Volume Title: Annals of Economic and Social Measurement, Volume 5, number 4

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm76-4

Publication Date: October 1976

Chapter Title: Numerical Aspects of Multivariate Normal Probabilities in Econometric Models

Chapter Author: J. E. Dutt

Chapter URL: http://www.nber.org/chapters/c10495

Chapter pages in book: (p. 547 - 561)

Annals of Economic and Social Measurement 5/4, 1976

NUMERICAL ASPECTS OF MULTIVARIATE NORMAL PROBABILITIES IN ECONOMETRIC MODELS*

BY J. E. DUTT

The role of Multivariate Normal Probabilities in Econometric Models has in the past been somewhat restrictive because of the unavailability of useful computational formulas.

Using the author's recent integral representations for the Multivariate Normal Probability Integral, Dutt (1973) and (1975), highly accurate and efficient computational formulas are now available for computing normal probabilities of dimension up to 6. These formulas have direct application to the Maximum Likelihood procedures which are of interest in econometric modelling.

1. INTRODUCTORY SUMMARY

Prior to 1972 and after years of considerable effort, the only known general representation for multivariate normal upper and lower probabilities consisted of Pearson's tetrachoric series (Kendall, 1941) which is well-known to be computationally unattractive for dimension $K>2$. A reasonably complete bibliography relating to multivariate normal probabilities up to 1972 can be found in Johnson and Kotz (1972). Milton (1972) applied a method based on a multidimensional iterated Simpson's quadrature to the customary iterated form for either an upper or lower probability integral. Milton's computerized procedure, however, appears to be at least one order of magnitude in running time slower than what is now available.

In the recent paper Dutt (1973), this author obtained an integral transform representation over $(0, \infty)$ for upper and lower multivariate normal probabilities using Pearson's tetrachoric or orthogonal series, Kendall (1941), as a starting point. A simplified representation for the normal and an extension to the multivariate t are given in Dutt (1975). The representations are for arbitrary normal and t probabilities of arbitrary dimension and correlation matrix.

The integral transform representation for multivariate normal probabilities is very useful when numerical evaluation is by the Gauss-Hermite quadrature method. A short table based on the integral transform representation for the quadravariate normal orthant probability P_4 which, except for nearly singular correlation matrices, is accurate to 7+ significant digits, is found in Dutt and Lin (1975). A more extensive table for P_4, Dutt and Lin (1975a) and a short table for the trivariate normal, Dutt, Lin and Desai (1976) will be available shortly. Accurate computational formulas are also derived for the exponential, error and arcsin functions, Dutt, Lin and Tao (1973). Integral transform representations over $(0, \infty)$ for arbitrary upper and lower multivariate probabilities with application for computing bivariate and equicorrelated trivariate χ^2 probabilities is discussed in Dutt and Soms (1976). A table of the trivariate t for unequal correlations is found in Dutt, Mattes, and Tao (1975).

* Presented at the NBER-NSF Conference on Decision Making Under Uncertainty, University of Chicago, 16–17 May, 1975.

Attention here is focused on properties of the integral transform representation over $(0, \infty)$ for multivariate normal probabilities which might be of interest in econometric models. Numerical results are discussed for several correlation structures and dimensions up to six.

2. Integral Transform Representations over $(0, \infty)$ for Upper and Lower Multivariate Probabilities

Integral transform representations over $(0, \infty)$ are here summarized for an arbitrary continuous multivariate distribution and in particular for the multivariate normal. The integral representation follows in the general case from a slight modification of a theorem of Gurland, Gurland (1948), Dutt and Soms (1976). That such a modification was possible in general was only realized after the integral representation for the multivariate normal was derived from the tetrachoric series, Dutt (1973, 1975). Both approaches however, follow either directly or indirectly from the Inversion theorem.

Let $X_1, \ldots, X_K$ have the K dimensional cdf $F_K(\underset{\sim}{x})$ and corresponding characteristic function $\phi_K(\underset{\sim}{t})$. For $k \leq K$, let $\phi_{k:j_1,\ldots,j_k}(\underset{\sim}{t})$ be the characteristic function corresponding to the marginal distribution of $X_{j_1}, \ldots, X_{j_k}$ where $j_1, \ldots, j_k$ is a subset of the integers $1, \ldots, K$.

Now define $I_{k:j_1,\ldots,j_k}$ as the integral transform

$$(2.1) \quad I_{k:j_1,\ldots,j_k} = (1/2\pi)^k \int_0^\infty \cdots \int_0^\infty \{\text{Real}\,[\Delta_{k:j_1,\ldots,j_k}/i^k]/\prod_{r=1}^k t_r\}\,dt,$$

where $\Delta_{k:j_1,\ldots,j_k} = \Delta_k[e^{-it'\hat{a}}\phi_{k:j_1,\ldots,j_k}(\underset{\sim}{t})]$ and $\Delta_k[f(t_1, \ldots, t_k)]$ is the kth central difference about 0 of $f(t_1, \ldots, t_k)$

$$\begin{aligned}\Delta_k[f(t_1, \ldots, t_k)] &= f(t_1, \ldots, t_k) - f(-t_1, t_2, \ldots, t_k) \\ &\quad - \cdots - t(t_1, t_2, \ldots, t_{k-1}, -t_k) + f(-t_1, -t_2, t_3, \ldots, t_k) \\ &\quad + \cdots + (-1)^k f(-t_1, \ldots, -t_k).\end{aligned}$$

and $\underset{\sim}{\hat{a}}$ is a continuity point of the distribution of $x_{j_1}, \ldots, x_{j_k}$.

Then, from Dutt and Soms (1976, equation 2.3), if $\underset{\sim}{a} = (a_1, \ldots, a_k)$ is a continuity point of F_K, the integral transform representation over $(0, \infty)$ for an arbitrary continuous multivariate lower probability is

$$\begin{aligned}(2.2) \qquad F_K(\underset{\sim}{a}) &= (\tfrac{1}{2})^K - (\tfrac{1}{2})^{K-1} \sum_j I_1(a_j) \\ &\quad + (\tfrac{1}{2})^{K-2} \sum\sum_{i<j} I_2(a_i, a_j) \\ &\quad - (\tfrac{1}{2})^{K-3} \sum\sum\sum_{i<j<k} I_3(a_i, a_j, a_k) \\ &\quad + \cdots + I_K(a_1, \ldots, a_K).\end{aligned}$$

with $I_{k:j_1,\ldots,j_k} = I_k(a_{j_1}, \ldots, a_{j_k})$.

For upper probabilities, the negative signs are simply changed to positive ones. In Gurland's work, $I_{k;j_1,\ldots,j_k}$ relates to the real part of

$$(2.3)\qquad (2\pi i)^{-k}\oint\cdots\oint e^{-i\underline{t}'\hat{a}}\left\{\phi_{k;j_1,\ldots,j_k}(\underline{t})\Big/\prod_{r=1}^{k} t_r\right\}d\underline{t},$$

where $\hat{a}$ is a continuity point of the marginal distribution of $X_{j_1},\ldots,X_{j_k}$ and for any function $g(\underline{t})$, using the notation of Gurland (1948).

$$\oint\cdots\oint g(\underline{t})\,d\underline{t}=\lim_{\substack{\epsilon_r\to 0\\ T_r\to\infty}}\int\cdots\int_{\substack{\epsilon_r<|t_r|<T_r\\ r=1,\ldots,k}} g(\underline{t})\,d\underline{t}.$$

As they stand, the Cauchy principal value integrals in equation (2.3) are divergent. This can be seen in the case of the bivariate normal. One of the integrals in equation (2.3) is of the form

$$(2.4)\qquad \int_{\epsilon}^{T}\cos a(t_1+t_2)\phi_2(t_1,t_2;\rho)\,dt_1\,dt_2/t_1t_2$$

which as $\epsilon\to 0$, $T\to\infty$ is divergent. On the other hand, using equation (2.1) the integrand is bounded at the origin and the integrals can be used in numerical integration.

The integral representation over $(0,\infty)$ for the multivariate normal may be either treated as a special case of (2.2) or obtained from the tetrachoric series in the following way.

The K dimensional normal probability integral is defined as

$$L_K(x_1,\ldots,x_K;R)=\int_{x_1}^{\infty}\cdots\int_{x_K}^{\infty} n_K(\underline{y}|0,R)\,d\underline{y}$$

for any real numbers $x_1,\ldots,x_K$. The integrand $n_K(\underline{y}|0,R_K)$ denotes the K-dimensional standardized normal density with correlation matrix R.

Consider the representation of L_K by the tetrachoric series (Kendall, 1941)

$$(2.5)\qquad L_K(x_1,\ldots,x_K;R)=\sum_{j_{12}=0}^{\infty}\cdots\sum_{j_{(K-1)K}=0}^{\infty}A_{j_{12},\ldots,j_{(K-1)K}},$$

where

$$A_{j_{12},\ldots,j_{(K-1)K}}=\prod_{\substack{m,n=1\\ m<n}}^{K}(r_{mn}^{j_{mn}}/j_{mn}!)\prod_{k=1}^{K}(n_k!)\tau_{n_k}(x_k),$$

$$n_K=\sum_{i=1}^{K}j_{Ki}\quad(i\neq K),\ \sum_{r=1}^{K}n_r=2\bar{n},$$

say, and $\tau_m(x)$ is the mth tetrachoric function (Abramowitz & Stegun, 1964, p. 934).

$$\tau_m(x) = Z(x)He_{m-1}(x)/(m!)^{1/2}, \qquad m = 1, 2, \ldots, \tag{2.6}$$

with

$$Z(x) = (1/(2\pi)^{1/2}) \exp(-x^2/2)$$

and $He_n(x)$ is the nth degree Hermite polynomial

$$He_n(x) = [(-1)^n/Z(x)]\left(\frac{d}{dx}\right)^n Z(x), \qquad n = 0, 1, \ldots.$$

From the integral representation of the Hermite polynomial (Abramowitz & Stegun, 1964, p. 786), an integral representation of the tetrachoric function $\tau_m(x)$ is obtained as

$$\tau_m(x) = (1/\pi(m!)^{1/2}) \int_0^\infty \exp(-s^2/2)s^{m-1} \cos(xs - (m-1)\pi/2)\, ds, \qquad m = 1, 2, \ldots. \tag{2.7}$$

The direct substitution of $\tau_m(x)$ given by (2.7) into the tetrachoric series (2.5) yields after considerable manipulation Dutt (1973, 1975)

$$L_K(x_1, \ldots, x_K; R) = \tag{2.8}$$

$$(\tfrac{1}{2})^K - (\tfrac{1}{2})^{K-1} \sum_{i=1}^{K} D^*_{1;i} + (\tfrac{1}{2})^{K-2} \sum_{i<j=1}^{K} D^*_{2;ij} + (\tfrac{1}{2})^{K-3} \sum_{i<j<k=1}^{K} D^*_{3;ijk} + \ldots$$
$$+ D^*_{K:1,\ldots,K}.$$

The D^* functions are defined by

$$D^*_K(x; R) = 2(2\pi)^{-K} \int_0^\infty ds_1 \ldots \int_0^\infty ds_K\, e^{-s's/2}\, d^*_K(\underline{s}; x; R) \Big/ \prod_{k=1}^{K} s_k, \tag{2.9}$$

where, for the first few, K,

$d^*_1 = \sin_1 = \sin(x_1 s_1), \qquad d^*_2 = e_{-12} \cos_{1-2} - e_{12} \cos_{1+2},$

$d^*_3 = e_{12+13+23} \sin_{1+2+3} - e_{-12-13+23} \sin_{-1+2+3} - e_{-12+13-23} \sin_{1-2+3} - e_{12-13-23} \sin_{1+2-3},$

$d^*_4 = e_{12+13+23+14+24+34} \cos_{1+2+3+4} + e_{12-13-23-14-24+34} \cos_{-1-2+3+4}$

$\quad + e_{-12+13-23-14+24-34} \cos_{-1+2-3+4} + e_{-12-13+23+14-24-34} \cos_{1-2-3+4}$

$\quad - e_{-12-13+23-14+24+34} \cos_{-1+2+3+4} - e_{-12+13-23+14-24+34} \cos_{1-2+3+4}$

$\quad - e_{12-13-23+14+24-34} \cos_{1+2-3+4} - e_{12+13+23-14-24-34} \cos_{1+2+3-4};$

and $D^*_{k:j_1,\ldots,j_k} = D^*_k(x_{j_1}, \ldots, x_{j_k})$. For notation,

$$e_{p_1q_1+\ldots+p_mq_m} = \exp\left\{-(r_{p_1q_1}s_{p_1}s_{q_1} + \ldots + r_{p_mq_m}s_{p_m}s_{q_m})\right\},$$
$$\sin_{p_1+\ldots+p_m} = \sin(x_{p_1}s_{p_1} + \ldots + x_{p_m}s_{p_m}),$$
$$\cos_{p_1+\ldots+p_m} = \cos(x_{p_1}s_{p_1} + \ldots + x_{p_m}s_{p_m}),$$

where $R = ((r_{ij}))$

A negative sign on the index p_1q_1 corresponds to $+r_{p_1q_1}s_{p_1}s_{q_1}$ and $-p_1$ corresponds to $-x_{p_1}s_{p_1}$.

The k dimensional normal cumulative probability is defined by

$$\Phi_k(x_1, \ldots, x_k; R) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} n_k(y|0, R)\, d\underline{y}. \tag{2.10}$$

In terms of the upper tail probability L_k,

$$\Phi_k(x_1, \ldots, x_k; R) = L_k(-x_1, \ldots, -x_k; R) \tag{2.11}$$

3. Properties of D_k^* Functions

Consider again the integral representation over $(0, \infty)$ for L_k

$$L_k = (\tfrac{1}{2})^k - (\tfrac{1}{2})^{k-1} \sum_{i=1}^{k} D_{1;i}^* + (\tfrac{1}{2})^{k-2} \sum_{i<j=1}^{k} D_{2;ij}^* + \ldots + D_{k;1,\ldots,k}^*. \tag{2.8}$$

It is noted that in particular

$$D_1^*(x) = \tfrac{1}{2}\,\mathrm{erf}\,(x/\sqrt{2}) \tag{3.1}$$

with

$$D_1^*(0) = 0 \tag{3.2}$$

and

$$D_1^*(\infty) = \tfrac{1}{2}. \tag{3.3}$$

Since $D_k^*(\underline{x})$ for k odd, is a sine transform with argument $\underline{x}'\underline{s}$ then equation (3.2) generalizes to

$$D_k^*(\underline{0}) \equiv 0, \text{for } k \text{ odd}. \tag{3.4}$$

Moreover, it can be shown easily with equation (2.8) that by mathematical induction on k, equation (3.3) generalizes to

$$D_k^*(\underline{\infty}) = (-\tfrac{1}{2})^k, \text{for any } k > 1. \tag{3.5}$$

From a practical point of view, equations (3.5) means that to at least 3 digits of accuracy

$$D_k^*(\underline{4}) \doteq (-\tfrac{1}{2})^k, \text{for any } k > 1.$$

In general, $D_k^*(\underline{x})$, for $k > 1$ is non-negative and monotonically increasing for k even and, non-positive and monotonically decreasing for k odd. There is a slight inconsistency for $k = 1$ in that $D_1^*(x)$ is defined in equation (3.1) as a positive function for positive x.

For the orthant case (i.e. $\underline{x} = 0$) in addition to identity (3.4), it was previously noted, Dutt (1975) that

$$D_2^*(0, 0; r) = (\arcsin r)/2\pi. \tag{3.6}$$

If $P_k = L_k(0, \ldots, 0)$ and $D_k^{*0} = D_k^*(0, \ldots, 0)$ then from equation (2.8)

$$P_k = (\tfrac{1}{2})^k + (\tfrac{1}{2})^{k-2} \sum_{i<j=1}^{k} (\arcsin r_{ij})/2\pi$$

$$+ (\tfrac{1}{2})^{k-4} \sum_{i_1<i_2<i_3<i_4=1}^{k} D_{4,i_1i_2i_3i_4}^{*0}$$

$$+ \ldots + D_{k;1,\ldots,k}^{*0}.$$

3.1 *Transfer of Sign Changes from x_i to ρ_{ij}*

For $k = 1$ it is clear that

(3.1.1) $$D_1^*(-x) = -D_1^*(x)$$

For $k > 1$, attention need be focused on only d_k^*. For $k = 2$, consider

$$d_2^* = e^{+r_{12}s_1s_2} \cos(x_1s_1 - x_2s_2) - e^{-r_{12}s_1s_2} \cos(x_1s_1 + x_2s_2)$$

for a single sign change (i.e. either x_1 or x_2) and then for a double sign change (i.e. both x_1 and x_2).

For a single sign change

(3.1.2) $$D_2^*(-x_1, x_2; r_{12}) = D_2^*(x_1, -x_2; r_{12}) = -D_2^*(x_1, x_2; -r_{12})$$

while for the double sign change

(3.1.3) $$D_2^*(-x_1, -x_2; \rho_{12}) = D_2^*(x_1, x_2; \rho_{12}).$$

Therefore, for a single sign change a negative sign is transferred from either x_1 or x_2 to r_{12} with a negative sign in front of D_2^*. A double sign change of x_1 and x_2 leaves D_2^* unchanged.

For $k = 3$ and considering a sign change of x_1 to $-x_1$,

(3.1.4) $$D_3^*(-x_1, x_2, x_3; r_{12}, r_{13}, r_{23}) = -D_3^*(x_1, x_2, x_3; -r_{12}, -r_{13}, r_{23}).$$

The sign change is transferred to the correlation in which one of the subscripts $i = 1$.

For the double sign change $x_1 \to -x_1, x_2 \to -x_2$

(3.1.5) $$D_3^*(-x_1, -x_2, x_3; r_{12}, r_{13}, r_{23}) = D_3^*(x_1, x_2, x_3; r_{12}, -r_{13}, -r_{23})$$

and the triple sign change

(3.1.6) $$D_3^*(-x_1, -x_2, -x_3; r_{12}, r_{13}, r_{23}) = -D_3^*(x_1, x_2, x_3; r_{12}, r_{13}, r_{23}).$$

From equation (3.1.3) and (3.1.6) it should be clear that in general

$$D_k^*(\underline{x}; r_{ij}) = (-1)^k D_k^*(-\underline{x}; r_{ij}).$$

Moreover, for any r and k if $\underline{x} = (x_{j_1}, \ldots, x_{j_k})'$ then

$$D_k^*(-x_{j_1}, \ldots, -x_{j_t}, +x_{j_{t+1}}, \ldots, +x_{j_k}; r_{ij})$$
$$= (-1)^t D_k^*(\underline{x}; -r_{ij_1}, \ldots, -r_{ij_t}, +r_{ij_{t+1}}, \ldots, +r_{ij_k})$$

provided $1 \leq i \leq k$, $i \neq j_m$ and $1 \leq m \leq t$. In other words, a change of sign occurs only among the r_{ij} for which one subscript is from the set $j_1, \ldots, j_t$.

3.2 *Symmetry Considerations*

For $k = 2$, observe that

$$D_2^*(x_1, x_2; r_{12}) = D_2^*(x_2, x_1; r_{12})$$

which implies that for fixed x_1 and x_2, there are two equivalent probabilities

$$L_2(x_1, x_2; \rho_{12}) = L_2(x_2, x_1; \rho_{12}).$$

For $k = 3$, there are the following six equivalent D_3^*'s:

$$\begin{aligned} D_3^*&(x_1, x_2, x_3; r_{12}, r_{13}, r_{23}) \\ &= D_3^*(x_1, x_3, x_2; r_{13}, r_{12}, r_{23}) \\ &= D_3^*(x_2, x_1, x_3; r_{12}, r_{23}, r_{13}) \\ &= D_3^*(x_2, x_3, x_1; r_{23}, r_{12}, r_{13}) \\ &= D_3^*(x_3, x_1, x_2; r_{13}, r_{23}, r_{12}) \\ &= D_3^*(x_3, x_2, x_1; r_{23}, r_{13}, r_{12}). \end{aligned}$$

The six equivalent D_3^*'s lead to six equivalent L_3's.

For $k = 4$, the general pattern is readily apparent. For a fixed x_1, x_2, x_3 and x_4 and fixed $\{r_{ij}\}$, there are $4! = 24$ equivalent D_4^*'s relating to the permutations of (1, 2, 3, 4). However, for a given set of the six correlation coefficients there are $6! = 720$ corresponding D_4^*'s which taken together with the above mentioned 24 yields a total of 30 distinct D_4^*'s with each occurring in 24 equivalent ways. The permutations of $\{\rho_{ij}\}$ yield, therefore, the total number of probabilities and the permutations of $\{x_i\}$, the subset of equivalent probabilities.

3.3 *Mixed (Upper and Lower) Probability Integral*

If the mixed probability integral is defined as

$$J_k = \int_{-\infty}^{x_{j_1}} \cdots \int_{-\infty}^{x_{j_t}} \int_{x_{j_{t+1}}}^{\infty} \cdots \int_{x_{j_k}}^{\infty} n_k(\underline{y}/\underline{\rho}, \{r_{ij}\})\, d\underline{y}$$

then

$$J_k = L_k(-x_{j_1}, \ldots, -x_{j_t}, x_{j_{t+1}}, \ldots, x_{j_k}; \quad (3.2.1)$$
$$-r_{ij_1}, \ldots, -r_{is_{t+1}}, \ldots, r_{ij_k})$$

provided $1 \le i \le k$, $i \ne j_m$ and $1 \le m \le t$.

The results of Sections 3.1 and 3.2 would apply also to the mixed case.

4. Summary of Computing Forms for L_j

There are a variety of different options in computing the D_j^* functions depending on the primary needs of the user. There may be, for example, interest in high accuracy (of the order of 7–8 digits) for a relatively small list of probabilities (less than 100), or moderate accuracy (of the order of 3–4 digits) for a large number of iterations (1,000+), or interest in dimensions greater than five where overflow can be a problem. Clearly, the more specific the case of interest the greater would be the advantages in computer running time and accuracy. Moreover, it would not be prudent to use the geneal form of L_k for the equcorrelated case or for one of the orthant cases.

A summary of the computing formulas for L_j are therefore presented in their most general form with the integration coefficients specified so as to anticipate overflow in the higher dimensions.

For $k = 1$ in equation (2.8)

$$L_1(x_1) = \tfrac{1}{2} - D^*_{1:1} \tag{4.1}$$

where

$$D^*_{1:1} = (1/\pi) \int_0^\infty \sin(x_1 s_1)\, e^{-s_1^2/2}\, ds_1/s_1.$$

Application of the Gaussian quadrature formula (Abramowitz and Stegun, 1964, p. 924) yields the computing formula for $D^*_{1:1}$ as

$$D^*_{1:1} \doteq (\sqrt{2}/\pi) \sum_{i=1}^{N} y_i \sin(\lambda_i x_1)$$

where $y_i = w_i/\lambda_i$, $\lambda_i = z_i\sqrt{2}$, $\{w_i\}$ are the Christoffel weight factors, and $\{z_i\}$ are the zeros of the Mth degree Hermite polynomial. With M even, $N = M/2$ so that N denotes only the positive zeros. The $\{y_i\}$ and $\{\lambda_i\}$ are used for all D^*_k. The $\{z_i\}$ and $\{w_i\}$ are found in Stroud and Secrest (1966).

For $k = 2$, equation (2.8) yields

$$L_2(x_1, x_2; \rho_{12}) = \tfrac{1}{4} - \tfrac{1}{2}[D^*_{1:1} + D^*_{1:2}] + D^*_{2:12} \tag{4.2}$$

where

$$D^*_{2:12} = (1/\pi^2) \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j d^*_2$$

$$d^*_2 = \exp[\rho_{12}\lambda_1\lambda_2] \cos(x_1\lambda_2 - x_2\lambda_2) - \exp[-\rho_{12}\lambda_1\lambda_2] \cos(x_1\lambda_1 + x_2\lambda_2).$$

For $k = 3$

$$L_3(x_1, x_2, x_3; \rho_{12}, \rho_{13}, \rho_{23}) = 1/8 - \tfrac{1}{4}[D^*_{1:1} + D^*_{1:2} + D^*_{1:3}] + \tfrac{1}{2}[D^*_{2:12} + D^*_{2:13} + D^*_{2:23}] + D^*_{3:123} \tag{4.3}$$

where

$$D^*_{3:123} = (1/\sqrt{2}\pi^3) \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} y_i y_j y_k d^*_3.$$

For $k = 4$

$$\begin{aligned} L_4(x_1, x_2, x_3; \rho_{12}, \ldots, \rho_{34}) &= 1/16 - \tfrac{1}{8}[D^*_{1:1} + D^*_{1:2} + D^*_{1:3} + D^*_{1:4}] \\ &+ \tfrac{1}{4}[D^*_{2:12} + D^*_{2:13} + D^*_{2:23} + D^*_{2:14} + D^*_{2:24} + D^*_{2:34}] \\ &+ \tfrac{1}{2}[D^*_{3:123} + D^*_{3:124} + D^*_{3:134} + D^*_{3:234}] \\ &+ D^*_{4:1234} \end{aligned} \tag{4.4}$$

where

$$D^*_{4:1234} = (1/2\pi^4) \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{r=1}^{N} y_i y_j y_k y_r \, d^*_4.$$

For $k = 5$,

$$(4.5) \qquad \begin{aligned} &L_5(x_1, x_2, x_3, x_4, x_5; \rho_{12}, \ldots, \rho_{45}) \\ &= 1/32 - \tfrac{1}{16}[D^*_{1:1} + D^*_{1;2} + D^*_{1;3} + D^*_{1;4} + D^*_{1;5}] \\ &\quad + \underbrace{\tfrac{1}{8}[D^*_{2:12} + D^*_{2:13} + \ldots + D^*_{2;45}]}_{10 \text{ terms}} \\ &\quad + \tfrac{1}{4}[D^*_{3;123} + D^*_{3:124} + D^*_{3:134} + D^*_{3:234} + D^*_{3:125} \\ &\quad + D^*_{3;135} + D^*_{3:235} + D^*_{3;145} + D^*_{3:245} + D^*_{3;345}] \\ &\quad + \tfrac{1}{2}[D^*_{4:1234} + D^*_{4:1235} + D^*_{4:1245} + D^*_{4;1345} + D^*_{4;2345}] \\ &\quad + D^*_{5:12345} \end{aligned}$$

where

$$D^*_{5;12345} = (1/2\sqrt{2}\pi^5) \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{r=1}^{N} \sum_{s=1}^{N} y_i y_j y_k y_r y_s \, d^*_5$$

$$\begin{aligned} &d^*_5(\lambda_i, \lambda_j, \lambda_k, \lambda_r, \lambda_s; x_1, x_2, x_3, x_4, x_5) \\ &= \{\exp[- - - - - - - - - -] \sin(+ + + + +) \\ &\quad - \exp[+ + - + - - + - - -] \sin(- + + + +) \\ &\quad - \exp[+ - + - + - - + - -] \sin[+ - + + +) \\ &\quad - \exp[- + + - - + - - + -] \sin(+ + - + +) \\ &\quad - \exp[- - - + + + - - - +] \sin(+ + + - +) \\ &\quad - \exp[- - - - - - + + + +] \sin(+ + + + -) \\ &\quad + \exp[- + + + + - + + - -] \sin(- - + + +) \\ &\quad + \exp[+ - + + - + + - + -] \sin(- + - + +) \\ &\quad + \exp[+ + - - + + - + + -] \sin(+ - - + +) \\ &\quad + \exp[+ + - - + + + - - +] \sin(- + + - +) \\ &\quad + \exp[+ - + + - + - + - +] \sin(+ - + - +) \\ &\quad + \exp[- + + + + - - - + +] \sin(+ + - - +) \\ &\quad + \exp[+ + - + - - - + + +] \sin(- + + + -) \\ &\quad + \exp[+ - + - + - + - + +] \sin(+ - + + -) \\ &\quad + \exp[- + + - - + + + - +] \sin(+ + - + -) \\ &\quad + \exp[- - - + + + + + + -] \sin(+ + + - -)\} \end{aligned}$$

where the order of correlation is

$$\rho_{12}, \rho_{13}, \rho_{23}, \rho_{14}, \rho_{24}, \rho_{34}, \rho_{15}, \rho_{25}, \rho_{35}, \rho_{45}.$$

5. Numerical Results

As was mentioned in the Introduction, a variety of tables and numerical results are now available based on integral transform representations of multivariate probabilities. In the normal case, these evaluations cover primarily L_3, L_4, P_4 and P_5 and, are probably far more accurate than necessary for most statistical applications. An accuracy of four significant digits seems reasonable particularly for higher dimensions. It is also clear that for higher dimensions excessive computer running time becomes a serious problem and some alternate approach is needed.

The value of N, the number of positive Hermite zeros, needed for a specified accuracy depends to a large degree on the determinant $|R|$ of the correlation matrix R and the limits $\{x_i\}$ of the probability integral. A guide for choosing N based on $|R|$ and max x_i so as to achieve 4 digit accuracy (i.e., four correct digits after the decimal point) in D_k^* is available in Table 5.1. Any rule based solely on $|R|$ would not be completely adequate. However, for a given accuracy, the required N does not seem to vary significantly for $D_3^*, \ldots, D_6^*$ although larger x_i require higher N. Somewhat larger N are needed for D_2^* while generally smaller N would be more satisfactory for D_6^* than for D_4^*, and more satisfactory for D_5^* than for D_3^*.

In the four variate orthant case, a table of P_4 is available for the 21 correlation sets of Bacon (1963) as a function of N and $|R|$, Dutt (1973). An enlargement of that table to include Φ_4 for $x_i = 1, 2$ and 3, as a function of $|R|$ and the N which are sufficient for four digit accuracy appears in Table 5.2. The required N appears in parentheses. It is also noted that a given N is satisfactory for both $+x_i$ and $-x_i$ so that the Φ_4's could be replaced by a corresponding set of upper tail probabilities, L_4's.

TABLE 5.1

A Guide Based on $|R|$ and max x_i for Choosing N in $D_k^*(k>2)$ for Four Digit Accuracy

$\lvert R\rvert$ \ max x_i	0*	1	2	3
$0.7<\lvert R\rvert\le 1$	1–2	2–3	3–4	5–6
$0.5<\lvert R\rvert\le 0.7$	2–3	2–3	3–4	5–6
$0.3<\lvert R\rvert\le 0.5$	2–4	3–4	3–4	5–6
$0.2<\lvert R\rvert\le 0.3$	3–4	3–4	4–6	5–6
$0.1<\lvert R\rvert\le 0.2$	3–6	4–6	4–8	6–8
$0.05<\lvert R\rvert\le 0.1$	4–6	4–6	4–8	6–8
$0<\lvert R\rvert\le 0.05$	4–12	4–12	4–12	6–12

* For the orthant case $(x_i=0)$, $D_3^* = D_5^* \equiv 0$

TABLE 5.2

Φ_4 FOR THE 21 SETS IN BACON AS A FUNCTION OF N AND x*

Case	r_{12}	r_{23}	r_{34}	r_{14}	r_{13}	r_{24}	$X = 1$	$x = 2$	$x = 3$	$\|R\|$
18	0.50	0.50	0.309 . . .	0	0	0	0.5657(4)	0.9201(4)	0.9948(6)	0.42838137
20	0.50	0.50	0.50	0	0.50	0	0.5930(4)	0.9241(4)	0.9949(6)	0.31250000
14	0.50	0.50	0.50	0	0	0	0.5773(4)	0.9219(4)	0.9948(6)	0.31250000
21	0.612 . . .	0.666 . . .	0.612 . . .	0.408 . . .	0.25	0.408 . . .	0.6322(4)	0.9302(4)	0.9951(6)	0.21701389
17	0.309 . . .	0.809 . . .	0.309 . . .	0	0	0	0.5792(4)	0.9239(4)	0.9950(6)	0.16362712
9	0.50	0.309 . . .	0.809 . . .	0	0	0	0.5938(4)	0.9259(4)	0.9951(6)	0.16362712
8	0.50	0.50	0.707 . . .	0	0	0	0.5939(4)	0.9251(4)	0.9950(6)	0.12500000
16	0.50	0.809 . . .	0.135 . . .	0	0	0	0.5826(6)	0.9249(6)	0.9951(6)	0.08181356
13	0.50	0.790 . . .	0.25	0	0	0	0.5860(4)	0.9248(6)	0.9951(6)	0.07812500
3	0.790 . . .	0.25	0.790 . . .	0	0	0	0.6150(6)	0.9303(4)	0.9953(6)	0.07812500
11	0.809 . . .	0.50	0.309 . . .	0	0	0	0.5925(4)	0.9258(6)	0.9951(6)	0.06250000
10	0.50	0.707 . . .	0.50	0	0	0	0.5929(4)	0.9251(4)	0.9950(6)	0.06250000
1	0.925 . . .	0.135 . . .	0.809 . . .	0	0	0	0.6324(4)	0.9351(8)	0.9957(8)	0.03125000
7	0.50	0.50	0.709 . . .	0	0	0	0.6029(6)	0.9271(6)	0.9951(6)	0.03125000
15	0.309 . . .	0.925 . . .	0.135 . . .	0	0	0	0.5873(8)	0.9273(6)	0.9953(10)	0.03125000
12	0.50	0.809 . . .	0.309 . . .	0	0	0	0.5915(6)	0.9257(6)	0.9951(6)	0.02387287
2	0.809 . . .	0.309 . . .	0.809 . . .	0	0	0	0.6224(4)	0.9316(6)	0.9954(6)	0.02387287
5	0.50	0.309 . . .	0.925 . . .	0	0	0	0.6113(10)	0.9302(8)	0.9954(10)	0.01193643
6	0.50	0.50	0.809 . . .	0	0	0	0.6052(8)	0.9276(6)	0.9951(8)	0.00911862
4	0.707 . . .	0.50	0.707 . . .	0	0	0	0.6115(8)	0.9285(6)	0.9952(6)	0
19	0.50	0.50	0.50	0.50	0	0	0.6028(4)	0.9250(4)	0	0

* In a number of cases the listed N in ( ) is satisfactory for five significant digits.

Work is now in progress to complete the computer routines for D_7^* and D_8^* which would then permit computation of the general normal probability L_k for all $k \leq 8$ and for the orthant probability P_k for all $k \leq 9$. The approach then would be to look for simple approximations for $D_2^*, \ldots, D_8^*$ which would be accurate to about three or four digits and which would be generally useful in iterative maximum likelihood procedures. Numerical properties of $D_2^*, \ldots, D_6^*$ are examined graphically in an appendix available on request from the author for four correlation matrices which are identified as Equicorrelated, Markov, Toeplitz and Nested. The corresponding probabilities $L_2, \ldots, L_6$ for the first three matrices are available in that paper.

6. Applications

Two applications where computational formulas of multivariate normal probabilities would be useful are briefly discussed. The first application relates to a model of contraception discussed by Heckman and Willis (1973) in which the mathematical details are here presented in a slightly more general way. The second application pertains to the multivariate probit problem, Ashford and Sowden (1970). Other applications might be inferred from McFadden (1974). The maximum likelihood method is used for illustration purposes although other estimation methods are available, Amemiya (1972). See also Tobin (1955, 1958) for an application in economics.

6.1 *Application # 1—A Model of Contraception*

Consider a set of continuous dependent random variables $S_1, S_2, \ldots$ where the index refers to time. Specifically, let S_j denote a woman's "level of contraception" at month j and consider M relevant economic variables $E_1, \ldots, E_M$. E_1 may relate to education level, E_2 to income level, etc. The event that a woman becomes pregnant in the jth month and leaves the sample is defined under this model by $S_j < \lambda$, where $\lambda = \alpha_0 + \sum_{m=1}^{M} \alpha_m E_m$ and $\{\alpha_j\}$ are unknown parameters relating to for example, first pregnancies only. The $\{\alpha_j\}$ would presumably change for second, third, etc. pregnancies. The inequality is reversed when she does not become pregnant, and hence remains in the sample.

The probability of a woman becoming pregnant in the kth month is

$$P_r[S_1 > \lambda, \ldots, S_{k-1} > \lambda\,;\, S_k < \lambda] = p_k(\lambda) \tag{6.1.1}$$

If now there is an independent sample of such women with different birth intervals, the method of maximum likelihood in principle may be used to estimate $\alpha_0, \alpha_1, \ldots$ by choosing those parameter values which maximize the joint probability of observing the sample distribution of birth intervals. To carry out the *ML* method, it is necessary to specify a probability distribution for $S_1, \ldots, S_k$.

To put this in a somewhat more general framework, let S_i be represented as the sum of two independent random variables $S_i = U_i + \epsilon$, where $(U_1, \ldots, U_k)$ is distributed as the multivariate normal $n_k(\underline{\mu}|\underline{0}, \{\sigma_{ij}\})$ and ϵ as $n_1(\epsilon|0, \sigma_\epsilon^2)$. It then follows that $(S_1, \ldots, S_k)$ is distributed as the multivariate normal $n_k(\underline{S}|\underline{0}, \{\sigma_{ij} +$

$\sigma_\epsilon^2\}$). In particular the correlation coefficient between S_i and S_j is given by

$$\rho_{ij} = (\sigma_{ij} + \sigma_\epsilon^2)/\sqrt{(\sigma_{ii} + \sigma_\epsilon^2)(\sigma_{jj} + \sigma_\epsilon^2)}$$
$$= \hat{\sigma}_{ij}/\hat{\sigma}_i\hat{\sigma}_j$$

In this context then the probability in equation (6.1.1) would be conditional on ϵ and interest would be in the probability

$$\int_{-\infty}^{\infty} p_k(\lambda|\epsilon)n_1(\epsilon)\,d\epsilon$$

which in terms of the L_k notation takes the appearance

$$\int_{-\infty}^{\infty} L_k(x_1, \ldots, x_{k-1}, -x_k)n_1(\epsilon)\,d\epsilon \tag{6.1.2}$$

with $x_i = (\lambda - \epsilon)/\hat{\sigma}_i$ for $i = 1, \ldots, k$.

6.2 *Application #2—Multivariate Probit Model*

Consider k response systems $S_1, \ldots, S_k$ in which the reaction of system S_i is defined to be of the form

$$y_i = x_i(z) - \psi_i \qquad \text{for } i = 1, \ldots, k$$

where $x_i(z)$ is a suitable response and ψ_i is referred to as the tolerance for system S_i. In other words, if $x_i(z) > \psi_i$, a toxic effect occurs.

It is reasonable to assume that the tolerance vector $\psi = (\psi_1, \ldots, \psi_k)'$ is distributed as multivariate normal. The response functions $x_i(z)$ are so chosen that all univariate marginals associated with ψ are standardized normals.

Let $\Phi_k(x_1, \ldots, x_k) = L_k(-x_1, \ldots, -x_k)$ where the subscript is dropped for $k = 1$. Then the probabilities of quantal response (+) and non-response (−) for system S_i are

$$p_i^+(z) = \Phi[x_i(z)]$$
$$p_i^-(z) = \Phi[-x_i(z)] = 1 - \Phi[x_i(z)].$$

The probabilities that systems S_i and S_j both have positive responses is

$$p_{ij}^{++}(z) = \Phi_2[x_i(z), x_j(z)].$$

In the bivariate case the other three probabilities of interest are

$$p_{ij}^{+-}(z) = \Phi_2[x_i(z), -x_j(z)] = p_i^+(z) - p_{ij}^{++}(z)$$
$$p_{ij}^{-+}(z) = \Phi_2[-x_i(z), x_j(z)] = p_j^+(z) - p_{ij}^{++}(z)$$

and

$$p_{ij}^{--}(z) = 1 - (p_{ij}^{++} + p_{ij}^{+-} + p_{ij}^{-+}).$$

For the k dimensional case, interest is in an expression of the form

$$p_{1,\ldots,k}^{\pm,\ldots,\pm} = \Phi_k[\pm x_1(z), \ldots, \pm x_k(z)].$$

Now, let each response function $x_i(z)$ be of the form

$$x_i(z) = \beta_i' \underset{\sim}{C}_z \qquad for\ i = 1, \ldots, k$$

where $\underset{\sim}{C}_z$ is a k-dimensional vector of known constants and $\underset{\sim}{\beta}_i$ for i fixed, is a k-dimensional vector of unknown parameters. Let $r_{1,\ldots,k}^{\pm,\ldots,\pm}$ denote the number of organisms in which the systems $S_1, \ldots, S_k$ exhibit the responses $(\pm, \ldots, \pm)$. The parameter vectors $\beta_1, \ldots, \beta_k$ can then be estimated by the log likelihood function of a given set of independent samples,

$$\mathscr{L} = \sum_{\text{(all groups)}} \sum_{\substack{\text{all sets of} \\ \pm,\ldots,\pm}} r_{1,\ldots,k}^{\pm,\ldots,\pm} \log p_{1,\ldots,k}^{\pm,\ldots,\pm} + \text{constant}.$$

7. Concluding Remarks

The computational formulas based on integral transform representations over $(0, \infty)$ for multivariate normal probabilities have been summarized with specific emphasis given to properties of the D^* functions in the representations. In an appendix available on request from the author, numerical aspects of $D_2^*, \ldots, D_6^*$ are examined for four important correlation matrices identified as Equicorrelated, Markov, Toeplitz and Nested. The general curve shapes of D_k^* in these four cases suggest the possibility of obtaining simple approximations in more general cases.

A variety of numerical results, most of which were previously unavailable are given for the multivariate normal probabilities $L_2, \ldots, L_6$, in an appendix available on request from the author. Two specific applications to econometric models are noted.

8. Acknowledgement

Acknowledgement is given to Dr. Peter Cheng for helpful numerical improvements in the computer routines and to Li Chun Tao for the original computer programs.

E. M. Laboratories, Inc.
Elmsford, N.Y. 10523

References

Abramowitz, M. and I. A. Stegun (Eds.) (1964), *Handbook of Mathematical Functions*, Washington. National Bureau of Standards.

Amemiya, T. (1972), "Bivariate Probit Analysis: Minimum Chi-Square Methods," Technical Report 76. The Economics Series. Institute for Mathematical Studies in the Social Sciences. Stanford University, Stanford.

Ashford, J. R. and R. R. Sowden (1970), "Multi-variate Probit Analysis," *Biometrics* 25, 535–546.

Bacon, R. H. (1963), "Approximations to Multivariate Normal Orthant Probabilities," *Ann. Math. Statist.* 40, 152–161.

Cramer, H. (1964), *Mathematical Methods of Statistics*, Princeton University Press.

Dutt, J. E. (1973), "A Representation of Multivariate Normal Probability Integrals by Integral Transforms," *Biometrika* 60, 637–645.

Dutt, J. E., T. K. Lin and L. C. Tao (1973), "New Appproximations to Familiar Functions," *Math. Comp.* 17, 939–942.
Dutt, J. E. (1975), "On Computing the Probability Integral of a General Multivariate *t*," *Biometrika* 62, 201–205.
Dutt, J. E. and T. K. Lin (1975), "A Short Table for Computing Orthant Probabilities of Dimensions 4 and 5," *Journal of Statistical Computation and Simulation* 4, 95–120.
Dutt, J. E. and T. K. Lin (1975a), "A Table for Normal Orthant Probabilities of Dimensions 4 and 5," Math/Stat TR-2, G. D. Searle & Co., Chicago, Illinois (October).
Dutt, J. E., K. D. Mattes and L. C. Tao (1975), "Tables of the Trivariate *T* for Comparing Three Treatments to a Control with Unequal Sample Sizes," Math/Stat TR-3, G. D. Searle & Co., Chicago, Illinois (October).
Dutt, J. E., T. K. Lin and H. Desai (1976), "A Short Table of Trivariate Normal Probabilities," Math/Stat TR-4, G. D. Searle & Co., Chicago, Illinois.
Dutt, J. E. and A. P. Soms (1976), "An Integral Representation for Calculating General Multivariate Probabilities with an Application to Multivariate X^2," *Commun. Statis.* A5(4), 377–388.
Gurland, J. (1948), "Inversion Formulae for the Distribution of Ratios," *Ann. Math. Statist.* 19, 228–237
Heckman, J. J. and R. J. Willis (1973), "Contraception Strategies and Fertility Outcomes: Estimates of a Stochastic Model of Reproduction," Proceedings of American Statistical Association--Social Statistics Section.
Johnson, N. L. and S. Kotz (1972), *Distributions in Statistics: Continuous Multivariate Distributions*, New York, Wiley and Sons.
Kendall, M. G. (1941), "Proof of Relations Connected with the Tetrachoric Series and its Generalizations," *Biometrika* 32, 196–198.
McFadden, D. (1974), *Quantal Choice Analysis: A Survey*, University of California, Berkeley.
Milton, R. C. (1972), "Computer Evaluation of the Multivariate Normal Integral," *Technometrics* 14, 881–889.
Stroud, A. H. and D. Secrest (1966), *Gaussian Quadrature Formulas*, Englewood Cliffs, New Jersey. Prentice-Hall.
Tobin, J. (1955), "The Application of Multivariate Probit Analysis to Economic Survey Data," Cowles Foundation Discussion Paper No. 1.
Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, 24–36.