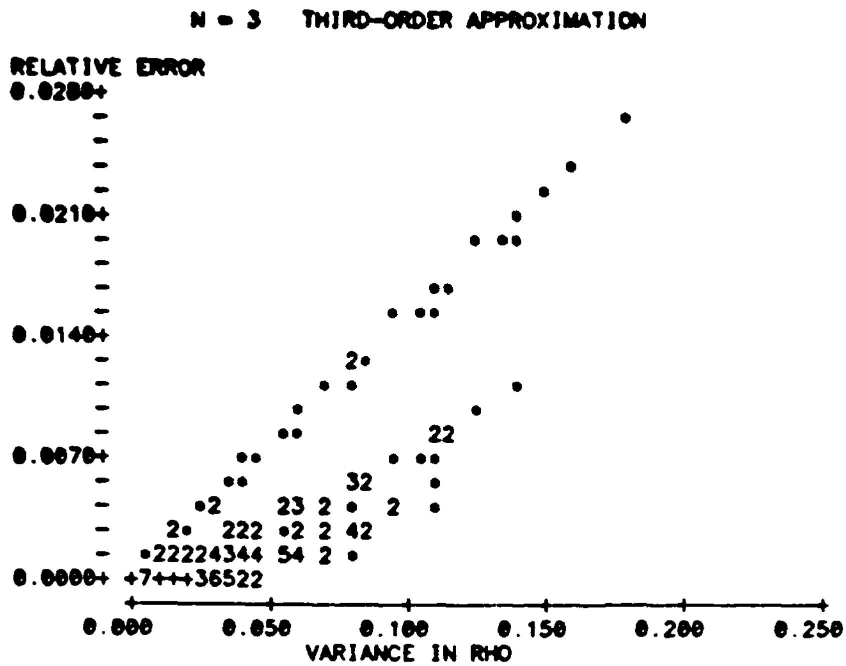
N = 3 THIRD-ORDER APPROXIMATION
RELATIVE ERROR
0.0280
0.0210
0.0140
0.0070
0.0000
0.000
0.050
0.100
0.150
0.200
0.250
VARIANCE IN RHO

Figure 3.

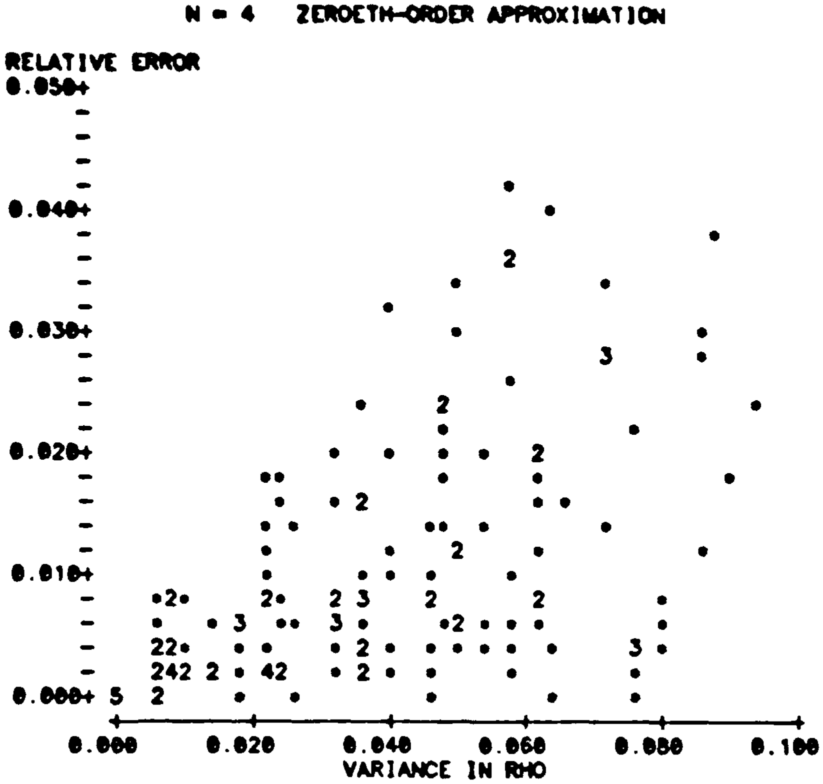
N = 4 ZEROETH-ORDER APPROXIMATION
RELATIVE ERROR
0.050+
0.040+
0.030+
0.020+
0.010+
0.000+
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

Figure 4.

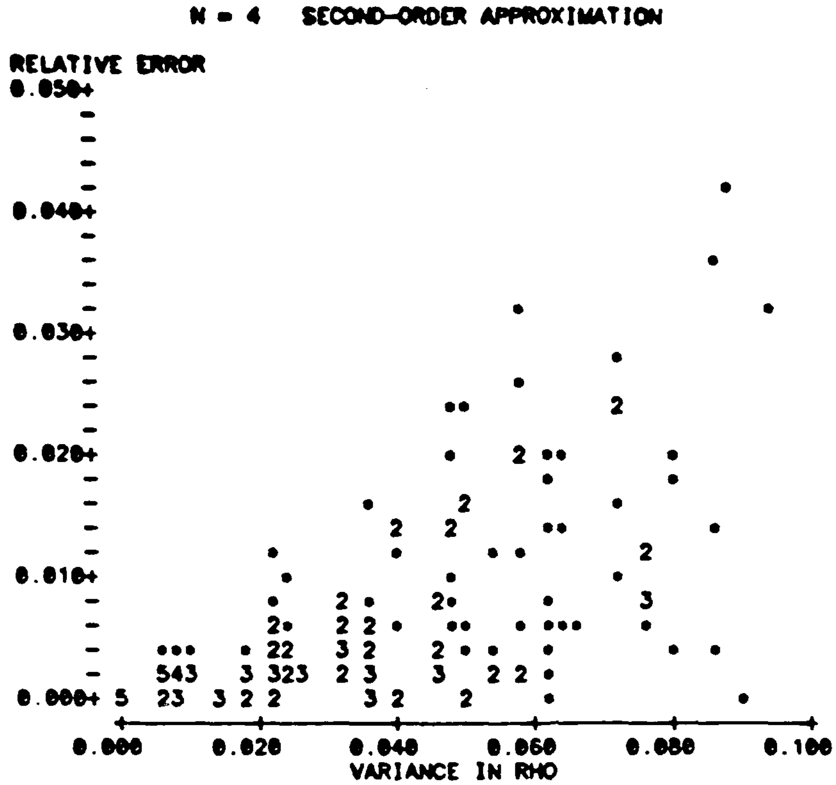
N = 4 SECOND-ORDER APPROXIMATION
RELATIVE ERROR
0.050
0.040
0.030
0.020
0.010
0.000
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

Figure 5.

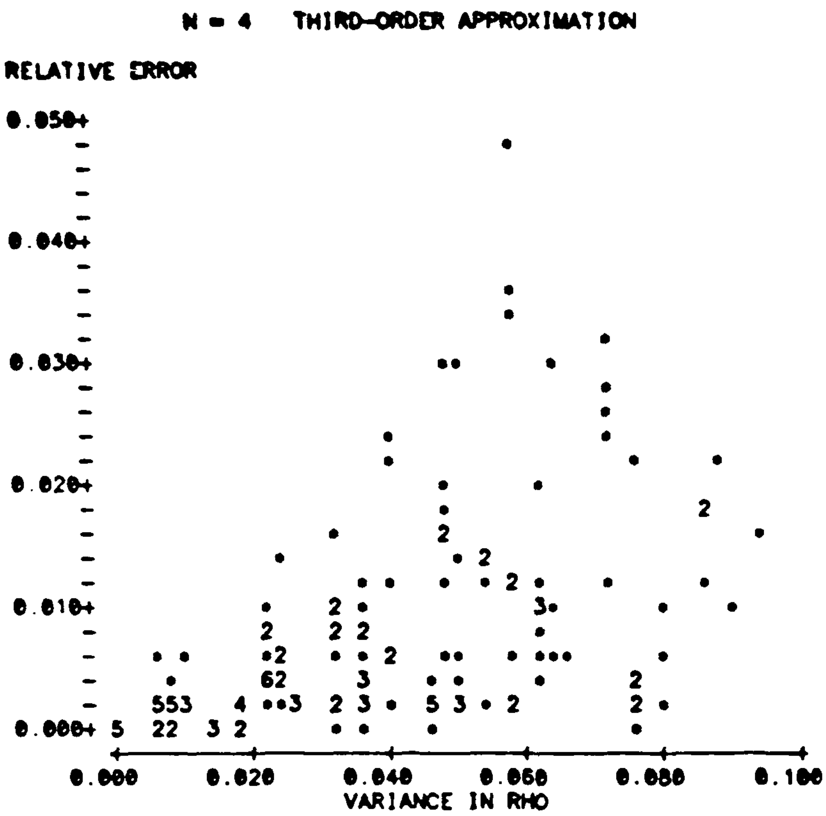
N = 4 THIRD-ORDER APPROXIMATION
RELATIVE ERROR
0.050
0.040
0.030
0.020
0.010
0.000
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

Figure 6.

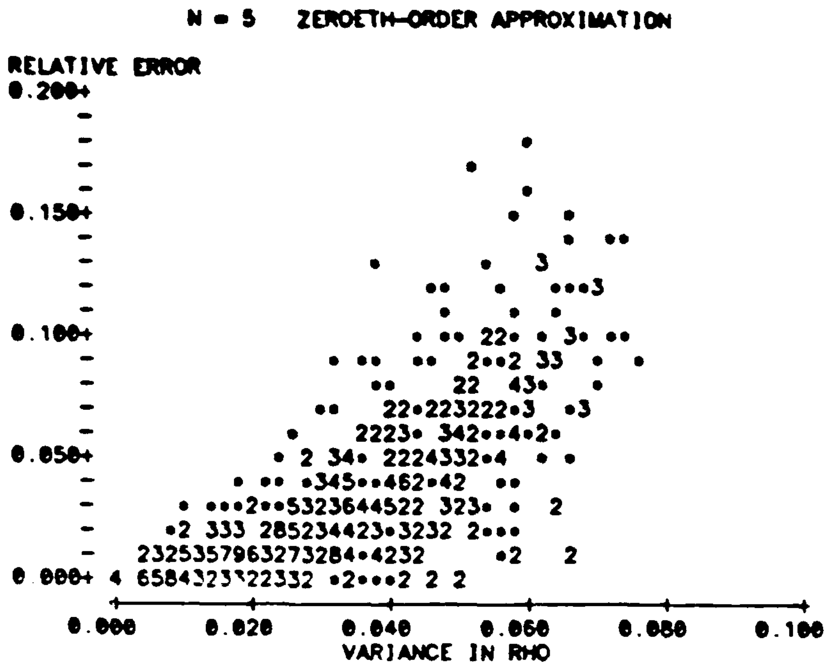
N = 5 ZEROETH-ORDER APPROXIMATION
RELATIVE ERROR
0.200
0.150
0.100
0.050
0.000
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

Figure 7.

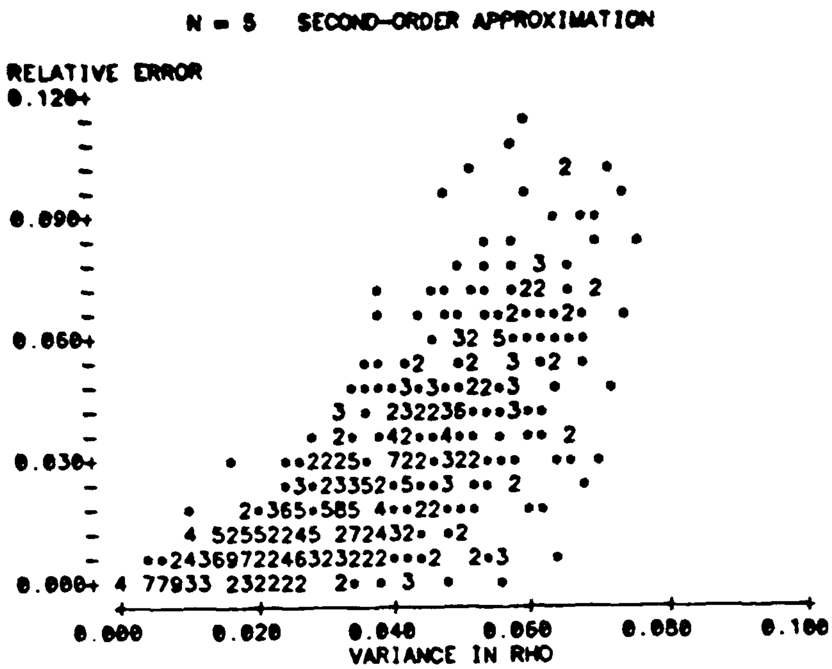
N = 5 SECOND-ORDER APPROXIMATION
RELATIVE ERROR
0.120
0.090
0.060
0.030
0.000
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

Figure 8.

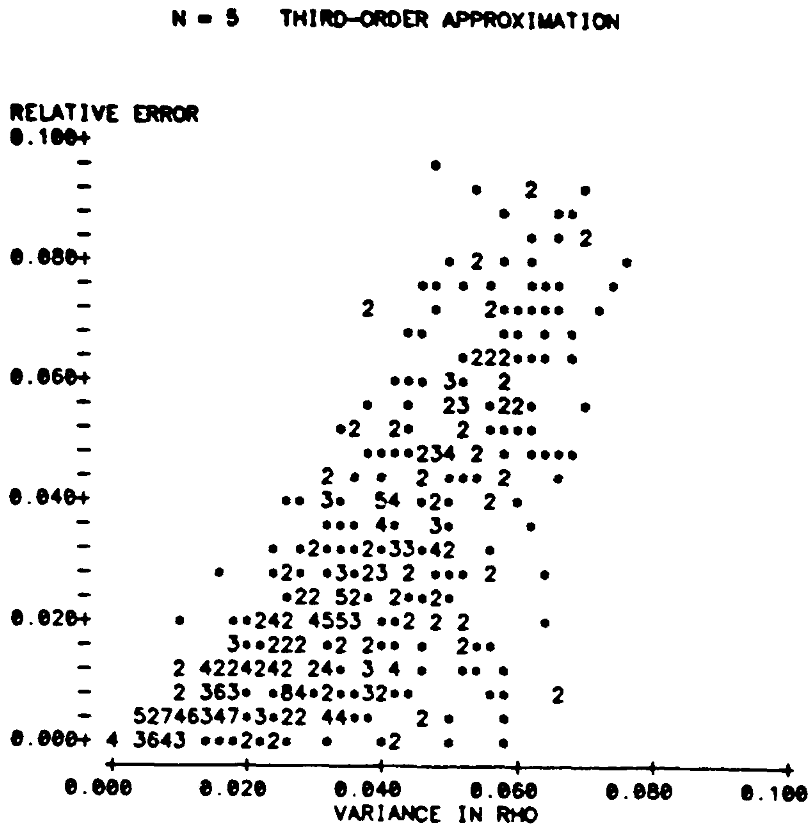
N = 5 THIRD-ORDER APPROXIMATION
RELATIVE ERROR
0.100
0.080
0.060
0.040
0.020
0.000
0.000
0.020
0.040
0.060
0.080
0.100
VARIANCE IN RHO

FIGURE 9.

N = 3

HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S LOWER BOUND - REL. ERROR OF APPROXIMATION (6))

```
0.00   10   ••••••••••
0.05   17   •••••••••••••••••
0.10   20   ••••••••••••••••••••
0.15   26   ••••••••••••••••••••••••••
0.20   30   ••••••••••••••••••••••••••••••
0.25   25   •••••••••••••••••••••••••
0.30   26   ••••••••••••••••••••••••••
0.35   22   ••••••••••••••••••••••
0.40   13   •••••••••••••
0.45    7   •••••••
```

HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S UPPER BOUND - REL. ERROR OF APPROXIMATION (6))

```
0.0    18   ••••••••••••••••••
0.1    43   •••••••••••••••••••••••••••••••••••••••••••
0.2    44   ••••••••••••••••••••••••••••••••••••••••••••
0.3    37   •••••••••••••••••••••••••••••••••••••
0.4    26   ••••••••••••••••••••••••••
0.5    15   •••••••••••••••
0.6    10   ••••••••••
0.7     2   ••
0.8     1   •
```

Figure 10.

N = 4

HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S LOWER BOUND - REL. ERROR OF APPROXIMATION (6))

0.00	5	•••••
0.05	3	•••
0.10	4	••••
0.15	5	•••••
0.20	10	••••••••••
0.25	9	•••••••••
0.30	12	••••••••••••
0.35	16	••••••••••••••••
0.40	14	••••••••••••••
0.45	19	•••••••••••••••••••
0.50	16	••••••••••••••••
0.55	14	••••••••••••••
0.60	10	••••••••••
0.65	10	••••••••••
0.70	2	••

HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S UPPER BOUND - REL. ERROR OF APPROXIMATION (6))

0.0	12	••••••••••••
0.2	24	••••••••••••••••••••••••
0.4	33	•••••••••••••••••••••••••••••••••
0.6	29	•••••••••••••••••••••••••••••
0.8	17	•••••••••••••••••
1.0	15	•••••••••••••••
1.2	8	••••••••
1.4	4	••••
1.6	4	••••
1.8	2	••
2.0	0	
2.2	1	•

Figure 11.

N = 5

```
   HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S LOWER BOUND - REL. ERROR OF APPROXIMATION (6))

     0.0       6    •••
     0.1      12    ••••••
     0.2      29    •••••••••••••••
     0.3      46    •••••••••••••••••••••••
     0.4      76    ••••••••••••••••••••••••••••••••••••••
     0.5      82    •••••••••••••••••••••••••••••••••••••••••
     0.6     100    ••••••••••••••••••••••••••••••••••••••••••••••••••
     0.7      86    •••••••••••••••••••••••••••••••••••••••••••
     0.8      37    •••••••••••••••••••


   HISTOGRAM OF THE DIFFERENCE
(REL. ERROR OF SLEPIAN'S UPPER BOUND - REL. ERROR OF APPROXIMATION (6))

     0.0      38    ••••••••
     0.5     140    ••••••••••••••••••••••••••••
     1.0     144    •••••••••••••••••••••••••••••
     1.5      69    ••••••••••••••
     2.0      42    •••••••••
     2.5      22    •••••
     3.0       9    ••
     3.5       4    •
     4.0       4    •
     4.5       1    •
     5.0       1    •
```

Figure 12.

Table 1

N	$\underset{\sim}{\rho}$	T_1	T_2
4	(.2,.2,.2,.2,.2,.2)	0.81	1.51
4	(.2,.4,.4,.6,.8,.8)	0.83	7.93
4	(.2,.2,.4,.4,.4,.8)	0.83	8.01
5	(.2,.2,.2,.2,.2,.2,.2,.2,.2,.2)	0.91	14.60
5	(.0,.0,.0,.2,.2,.2,.4,.4,.4,.4)	0.88	16.80
5	(.2,.2,.2,.4,.6,.6,.6,.6,.6,.8)	0.88	44.90

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 433	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Evaluation Of Normal Probabilities Of Symmetric Regions		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Satish Iyengar		8. CONTRACT OR GRANT NUMBER(s) N00014-89-J-1627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE July 26, 1990
		13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

multivariate normal probability, permutation-symmetric region, equicorrelated normal, positive orthant, exceedance distribution, mixture, Gram-Charlier expansion.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE FOLLOWING PAGE.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TECHNICAL REPORT NO. 433

20. ABSTRACT

Let $X = (X_1,\ldots,X_n)'$ be a standardized multivariate normal random vector and let $A \varepsilon \mathbb{R}^n$ be a permutation-symmetric region. We provide and justify an approximation to $P(X \varepsilon A)$ which is easy to compute, and derive correction terms based on a Gram-Charlier expansion. We then assess the performance of our technique numerically by evaluating positive orthant probabilities, and we prove some results for exceedance probabilities.