Biometrika (1974), **61**, 3, *p.* 615
Printed in Great Britain

An iterated integral representation for a multivariate normal integral having block covariance structure

By ROBERT E. BECHHOFER and AJIT C. TAMHANE
Department of Operations Research, Cornell University, Ithaca, New York

Summary

It is shown that a km-variate normal probability integral over a rectangular region can be expressed as an iterated k-variate normal integral when the k sets of m variates each have a certain commonly realized block covariance structure. The latter representation is much easier to evaluate numerically than is the former. This result generalizes previous results for $k = 1$ of Dunnett & Sobel and Steck & Owen.

Some key words: Block covariance structure; Multivariate normal integral; Multivariate normal probabilities; Numerical integration; Ranking and selection procedures.

1. Introduction

The problem of evaluating multivariate normal probabilities over rectangular regions has received considerable attention; see Dutt (1973) for recent relevant references. Such probabilities arise, e.g. in connexion with studies of the performance characteristics of ranking and selection procedures involving means of normal distributions. In most cases the evaluation of these probabilities is time consuming and costly, even on modern computers, and thus the implementation of ranking and selection procedures, e.g. the computation of tables to facilitate their use, has been generally inhibited. However, it is known that the equicorrelated case, which is a very important one in applications, yields iterated integrals which are particularly tractable. The purpose of the present paper is to show that similar simplifications arise if the variates have a certain block covariance structure.

Let $X' = (X_{11}, \ldots, X_{1m}, X_{21}, \ldots, X_{2m}, \ldots, X_{k1}, \ldots, X_{km})$ denote a vector consisting of k sets of m variates each. We assume that X' has a km-variate standard normal distribution with

$$\operatorname{corr}(X_{ij_1}, X_{ij_2}) = \rho_i \quad (1 \leqslant i \leqslant k;\ j_1 \neq j_2, 1 \leqslant j_1, j_2 \leqslant m),$$

$$\operatorname{corr}(X_{i_1 j}, X_{i_2 j}) = \eta_{i_1 i_2} \quad (i_1 \neq i_2, 1 \leqslant i_1, i_2 \leqslant k; 1 \leqslant j \leqslant m),$$

$$\operatorname{corr}(X_{i_1 j_1}, X_{i_2 j_2}) = \xi_{i_1 i_2} \quad (i_1 \neq i_2; 1 \leqslant i_1, i_2 \leqslant k;\ j_1 \neq j_2, 1 \leqslant j_1, j_2 \leqslant m);$$

see (6). For given constants a_{ij} and b_{ij} $(-\infty \leqslant a_{ij} \leqslant b_{ij} \leqslant \infty;\ 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant m)$ we are interested in the probability $H\{(a_{1j}, \ldots, a_{kj}), (b_{1j}, \ldots, b_{kj}); 1 \leqslant j \leqslant m\}$, where

$$H = \operatorname{pr}\{a_{ij} < X_{ij} < b_{ij} \,(1 \leqslant i \leqslant k, 1 \leqslant j \leqslant m)\}. \tag{1}$$

Now H can be found by determining the volume under a particular km-variate normal surface. In this paper we develop an equivalent iterated integral representation of (1); this latter representation is much easier to evaluate numerically than is the km-variate normal one. Our result is valid for $0 \leqslant \rho_i < 1$ $(1 \leqslant i \leqslant k)$ and under certain restrictions on the $\xi_{i_1 i_2}$ and $\eta_{i_1 i_2}$, implied by the positive-definiteness of the matrices Λ_0, Λ_1, and Ω defined by (3), and (6).

For $k = 1$, Cacoullos & Sobel (1966, p. 454) gave the general result

$$\mathrm{pr}\{a_j < X_j < b_j \ (1 \leqslant j \leqslant m)\} = \int_{-\infty}^{\infty} \prod_{j=1}^{m} \left[F\left\{\frac{b_j+\rho^{\frac{1}{2}}y}{(1-\rho)^{\frac{1}{2}}}\right\} - F\left\{\frac{a_j+\rho^{\frac{1}{2}}y}{(1-\rho)^{\frac{1}{2}}}\right\}\right] f(y)\,dy, \tag{2}$$

which they state can be shown to hold, using the proof of Steck & Owen (1962), for corr $(X_{j_1}, X_{j_2}) = \rho > -1/(m-1)$ $(j_1 \neq j_2; 1 \leqslant j_1, j_2 \leqslant m)$; here $F(.)$ is the standard normal distribution function and $f(.)$ is the corresponding density function. Earlier work on this problem was done by Dunnett & Sobel (1955) who developed (2) for the special case $a_j = -\infty$ $(1 \leqslant j \leqslant m)$, $\rho \geqslant 0$; Steck & Owen (1962) extended that result to the case $a_j = -\infty$ $(1 \leqslant j \leqslant m)$, $\rho > -1/(m-1)$ and also gave three other equivalent representations of this probability when $b_j = b$ $(1 \leqslant j \leqslant m)$.

In §3 of this paper we show that when $\eta_{ij} = \xi_{ij}$ $(i \neq j;\ 1 \leqslant i,j \leqslant k)$, a method proposed earlier by Das (1956) for reducing the size of a multivariate normal integral yields the same result as is obtained by our method.

Some situations in which the correlation matrix has the special block structure which we are considering are mentioned in §5.

2. Derivation of the iterated integral representation

Let $Y'_j = (Y_{1j}, \ldots, Y_{kj})$ $(0 \leqslant j \leqslant m)$ be a k-vector having a standard multivariate normal distribution F_j with associated correlation matrix $\Lambda_j = ((\lambda^{(j)}_{i_1 i_2}))$. We assume that the Y'_j $(0 \leqslant j \leqslant m)$ are independent, and that the Y'_j $(1 \leqslant j \leqslant m)$ are identically distributed. The nondiagonal elements of Λ_0 and Λ_1 are given for $i_1 \neq i_2$; $1 \leqslant i_1, i_2 \leqslant k$ by

$$\lambda^{(0)}_{i_1 i_2} = E(Y_{i_1 0} Y_{i_2 0}) = \frac{\xi_{i_1 i_2}}{(\rho_{i_1}\rho_{i_2})^{\frac{1}{2}}}, \tag{3a}$$

$$\lambda^{(1)}_{i_1 i_2} = E(Y_{i_1 1} Y_{i_2 1}) = \frac{\eta_{i_1 i_2} - \xi_{i_1 i_2}}{\{(1-\rho_{i_1})(1-\rho_{i_2})\}^{\frac{1}{2}}}. \tag{3b}$$

We further assume that Λ_0 and Λ_1 are positive-definite, and thus for $i_1 \neq i_2$ and $1 \leqslant i_1, i_2 \leqslant k$ we must have

$$(\xi_{i_1 i_2})^2 < \rho_{i_1}\rho_{i_2}, \quad (\eta_{i_1 i_2} - \xi_{i_1 i_2})^2 < (1-\rho_{i_1})(1-\rho_{i_2}). \tag{4}$$

In addition we assume for $1 \leqslant i \leqslant k$ that $0 \leqslant \rho_i < 1$. In the following development our proof is for the case $0 < \rho_i < 1$ $(1 \leqslant i \leqslant k)$, but the same final result would be obtained if, when one or more $\rho_i = 0$ in (5), we then define the corresponding $\lambda^{(0)}_{i_1 i_2}$ as being equal to zero in (3a).

For $k = 2$ the two conditions (4) are both necessary and sufficient that Λ_0 and Λ_1, respectively, are positive-definite. Also, if $\xi_{i_1 i_2} = \xi$, $\eta_{i_1 i_2} = \eta$, and $\rho_{i_1} = \rho$ for $(i_1 \neq i_2;\ 1 \leqslant i_1, i_2 \leqslant k)$, then Λ_0 and Λ_1 are positive-definite if

$$-1/(k-1) < \xi/\rho < 1, \quad -1/(k-1) < (\eta-\xi)/(1-\rho) < 1,$$

respectively.

We now consider the km-vector $X' = (X_1, \ldots, X_k) = (X_{11}, \ldots, X_{1m}, \ldots, X_{k1}, \ldots, X_{km})$ which is formed from the Y'_j $(0 \leqslant j \leqslant m)$ by the transformation

$$\begin{aligned} X_i &= (X_{i1}, \ldots, X_{im}) \\ &= (Y_{i1}, \ldots, Y_{im})(1-\rho_i)^{\frac{1}{2}} + (Y_{i0}, \ldots, Y_{i0})\rho_i^{\frac{1}{2}} \quad (1 \leqslant i \leqslant k). \end{aligned} \tag{5}$$

It is straightforward to show that X' has a standard multivariate normal distribution F with associated correlation matrix $\Omega = ((\Omega_{ij}))$, the elements of which are given by the $m \times m$ matrices

$$\Omega_{ii} = E(X_i' X_i) = ((\omega_{pq}^{(ii)})), \quad \omega_{pq}^{(ii)} = \begin{cases} 1 & (p = q), \\ \rho_i & (p \neq q), \end{cases} \tag{6a}$$

$$\Omega_{i_1 i_2} = E(X_{i_1}' X_{i_2}) = ((\omega_{pq}^{(i_1 i_2)})), \quad \omega_{pq}^{(i_1 i_2)} = \begin{cases} \eta_{i_1 i_2} & (p = q), \\ \xi_{i_1 i_2} & (p \neq q). \end{cases} \tag{6b}$$

Here $1 \leqslant i \leqslant k$, $1 \leqslant p, q \leqslant m$ and $i_1 \neq i_2$, $1 \leqslant i_1, i_2, \leqslant k$, $1 \leqslant p, q \leqslant m$. We also assume that Ω is positive-definite, which places additional restrictions on the $\xi_{i_1 i_2}$ and $\eta_{i_1 i_2}$. For example, for $k = 2$, $m \geqslant 2$ it can be shown that in order that $\Omega_{2m \times 2m}$ is positive-definite we must have

$$D_{m,2} = \{(1-\rho_1)(1-\rho_2) - (\eta_{12} - \xi_{12})^2\}^{m-1} \times [\{1 + (m-1)\rho_1\}\{1 + (m-1)\rho_2\} - \{m\xi_{12} + (\eta_{12} - \xi_{12})\}^2] > 0, \tag{7}$$

where $D_{m,2} = |\Omega_{2m \times 2m}|$.

For given constants a_{ij}, b_{ij} $(-\infty \leqslant a_{ij} \leqslant b_{ij} \leqslant \infty)$ $(1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant m)$ we first evaluate the probability

$$G\{(b_{1j}, \ldots, b_{kj})\ (1 \leqslant j \leqslant m)\} = \operatorname{pr}\{X_{ij} < b_{ij}\ (1 \leqslant i \leqslant k, 1 \leqslant j \leqslant m)\}. \tag{8}$$

We have, using (5), that

$$G = \operatorname{pr}\{X_i' < (b_{i1}, \ldots, b_{im}) \quad (1 \leqslant i \leqslant k)\} \tag{9}$$

$$= \operatorname{pr}[Y_j' < \{(1-\rho_1)^{-\frac{1}{2}}(b_{1j} - \rho_1^{\frac{1}{2}} Y_{10}), \ldots, (1-\rho_k)^{-\frac{1}{2}}(b_{kj} - \rho_k^{\frac{1}{2}} Y_{k0})\}\ (1 \leqslant j \leqslant m)]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{m} F_1\left\{\frac{b_{1j} - \rho_1^{\frac{1}{2}} y_{10}}{(1-\rho_1)^{\frac{1}{2}}}, \ldots, \frac{b_{kj} - \rho_k^{\frac{1}{2}} y_{k0}}{(1-\rho_k)^{\frac{1}{2}}}\right\} f_0(y_{10}, \ldots, y_{k0})\, dy_{10} \ldots dy_{k0}, \tag{10}$$

where F_j $(j = 0, 1)$ is the k-variate standard normal distribution with correlation matrix Λ_j, and f_0 is the k-variate standard normal density function corresponding to F_0.

Thus the desired probability (1) is given by

$$\begin{aligned} H = {} & G\{(b_{1j}, b_{2j}, \ldots, b_{kj})\ (1 \leqslant j \leqslant m)\} - G\{(a_{1j}, b_{2j}, \ldots, b_{kj})\ (1 \leqslant j \leqslant m)\} \\ & - G\{(b_{1j}, a_{2j}, \ldots, b_{kj})\ (1 \leqslant j \leqslant m)\} - \ldots - G\{(b_{1j}, b_{2j}, \ldots, a_{kj})\ (1 \leqslant j \leqslant m)\} \\ & + \ldots + (-1)^k G\{(a_{1j}, a_{2j}, \ldots, a_{kj})\ (1 \leqslant j \leqslant m)\}. \end{aligned} \tag{11}$$

When $k = 1$, (11) clearly reduces to (2).

3. Further simplification of the iterated integral representation

Since Λ_0 is assumed to be positive-definite we can find a nonsingular $k \times k$ matrix R such that $RR' = \Lambda_0$. Letting $Z = R^{-1} Y_0$ we note that the elements of the k-vector $Z' = (Z_1, \ldots, Z_k)$ are normally and independently distributed with zero means and unit variances. Thus the right-hand side of (10) reduces to

$$\frac{1}{(2\pi)^{\frac{1}{2}k}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{m} F_1\left\{\frac{b_{1j} - \rho_1^{\frac{1}{2}} \Sigma r_{1a} z_a}{(1-\rho_1)^{\frac{1}{2}}}, \ldots, \frac{b_{kj} - \rho_k^{\frac{1}{2}} \Sigma r_{ka} z_a}{(1-\rho_k)^{\frac{1}{2}}}\right\} \exp(-\tfrac{1}{2}\Sigma z_a^2)\, dz_1 \ldots dz_k, \tag{12}$$

sums with respect to a being over $1, \ldots, k$.

If $\eta_{i_1 i_2} = \xi_{i_1 i_2}$ ($i_1 \neq i_2$, $1 \leqslant i_1, i_2 \leqslant k$), then Λ_1 defined in (3b) becomes the identity matrix, and (11) can be written very simply, using (12), as

$$H = \frac{1}{(2\pi)^{\frac{1}{2}k}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{k} \prod_{j=1}^{m} \left[F\left\{ \frac{b_{ij} - \rho_i^{\frac{1}{2}} \Sigma r_{ia} z_a}{(1-\rho_i)^{\frac{1}{2}}} \right\} - F\left\{ \frac{a_{ij} - \rho_i^{\frac{1}{2}} \Sigma r_{ia} z_a}{(1-\rho_i)^{\frac{1}{2}}} \right\} \right] \exp\left(-\tfrac{1}{2}\Sigma z_a^2\right) dz_1, \ldots, dz_k. \tag{13}$$

In the next section we shall derive (13) using a method due to Das (1956).

4. APPLICATION OF THE METHOD OF DAS IN OUR SPECIAL CASE

Das's (1956) method as extended by Webster (1970) is as follows: Suppose that

$$X' = (X_1, \ldots, X_n)$$

is an n-vector having a standard multivariate normal distribution with correlation matrix Ω, and suppose that Ω can be expressed as $\Omega = C^2 + DD'$, where C is a $n \times n$ diagonal matrix with positive diagonal elements c_i ($1 \leqslant i \leqslant n$), and D is a $n \times k$ real matrix. Then

$$\mathrm{pr}\{a_i < X_i < b_i\,(1 \leqslant i \leqslant n)\} = \frac{1}{(2\pi)^{\frac{1}{2}k}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left\{ F\left(\frac{b_i - \Sigma d_{ia} z_a}{c_i} \right) - F\left(\frac{a_i - \Sigma d_{ia} z_a}{c_i} \right) \right\} \exp\left(-\tfrac{1}{2}\Sigma z_a^2\right) dz_1 \ldots dz_k. \tag{14}$$

In general, the difficulty in applying the method lies in finding the appropriate value of k, and the matrices C and D. However, in our special case this task is greatly simplified.

We define $Y' = (Y_{11}, \ldots, Y_{1m}, \ldots, Y_{k1}, \ldots, Y_{km})$ and $Y_0' = (Y_{10}, \ldots, Y_{k0})$, and let C be a $km \times km$ matrix and Q a $km \times k$ matrix given for $0 < \rho_r < 1$ ($1 \leqslant r \leqslant k$) by

$$c_{ij} = \begin{cases} (1-\rho_r)^{\frac{1}{2}} & (i = j = rm - m + s;\quad 1 \leqslant r \leqslant k, 1 \leqslant s \leqslant m), \\ 0 & \text{otherwise}; \end{cases}$$

$$q_{ij} = \begin{cases} \rho_j^{\frac{1}{2}} & (i = jm - m + s;\quad 1 \leqslant j \leqslant k, 1 \leqslant s \leqslant m), \\ 0 & \text{otherwise}. \end{cases}$$

Then (5) can be written as $X = CY + QY_0$. Since Y and Y_0 are independent we can write $\Omega_X = C\Omega_Y C' + Q\Omega_0 Q'$, where Ω_Y is the $km \times km$ correlation matrix of Y and Ω_0 is that for Y_0. For our special case we have $\Omega_Y = I$, and hence $\Omega_X = C^2 + Q\Omega_0 Q'$. Letting $QR = D$, where R is a $k \times k$ nonsingular matrix such that $RR' = \Omega_0$, we see that the special form of Ω applies. Therefore we can write (14) as (13).

5. APPLICATIONS

We now mention some examples involving multivariate normal integrals which have the special correlation structure studied in the present paper. This research was motivated by a similar integral which arose in a paper of Dunnett (1960, equation 2). Fairweather (1966) also encountered such an integral; one of the two $(2k-t)$-variate probabilities associated with (3.9) of his paper has this structure if k is even and $t = \frac{1}{2}k$. Both these authors were severely limited in the study of their problems due to the difficulties associated with numerical evaluation of the integrals. It is hoped that the results derived in the present paper will help ease this difficulty, e.g. Dunnett's $(2k-2)$-variate integrals can now be evaluated simply by using bivariate normal probabilities. We also remark that block covariance matrices occur commonly in multivariate data analysis, for instance in pooling of cross-sectional and time series data in econometric studies.

The work was supported by the U.S. Army Research Office and the Office of Naval Research.

References

Cacoullas, T. & Sobel, M. (1966). An inverse sampling procedure for selecting the most probable event in a multinomial distribution. In *Multivariate Analysis*, Ed. P. R. Krishnaiah, pp. 423–55. New York: Academic Press.

Das, S. C. (1956). The numerical evaluation of a class of integrals, II. *Proc. Camb. Phil. Soc.* **52**, 442–8.

Dunnett, C. W. (1960). On selecting the largest of k normal population means. *J. R. Statist. Soc.* B **22**, 1–40.

Dunnett, C. W. & Sobel, M. (1955). Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution. *Biometrika* **42**, 258–60.

Dutt, J. E. (1973). A representation of multivariate normal probability integrals by integral transforms. *Biometrika* **60**, 637–45.

Fairweather, W. R. (1966). Some extensions of Somerville's procedure for ranking means of normal populations. *Biometrika* **55**, 411–8.

Steck, G. P. & Owen, D. B. (1962). A note on the equicorrelated multivariate normal distribution. *Biometrika* **49**, 269–71.

Webster, J. T. (1970). On the application of the method of Das in evaluating a multivariate normal integral. *Biometrika* **57**, 657–60.

[*Received February* 1974. *Revised June* 1974]