Biometrika (1961), 48, 3 and 4, p. 409
Printed in Great Britain

On the evaluation of the probability integral of the multivariate t-distribution

BY S. JOHN
Indian Statistical Institute, Calcutta

1. INTRODUCTION AND SUMMARY

A random vector $t = (t_1, t_2, \ldots, t_p)$ is said to have the p-variate t-distribution with n degrees of freedom if its distribution in p-dimensional Euclidean space has for its density function the function

$$g_n(t_1, t_2, \ldots, t_p; \mathrm{P}) = \frac{|A|^{\frac{1}{2}}\Gamma(\frac{1}{2}[n+p])}{(n\pi)^{\frac{1}{2}p}\Gamma(\frac{1}{2}n)}\left[1+\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p}\alpha_{ij}t_it_j\right]^{-\frac{1}{2}(n+p)}$$
$$(-\infty < t_i < \infty;\ i = 1, 2, \ldots, p), \qquad (1)$$

with $\mathrm{P} = (\rho_{ij})$ a positive definite $p \times p$ matrix having every one of its diagonal elements equal to unity, $A = (\alpha_{ij}) = \mathrm{P}^{-1}$, and $|A|$ denoting the determinant of the matrix A. We shall introduce the symbol $G_n(h_1, h_2, \ldots, h_p; \mathrm{P})$ to denote the corresponding distribution function, i.e.

$$G_n(h_1, h_2, \ldots, h_p; \mathrm{P}) = \int_{-\infty}^{h_1}\int_{-\infty}^{h_2}\cdots\int_{-\infty}^{h_p} g_n(t_1, \ldots, t_p; \mathrm{P})\,dt_1\ldots dt_p. \qquad (2)$$

In the bivariate case, we shall write $g_n(t_1, t_2; \rho_{12})$ for $g_n(t_1, t_2; \mathrm{P})$ and $G_n(t_1, t_2; \rho_{12})$ for $G_n(t_1, t_2; \mathrm{P})$.

If the random variables $x_1, \ldots, x_p$ follow the multivariate normal law with means zero, common variance σ^2 and a correlation matrix (ρ_{ij}) and if $(ns^2)/\sigma^2$ is an independent χ^2 variable with n degrees of freedom, the random vector $t = (t_1, \ldots, t_p)$, where $t_i = x_i/s\ (i = 1, \ldots, p)$ will have $g_n(t_1, \ldots, t_p; \mathrm{P})$ as density function. Bechhofer, Dunnett & Sobel (1954) consider this distribution in connexion with a problem in the ranking of means of normal populations. Dunnett & Sobel (1954) give a formula for evaluating the probability integral when $p = 2$. Using this, they have prepared tables of the function $G_n(h, h; \pm 0{\cdot}5)$ and its inverse. Dunnett & Sobel (1955) provide approximations for $G_n(h_1, \ldots, h_p; \mathrm{P})$, valid when P satisfies certain conditions. In this paper we give an alternative formula for the evaluation of the probability integral. Though we too discuss only the bivariate case in detail, our method is of wider applicability in the sense that it can be adopted to get the probability integral of the multivariate t-distribution of any dimension.

We must mention here that the method of this paper is similar to that of Kendall (1941) for evaluating the probability integral of the multivariate normal distribution.

We have already mentioned that the multivariate t-distribution arises in the ranking of normal populations according to their means. We give more applications towards the end of this paper. It is shown how the multivariate t-distribution can be used in setting up simultaneous confidence bounds for the means of correlated normal variables. Other applications are in constructing simultaneous confidence bounds for the parameters in a linear model and for future observations from a multivariate normal distribution. Further applications will appear in a later paper.

2. The characteristic function

To derive our formula for the probability integral we require the characteristic function. For this purpose, we may, without loss of generality, put $\sigma = 1$. The characteristic function is given by

$$\phi(\theta_1, \theta_2, \ldots, \theta_p) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_0^{\infty} \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}} \exp(i\theta_1 x_1/s + i\theta_2 x_2/s + \ldots + i\theta_p x_p/s$$

$$\times \exp\left(-\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\alpha_{ij}x_i x_j\right) \frac{(ns^2)^{\frac{1}{2}n-1}}{2^{\frac{1}{2}n-1}\Gamma(\frac{1}{2}n)} e^{-\frac{1}{2}ns^2} ns\, ds\, dx_1 \ldots dx_p$$

$$= n^{\frac{1}{2}n}\int_0^{\infty} \frac{s^{n+p-1}}{2^{\frac{1}{2}n-1}\Gamma(\frac{1}{2}n)} e^{-\frac{1}{2}ns^2} ds \left[\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}} \exp(i\theta_1 z_1 + \ldots + i\theta_p z_p)\right.$$

$$\left. \times \exp\left(-\tfrac{1}{2}s^2 \sum_i \sum_j \alpha_{ij} z_i z_j\right) dz_1 \ldots dz_p\right]$$

$$= \int_0^{\infty} \frac{(ns^2)^{\frac{1}{2}n-1}}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)} e^{-\frac{1}{2}ns^2} d(ns^2) \exp\left(-\frac{1}{2s^2}\theta \mathrm{P} \theta'\right)$$

$$= \frac{1}{\Gamma(\frac{1}{2}n)}\int_0^{\infty} t^{\frac{1}{2}n-1} \exp\left(-t - \frac{n}{4t}\theta \mathrm{P}\theta'\right) dt, \tag{3}$$

where

$$\theta = (\theta_1, \ldots, \theta_p).$$

3. The probability integral

For the sake of simplicity we now restrict our attention to the case $p = 2$; exactly similar methods will give the probability integral whatever positive integral value p has.

In this case,

$$\phi(\theta_1, \theta_2) = \frac{1}{\Gamma(\frac{1}{2}n)}\int_0^{\infty} t^{\frac{1}{2}n-1} \exp\left[-t - \frac{n}{4t}(\theta_1^2 + \theta_2^2)\right]\left[\sum_{r=0}^{\infty}\frac{1}{r!}\left(-\frac{n\rho}{2t}\right)^r \theta_1^r \theta_2^r\right] dt, \tag{4}$$

where we have set $\rho = \rho_{12}$.

By the inversion theorem, the joint density function $g_n(t_1, t_2; \rho)$ of t_1 and t_2 is given by

$$g_n(t_1, t_2; \rho) = \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp(-i\theta_1 t_1 - i\theta_2 t_2)\, d\theta_1\, d\theta_2$$

$$\times \frac{1}{\Gamma(\frac{1}{2}n)}\int_0^{\infty} t^{\frac{1}{2}n-1}\exp\left[-t - \frac{n}{4t}(\theta_1^2+\theta_2^2)\right] dt \left[\sum_{r=0}^{\infty}\frac{1}{r!}\left(-\frac{n\rho}{2t}\right)^r \theta_1^r\theta_2^r\right]$$

$$= \frac{1}{\Gamma(\frac{1}{2}n)}\int_0^{\infty} e^{-t} t^{\frac{1}{2}n-1} dt \left[\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{n\rho}{2t}\right)^r \left(\frac{\partial}{\partial t_1}\right)^r \left(\frac{t}{n\pi}\right)^{\frac{1}{2}}\right.$$

$$\left.\times \exp\left(-\frac{1}{n}t_1^2 t\right)\left(\frac{\partial}{\partial t_2}\right)^r\left(\frac{t}{n\pi}\right)^{\frac{1}{2}}\exp\left(-\frac{1}{n}t_2^2 t\right)\right]. \tag{5}$$

We are now in a position to evaluate the probability integral. It is easy to see that

$$G_n(h_1, h_2; \rho) = y_{n,0}(h_1, h_2) + \frac{1}{2\pi}\sum_{r=1}^{\infty}\frac{\rho^r}{r!} y_{n,r}(h_1, h_2), \tag{6}$$

where

$$y_{n,r}(h_1, h_2) = \frac{2}{n\Gamma(\frac{1}{2}n)}\int_0^{\infty} e^{-t} t^{\frac{1}{2}n} dt \left(\frac{n}{2t}\right)^r \left[\left(\frac{\partial}{\partial h_1}\right)^{r-1}\exp\left(-\frac{1}{n}h_1^2 t\right)\right]\left[\left(\frac{\partial}{\partial h_2}\right)^{r-1}\exp\left(-\frac{1}{n}h_2^2 t\right)\right]$$

$$= \frac{1}{\Gamma(\frac{1}{2}n)}\int_0^{\infty}\exp\left[-t\left\{1+\frac{1}{n}(h_1^2+h_2^2)\right\}\right] t^{\frac{1}{2}n-1}\, dt\, H_{r-1}\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_1\right) H_{r-1}\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_2\right)$$

$$(r = 1, 2, \ldots). \tag{7}$$

Here $H_r(x)$ denotes the Hermite polynomial of degree r defined by

$$\left(-\frac{d}{dx}\right)^r e^{-\frac{1}{2}x^2} = H_r(x)\, e^{-\frac{1}{2}x^2}, \tag{8}$$

and

$$\begin{aligned} y_{n,0}(h_1,h_2) &= \frac{(n\pi)^{-1}}{\Gamma(\frac{1}{2}n)} \int_0^\infty e^{-t} t^{\frac{1}{2}n} \left[\int_{-\infty}^{h_1} \exp\left(-\frac{1}{n} t_1^2 t\right) dt_1\right] \left[\int_{-\infty}^{h_2} \exp\left(-\frac{1}{n} t_2^2 t\right) dt_2\right] dt \\ &= \frac{1}{\Gamma(\frac{1}{2}n)} \int_0^\infty e^{-t} t^{\frac{1}{2}n-1}\, dt\, G\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_1\right) G\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_2\right), \end{aligned} \tag{9}$$

where

$$G(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du. \tag{10}$$

We give below explicit expressions for $y_{n,r}(h_1,h_2)$ $(r = 1, 2, \ldots, 6)$.

If we set

$$z = 1 + (1/n)(h_1^2 + h_2^2), \tag{11}$$

$$y_{n,1}(h_1,h_2) = z^{-\frac{1}{2}n}, \tag{12}$$

$$y_{n,2}(h_1,h_2) = h_1 h_2 z^{-(\frac{1}{2}n+1)}, \tag{13}$$

$$y_{n,3}(h_1,h_2) = \{1 + (2/n)\} h_1^2 h_2^2 z^{-(\frac{1}{2}n+2)} - (h_1^2 + h_2^2)\, z^{-(\frac{1}{2}n+1)} + z^{-\frac{1}{2}n}, \tag{14}$$

$$y_{n,4}(h_1,h_2) = \left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right) h_1^3 h_2^3 z^{-(\frac{1}{2}n+3)} - 3\left(1+\frac{2}{n}\right) h_1 h_2 (h_1^2+h_2^2)\, z^{-(\frac{1}{2}n+2)} + 9 h_1 h_2 z^{-(\frac{1}{2}n+1)}, \tag{15}$$

$$\begin{aligned} y_{n,5}(h_1,h_2) &= \left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right)\left(1+\frac{6}{n}\right) h_1^4 h_2^4 z^{-(\frac{1}{2}n+4)} - 6\left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right) h_1^2 h_2^2 (h_1^2+h_2^2)\, z^{-(\frac{1}{2}n+3)} \\ &\quad + 3\left(1+\frac{2}{n}\right)(h_1^4 + 12 h_1^2 h_2^2 + h_2^4)\, z^{-(\frac{1}{2}n+2)} - 18(h_1^2+h_2^2)\, z^{-(\frac{1}{2}n+1)} + 9 z^{-\frac{1}{2}n}, \end{aligned} \tag{16}$$

$$\begin{aligned} y_{n,6}(h_1,h_2) &= \left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right)\left(1+\frac{6}{n}\right)\left(1+\frac{8}{n}\right) h_1^5 h_2^5 z^{-(\frac{1}{2}n+5)} \\ &\quad - 10\left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right)\left(1+\frac{6}{n}\right) h_1^3 h_2^3 (h_1^2+h_2^2)\, z^{-(\frac{1}{2}n+4)} \\ &\quad + 5\left(1+\frac{2}{n}\right)\left(1+\frac{4}{n}\right) h_1 h_2 (3h_1^4 + 20 h_1^2 h_2^2 + 3h_2^4)\, z^{-(\frac{1}{2}n+3)} \\ &\quad - 150\left(1+\frac{2}{n}\right) h_1 h_2 (h_1^2+h_2^2)\, z^{-(\frac{1}{2}n+2)} + 225 h_1 h_2 z^{-(\frac{1}{2}n+1)} \end{aligned} \tag{17}$$

The evaluation of $y_{n,0}(h_1,h_2)$ requires separate consideration.

4. Evaluation of $y_{n,0}(h_1, h_2)$

First we observe that

$$y_{n,0}(h_1,h_2) = G_n(h_1,h_2;\ 0). \tag{18}$$

Equation (9) can be used to evaluate $y_{n,0}(h_1,h_2)$. The integration involved has to be done numerically. Gauss's formula for numerical quadrature (Kopal, 1955, p. 371) will be found especially convenient. By this method we have prepared tables of $y_{n,0}(h_1,h_2)$ for $n = 11, 12$. The first of these, that for $n = 11$, is reproduced at the end of this paper as Table 2. The next section gives recurrence relations connecting these values of n with other values of n.

Other methods for the evaluation of $y_n(h_1,h_2)$ are given by John (1961). The table was computed on HEC-2M at the Indian Statistical Institute. The programme was prepared by

Mr P. K. Mitra, and Mr B. Mukherjee was responsible for the running of it off the computer. A similar table for the case $n = 12$ is available at the Institute.

In the tables, $y_{n,0}(h_1, h_2)$ has been tabulated only for values of h_1 and h_2 satisfying the inequalities $h_1 \geqslant 0$, $h_2 \geqslant 0$, $h_2 \geqslant h_1$. Values of $y_{n,0}(h_1, h_2)$ for h_1 or h_2 negative can be found from the formulae

$$y_{n,0}(h_1, h_2) = G_n(h_2) - y_{n,0}(-h_1, h_2), \tag{19}$$

$$y_{n,0}(h_1, h_2) = G_n(h_1) - y_{n,0}(h_1, -h_2), \tag{20}$$

where $G_n(x)$ is the probability that a random variable having Student's t-distribution with n degrees of freedom has a value less than or equal to x. If both h_1 and h_2 are negative, we may use the formula

$$y_{n,0}(h_1, h_2) = 1 + y_{n,0}(-h_1, -h_2) - G_n(-h_1) - G_n(-h_2). \tag{21}$$

Further, there is no loss of generality in assuming $h_2 \geqslant h_1$ since

$$y_{n,0}(h_1, h_2) = y_{n,0}(h_2, h_1). \tag{22}$$

5. Some useful recurrence relations

From equation (9), we get by integration by parts the following recurrence relation:

$$y_{n,0}(h_1, h_2) = y_{n+2,0}\left(\left[1+\frac{2}{n}\right]^{\frac{1}{2}} h_1, \left[1+\frac{2}{n}\right]^{\frac{1}{2}} h_2\right) - (n^3\pi)^{-\frac{1}{2}} \frac{\Gamma(\frac{1}{2}n+\frac{1}{2})}{\Gamma(\frac{1}{2}n)} \left[\frac{h_1 G_{n+1}\left(\left[\frac{n+1}{n+h_1^2}\right]^{\frac{1}{2}} h_2\right)}{\{1+(1/n)h_1^2\}^{\frac{1}{2}(n+1)}} + \frac{h_2 G_{n+1}\left(\left[\frac{n+1}{n+h_2^2}\right]^{\frac{1}{2}} h_1\right)}{\{1+(1/n)h_2^2\}^{\frac{1}{2}(n+1)}}\right]. \tag{23}$$

In getting equation (23), $t^{\frac{1}{2}n-1}$ was the factor which we selected for the first integration. If we now select the factor e^{-t} for the first integration, we get the recurrence relation given below:*

$$y_{n,0}(h_1, h_2) = y_{n-2,0}\left(\left[1-\frac{2}{n}\right]^{\frac{1}{2}} h_1, \left[1-\frac{2}{n}\right]^{\frac{1}{2}} h_2\right) + \tfrac{1}{2}(n\pi)^{-\frac{1}{2}} \frac{\Gamma(\frac{1}{2}n-\frac{1}{2})}{\Gamma(\frac{1}{2}n)} \left[\frac{h_1 G_{n-1}\left(\left[\frac{n-1}{n+h_1^2}\right]^{\frac{1}{2}} h_2\right)}{\{1+(1/n)h_1^2\}^{\frac{1}{2}(n-1)}} + \frac{h_2 G_{n-1}\left(\left[\frac{n-1}{n+h_2^2}\right]^{\frac{1}{2}} h_1\right)}{\{1+(1/n)h_2^2\}^{\frac{1}{2}(n-1)}}\right] \quad (n > 2). \tag{24}$$

By the same procedure we can express $y_{2,0}(h_1, h_2)$ entirely in terms of the probability integral of the univariate t-distribution. This relationship is given by the following equation:

$$y_{2,0}(h_1, h_2) = \tfrac{1}{4} + 8^{-\frac{1}{2}} \left[\frac{h_1 G_1([2+h_1^2]^{-\frac{1}{2}} h_2)}{(1+\frac{1}{2}h_1^2)^{\frac{1}{2}}} + \frac{h_2 G_1([2+h_2^2]^{-\frac{1}{2}} h_1)}{(1+\frac{1}{2}h_2^2)^{\frac{1}{2}}}\right]. \tag{25}$$

This, together with (24), shows that $y_{n,0}(h_1, h_2)$ for all even n can be built up from tables of the probability integral of the univariate t-distribution. If n is a small even integer this procedure is to be preferred. Similarly if n is a fairly large integer (odd or even) it may be advantageous to connect it with the probability integral of the bivariate normal distribution through the formula (23).

* Relation (24) can be derived also from (23).

5. Comparisons with the values given by Bechhofer, Dunnett and Sobel (1954)

As a check on the correctness of our formula and as a means of seeing how good an approximation is obtained by taking the sum of a few terms from the beginning in the series expansion (6), we carried out some computations. The results are presented in Table 1, where we have taken $n = 12$ and $h_1 = h_2 = h$; in the approximation, we have used the first six terms in summing the second expression on the right-hand side of equation (6). It is seen that for $\rho = \pm 0{\cdot}5$, the sum of the first seven terms of the series expansion (6) provides an approximation to the correct value which is satisfactory for most purposes. If the value of $|\rho|$ is larger, more terms have to be included in the summation to get the same accuracy. We have developed simpler methods for computing the probability integral in such cases. These methods will be published when the required tables are ready.

Table 1. *Values of* $G_{12}(h, h; \pm 0{\cdot}5)$

	Case $\rho = 0{\cdot}5$, $n = 12$		Case $\rho = -0{\cdot}5$, $n = 12$	
h	Approx. from (6)	Exact value	Approx. from (6)	Exact value
0·00	0·3333	0·33333	0·1667	0 16667
·25	·4355	·43555	·2799	·27988
·50	·5421	·54150	·4131	·41366
·75	·6429	·64292	·5488	·54880
1·00	·7330	·73301	·6694	·66936

6. The multivariate case

We have discussed in detail how the probability integral of the bivariate t-distribution may be evaluated. It is possible to get an expression similar to (6) for the probability integral of a multivariate t-distribution of any dimension by methods similar to those which were adopted in the bivariate case. The formula is the following:

$$G_n(h_1, h_2, \ldots, h_p; \mathrm{P}) = G_n(h_1, h_2, \ldots, h_p; I) + (2\pi)^{-\frac{1}{2}p} \sum_{r=1}^{\infty} \frac{1}{r!} Y_{n,r}(h_1, \ldots, h_p; \mathrm{P}), \tag{26}$$

where

$$Y_{n,r}(h_1, \ldots, h_p; \mathrm{P}) = \frac{1}{\Gamma(\frac{1}{2}n)} \int_0^{\infty} t^{\frac{1}{2}n-1} \exp\left[-t\{1 + (1/n)(h_1^2 + \ldots + h_p^2)\}\right] U_{n,r}(t; h_1, \ldots, h_p; \mathrm{P})\, dt, \tag{27}$$

and

$$U_{n,r}(t; h_1, \ldots, h_p; \mathrm{P}) = \left\{\prod_{i=1}^{p} H^{-1}\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_i\right)\right\} \left[\sum_{i<j}^{p} \rho_{ij} H\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_i\right) H\left(\left[\frac{2t}{n}\right]^{\frac{1}{2}} h_j\right)\right]^r. \tag{28}$$

It is to be understood that in (28) after expansion of the integrand, $H^r([2t/n]^{\frac{1}{2}} h_i)$ is to be replaced by $H_r([2t/n]^{\frac{1}{2}} h_i)$ if $r \geqslant 0$ and by $(2\pi)^{\frac{1}{2}} G([2t/n]^{\frac{1}{2}} h_i) \exp\{(1/n)\, t h_i^2\}$ if $r = -1$.

For many of the terms in (26), explicit expressions similar to those in equations (12) to (17) can be given; but others have to be evaluated by numerical quadrature.

7. Applications

The multivariate t-distribution arises in many statistical problems. In this section we shall describe a few of these problems.

7·1. *Simultaneous confidence bounds for the means of correlated normal variables*

Suppose $x = (x_1, \ldots, x_p)$ follows a multivariate normal distribution with mean vector $\mu = (\mu_1, \ldots, \mu_p)$ and dispersion matrix σ^2P. P is the correlation matrix and we shall suppose that this is known. The parameters μ and σ^2 are unknown. It is desired to set up simultaneous confidence bounds for the μ_i's.

This can be done as follows: draw a sample of size N from the population of x's. Let the observations be x_{ij} ($i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, N$), the first subscript pertaining to the variate and the second subscript pertaining to the sample. Let

$$\bar{x}_i = \left(\sum_{r=1}^{N} x_{ir}\right) \Big/ N \quad (i = 1, 2, \ldots, p). \tag{29}$$

Then

$$s^2 = \left[\sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_{ij} \sum_{r=1}^{N} (x_{ir} - \bar{x}_i)(x_{jr} - \bar{x}_j)\right] \Big/ \{(N-1)p\} \tag{30}$$

is an unbiased estimator of σ^2. The random variable $(N-1)ps^2/\sigma^2$ is distributed independently of the $\bar{x}_i$'s as χ^2 with $(N-1)p$ degrees of freedom. We shall set $n = (N-1)p$. Determine h so that

$$\int_{-h}^{h} \cdots \int_{-h}^{h} g_n(t_1, \ldots, t_p; \mathrm{P})\, dt_1, \ldots, dt_p = \alpha, \tag{31}$$

where α is a positive number between zero and one. Then the inequalities

$$\bar{x}_i - N^{-\frac{1}{2}} hs < \mu_i < \bar{x}_i + N^{-\frac{1}{2}} hs \quad (i = 1, 2, \ldots, p) \tag{32}$$

will be simultaneously satisfied with probability α.

The condition that all the variates have the same variance can be slightly relaxed. It is enough that the ratios of the variances are known. Thus if $V(x_i) = c_i \sigma^2$ $(i = 1, 2, \ldots, p)$ and the c_i's are known, we have only to consider $y_i = c_i^{-\frac{1}{2}} x_i$ in the place of x_i. This will yield simultaneous confidence bounds for $c_i^{-\frac{1}{2}} \mu_i$'s, which can be readily converted into simultaneous confidence bounds for the μ_i's.

We must refer here to a paper by Olive Jean Dunn (1958) discussing methods of setting simultaneous confidence bounds for the expected values of correlated normal variables. She does not require that the correlation matrix be known. She achieves her result by considering p linear combinations of the observations which have the same expectations as the original variables. The first linear combination is a linear combination of the observations on the variable x_1; the second linear combination is a linear combination of the observations on x_2 and so on. In so far as the linear combinations are not unique, this procedure is unsatisfactory.

7·2. *Simultaneous confidence bounds for future observation*

Here again we will suppose that the population is multivariate normal, that all the variates have equal variances and that the correlation matrix is known. A sample of size N is available and it is required to set up simultaneous confidence bounds for the components of a future observation vector $x = (x_1, \ldots, x_p)$.

Consider the inequalities

$$\bar{x}_i - (1 + N^{-1})^{\frac{1}{2}} hs \leqslant x_i \leqslant \bar{x}_i + (1 + N^{-1})^{\frac{1}{2}} hs \quad (i = 1, 2, \ldots, p), \tag{33}$$

$\bar{x}_i$, h and s having the same meanings as in § 7·1. The probability that these inequalities will hold simultaneously is α. The inequalities (33) will thus serve as simultaneous confidence bounds for the components of x.

As in § 7·1 the condition that all the variates should have the same variance can be replaced by the weaker condition that ratios of the variances should be known.

7·3. *Simultaneous confidence bounds for regression coefficients**

In regression analysis we have N independent normal variables $y_1, y_2, \ldots, y_N$ with common variance σ^2 and expectations given by

$$E(y_r) = \beta_0 + \sum_{i=1}^{p} \beta_i x_{ir} \quad (r = 1, 2, \ldots, N). \tag{34}$$

The x_{ij}'s are known constants. The β's are unknown parameters.

Put

$$S_{ij} = \sum_{r=1}^{N} (x_{ir} - \bar{x}_i)(x_{jr} - \bar{x}_j) \quad (i, j = 1, 2, \ldots, p), \tag{35}$$

$$S = (S_{ij}), \tag{36}$$

$$T_{yi} = \sum_{r=1}^{N} y_r(x_{ir} - \bar{x}_i) \quad (i = 1, 2, \ldots, p), \tag{37}$$

$$T_{yy} = \sum_{r=1}^{N} (y_r - \bar{y})^2. \tag{38}$$

Let $(b_1, b_2, \ldots, b_p)$ be a solution of the equations

$$\left.\begin{array}{c} S_{11}b_1 + S_{12}b_2 + \ldots + S_{1p}b_p = T_{y1}, \\ S_{21}b_1 + S_{22}b_2 + \ldots + S_{2p}b_p = T_{y2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ S_{p1}b_1 + S_{p2}b_2 + \ldots + S_{pp}b_p = T_{yp}. \end{array}\right\} \tag{39}$$

Also let

$$s^2 = \frac{1}{N-p-1}(T_{yy} - b_1T_{y1} - b_2T_{y2} - \ldots - b_pT_{yp}). \tag{40}$$

Then $(b_1, b_2, \ldots, b_p)$ have a multivariate normal distribution with means $(\beta_1, \beta_2, \ldots, \beta_p)$ and variance-covariance matrix $(c_{ij})\sigma^2 = S^{-1}\sigma^2$.

s^2 is an unbiased estimator of σ^2. $(N-p-1)s^2/\sigma^2$ has a χ^2 distribution with $(N-p-1)$ degrees of freedom and is independent of $(b_1, \ldots, b_p)$. We shall set $n = N-p-1$.

The correlation matrix (ρ_{ij}) for $(b_1, b_2, \ldots, b_p)$ is given by

$$\rho_{ij} = (c_{ii}c_{jj})^{-\frac{1}{2}} c_{ij} \tag{41}$$

and is known. Determine h so that

$$\int_{-h}^{h} \int_{-h}^{h} g_n(t_1, \ldots, t_p; \mathrm{P})\, dt_1 \ldots dt_p = \alpha. \tag{42}$$

Then the probability that the inequalities

$$b_i - hc_{ii}^{\frac{1}{2}}s \leqslant \beta_i \leqslant b_i + hc_{ii}^{\frac{1}{2}}s \quad (i = 1, 2, \ldots, p) \tag{43}$$

are simultaneously satisfied is α. Thus inequalities (43) serve as simultaneous confidence bounds for $\beta_1, \beta_2, \ldots, \beta_p$. Simultaneous confidence bounds for any set of independent linear functions of the parameters can be set up in this fashion. These bounds appear to be more

* I am indebted to Dr C. R. Rao for pointing out the possibility of this application.

relevant than Scheffé's (1953) who gives methods for setting up simultaneous confidence bounds for all possible linear functions of the parameters, standardized in a particular way. Most often we are interested in a few parametric functions and not the whole family of such functions. These functions get a bad deal in Scheffé's method; the width of the confidence interval is unnecessarily large.*

REFERENCES

BECHHOFER, R. E., DUNNETT, C. W. & SOBEL, M. (1954). A two sample multiple decision procedure for ranking means of normal populations with a common variance. *Biometrika*, **41**, 170–6.

DUNN, O. J. (1958). Estimation of the means of dependent variables. *Ann. Math. Statist.* **29**, 1095–111.

DUNNETT, C. W. & SOBEL, M. (1954). A bivariate generalization of Student's t-distribution, with tables for certain special cases. *Biometrika*, **41**, 150–69.

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.

JOHN, S. (1961). Methods for the evaluation of probabilities of polygonal and angular regions when the distribution is bivariate t. (To be published, probably in *Sankhyā*.)

KENDALL, M. G. (1941). Proof of relations connected with the tetrachoric series and its generalization. *Biometrika*, **32**, 196–8.

KOPAL, Z. (1955). *Numerical Analysis*. London: Chapman and Hall Ltd.

SCHEFFÉ, H. (1953). A method of judging all contrasts in the analysis of variance. *Biometrika*, **40**, 87–104.

Table 2. *Values of* $y_{n.0}(h_1, h_2)$ *for* $n = 11$

h_2 \ h_1	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
0·0	0·25000									
·1	·26946	0·29051								
·2	·28871	·31133	0·33369							
·3	·30755	·33170	·35558	0·37894						
·4	·32579	·35141	·37675	·40154	0·42552					
0·5	0·34326	0·37029	0·39703	0·42318	0·44848	0·47269				
·6	·35983	·38819	·41624	·44368	·47022	·49562	0·51967			
·7	·37538	·40498	·43427	·46290	·49061	·51712	·54222	0·56575		
·8	·38983	·42059	·45100	·48075	·50953	·53707	·56314	·58759	0·61027	
·9	·40314	·43495	·46641	·40717	·52603	·55541	·58238	·60766	·63111	0·65267
1·0	0·41529	0·44805	0·48045	0·51213	0·54279	0·57212	0·59990	0·62593	0·65010	0·67230
1·2	·43616	·47053	·50452	·53777	·56094	·60072	·62988	·65721	·68258	·70590
1·4	·45272	·48835	·52350	·55806	·59141	·62333	·65357	·68191	·70823	·73243
1·6	·46552	·50210	·53828	·57367	·60791	·64070	·67175	·70088	·72792	·75280
1·8	·47516	·51244	·54932	·58539	·62029	·65372	·68538	·71509	·74267	·76805
2·0	0·48229	0·52008	0·55745	0·59401	0·62940	0·66328	0·69539	0·72551	0·75349	0·77924
2·2	·48747	·52561	·56334	·60025	·63597	·67019	·70261	·73303	·76129	·78731
2·4	·49119	·52958	·56755	·60470	·64067	·67511	·70775	·73838	·76685	·79305
2·6	·49382	·53239	·57053	·60785	·64398	·67858	·71138	·74215	·77070	·79709
2·8	·49568	·53430	·57262	·61006	·64630	·68101	·71391	·74479	·77349	·79991
3·0	0·49698	0·53574	0·57408	0·61160	0·64791	0·68270	0·71567	0·74662	0·77539	0·80187
4·0	·49948	·53838	·57686	·61452	·65097	·68589	·71900	·75008	·77897	·80557
5·0	·49989	·53881	·57732	·61499	·65147	·68641	·71953	·75063	·77953	·80615
6·0	·49997	·53890	·57740	·61508	·65156	·68650	·71963	·75073	·77964	·80626
7·0	·49999	·53892	·57742	·61510	·65158	·68653	·71966	·75076	·77967	·80629
8·0	0·49999	0·53892	0·57743	0·61511	0·65159	0·68653	0·71966	0·75076	0·77967	0·80629

* Scheffé's method is easier to apply, as a referee has pointed out.

Table 2 (*cont.*)

h_2 \ h_1	0·9	1·0	1·2	1·4	1·6	1·8	2·0	2·2	2·4
0·9	0·65267								
1·0	0·67230	0·69253							
1·2	·70590	·72714	0·76351						
1·4	·73243	·75448	·79224	0·82209					
1·6	·75280	·77547	·81430	·84502	0·86864				
1·8	·76805	·79118	·83083	·86222	·88637	0·90450			
2·0	0·77924	0·80271	0·84296	0·87485	0·89939	0·91784	0·93142		
2·2	·78731	·81103	·85172	·88396	·90881	·92749	·94125	0·95122	
2·4	·79305	·81695	·85795	·89045	·91551	·93437	·94827	·95834	0·96554
2·6	·79709	·82111	·86233	·89503	·92024	·93922	·95322	·96337	·97064
2·8	·79991	·82402	·86540	·89822	·92355	·94262	·95669	·96690	·97421
3·0	0·80187	0·82604	0·86753	0·90045	0·92585	0·94499	0·95911	0·96937	0·97671
4·0	·80557	·82985	·87154	·90463	·93019	·94946	·96369	·97403	·98144
5·0	·80615	·83045	·87217	·90530	·93088	·95017	·96442	·97477	·98220
6·0	·80626	·83056	·87229	·90542	·93101	·95030	·96455	·97491	·98234
7·0	·80629	·83059	·87231	·90545	·93104	·95033	·96459	·97495	·98238
8·0	0·80629	0·83059	0·87232	0·90545	0·93104	0·95034	0·96459	0·97495	0·98238

h_2 \ h_1	2·4	2·6	2·8	3·0	4·0	5·0	6·0	7·0	8·0
2·4	0·96554								
2·6	·97064	0·97578							
2·8	·97421	·97939	0·98302						
3·0	0·97671	0·98191	0·98556	0·98812					
4·0	·98144	·98670	·99040	·99299	0·99795				
5·0	·98220	·98747	·99118	·99378	·99876	0·99958			
6·0	·98234	·98762	·99133	·99393	·99892	·99974	0·99990		
7·0	·98238	·98765	·99136	·99396	·99895	·99978	·99994	0·99998	
8·0	0·98238	0·98766	0·99137	0·99397	0·99896	0·99979	0·99995	0·99999	1·00000