CANADIAN APPLIED
MATHEMATICS QUARTERLY
Volume 13, Number 1, Spring 2005

PROBABILITY INTEGRALS OF THE MULTIVARIATE t DISTRIBUTION

SARALEES NADARAJAH AND SAMUEL KOTZ

ABSTRACT. Results on probability integrals of multivariate t distributions are reviewed. We believe that this review will serve as an important reference and encourage further research activities in the area.

1 Introduction A p-dimensional random vector

$$\mathbf{X}^T = (X_1, \ldots, X_p)$$

is said to have the t distribution with degrees of freedom ν, mean vector $\boldsymbol{\mu}$ and correlation matrix $\mathbf{R}$ if its joint pdf is given by:

$$(1) \quad f(\mathbf{x}) = \frac{\Gamma((\nu+p)/2)}{(\pi\nu)^{p/2}\Gamma(\nu/2)\left|\mathbf{R}\right|^{1/2}} \times \left[1 + \frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{R}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\nu+p)/2}.$$

The degrees of freedom parameter ν is also referred to as the shape parameter, as the peakedness of (1) may be diminished, preserved or increased by varying ν (Jensen [**31**]). The distribution is said to be central if $\boldsymbol{\mu} = \mathbf{0}$. Note that if $p = 1$, $\boldsymbol{\mu} = 0$ and $\mathbf{R} = 1$, then (1) reduces to the univariate Student's t distribution. If $p = 2$, then (1) is a slight modification of the bivariate surface of Pearson [**43**]. If $\nu = 1$, then (1) is the p-variate Cauchy distribution. If $(\nu + p)/2 = m$, an integer, then (1) is the p-variate Pearson type VII distribution. The limiting form of (1) as $\nu \to \infty$ is the joint pdf of the p-variate normal distribution with

AMS subject classification: 62E99.
Keywords: Multivariate normal distribution, multivariate t distribution, probability integrals.

mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$. The particular case of (1) for $\boldsymbol{\mu} = 0$ and $\mathbf{R} = \mathbf{I}_p$ is a mixture of the normal density with zero means and covariance matrix $v\mathbf{I}_p$—in the scale parameter v.

Multivariate t distributions are of increasing importance in classical as well as in Bayesian statistical modeling; however, relatively little is known by means of mathematical properties or statistical methods. These distributions have been perhaps unjustly overshadowed by the multivariate normal distribution. Both the multivariate t and the multivariate normal are members of the general family of elliptically symmetric distributions. However, we feel that it is desirable to focus on these distributions separately for several reasons:

- Multivariate t distributions are generalizations of the classical univariate Student t distribution, which is of central importance in statistical inference. The possible structures are numerous, and each one possesses special characteristics as far as potential and current applications are concerned.
- Application of multivariate t distributions is a very promising approach in multivariate analysis. Classical multivariate analysis is soundly and rigidly tilted toward the multivariate normal distribution while multivariate t distributions offer a more viable alternative with respect to real-world data, particularly because its tails are more realistic. We have seen recently some unexpected applications in novel areas such as cluster analysis, discriminant analysis, multiple regression, robust projection indices, and missing data imputation.
- Multivariate t distributions for the past 20 to 30 years have played a crucial role in Bayesian analysis of multivariate data. They serve by now as the most popular prior distribution (because elicitation of prior information in various physical, engineering, and financial phenomena is closely associated with multivariate t distributions) and generate meaningful posterior distributions.

There has been some amount of research carried out on probability integrals of multivariate t distributions. Most of the work was done during the pre-computer era, but recently several computer programs have been written to evaluate probability integrals. The aim of this paper is to provide a comprehensive review of the known results. We believe that this review will serve as an important reference and encourage further research activities in the area.

The paper is organized as follows. Sections 2 to 8 by now may have lost some of their usefulness but are still of substantial historical interest in addition to their mathematical value. We have decided to record these results in some detail in spite of the fact that some of the expressions are quite lengthy and cumbersome. Sections 9 to 14 contain more practically relevant and modern results.

2 Dunnett and Sobel's probability integrals One of the earliest results on probability integrals is that due to Dunnett and Sobel [**14**]. Let (X_1, X_2) have the central bivariate t distribution with degrees of freedom ν and the equicorrelation structure $r_{ij} = \rho$, $i \neq j$. The corresponding bivariate pdf is

$$f(x_1, x_2; \nu, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \left\{ 1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)} \right\}^{-(\nu+2)/2} \tag{2}$$

with the probability integral

$$P\left(y_1, y_2; \nu, \rho\right) = \int_{-\infty}^{y_2} \int_{-\infty}^{y_1} f(x_1, x_2; \nu, \rho)\, dx_1\, dx_2. \tag{3}$$

Let

$$x\left(m, y_1, y_2\right) = \frac{(y_1 - \rho y_2)^2}{(y_1 - \rho y_2)^2 + (1-\rho^2)(m + y_2^2)},$$

and let

$$I_{x(m,y_1,y_2)}(a, b) = \int_0^{x(m,y_1,y_2)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, y^{a-1}(1-y)^{b-1}\, dy$$

denote the incomplete beta function. Dunnett and Sobel [**14**] evaluated exact expressions for (3) when ν takes on positive integer values. For even ν and odd ν, they obtained

$$\begin{aligned} P(y_1, y_2; \nu, \rho) = {} & \frac{1}{2\pi} \arctan \frac{\sqrt{1-\rho^2}}{-\rho} \\ & + \frac{y_2}{4\sqrt{\nu\pi}} \sum_{j=1}^{\nu/2} \frac{\Gamma(j-1/2)}{\Gamma(j)} \left(1 + \frac{y_2^2}{\nu}\right)^{1/2-j} \\ & \times \left[1 + \operatorname{sgn}(y_1 - \rho y_2) I_{x(\nu,y_1,y_2)}\left(\frac{1}{2}, j - \frac{1}{2}\right)\right] \\ & + \frac{y_1}{4\sqrt{\nu\pi}} \sum_{j=1}^{\nu/2} \frac{\Gamma(j-1/2)}{\Gamma(j)} \left(1 + \frac{y_1^2}{\nu}\right)^{1/2-j} \\ & \times \left[1 + \operatorname{sgn}(y_2 - \rho y_1) I_{x(\nu,y_2,y_1)}\left(\frac{1}{2}, j - \frac{1}{2}\right)\right] \end{aligned} \tag{4}$$

and

$$
\begin{aligned}
(5)\quad P(y_1, y_2; \nu, \rho) = \frac{1}{2\pi} \arctan \left\{ -\sqrt{\nu} \left[\frac{\alpha\beta + \gamma\delta}{\gamma\beta - \nu\alpha\delta} \right] \right\} & \\
+ \frac{y_2}{4\sqrt{\nu\pi}} \sum_{j=1}^{(\nu-1)/2} \frac{\Gamma(j)}{\Gamma(j+1/2)} \left(1 + \frac{y_2^2}{\nu}\right)^{-j} & \\
\times \left[1 + \operatorname{sgn}(y_1 - \rho y_2) I_{x(\nu, y_1, y_2)} \left(\frac{1}{2}, j\right) \right] & \\
+ \frac{y_1}{4\sqrt{\nu\pi}} \sum_{j=1}^{(\nu-1)/2} \frac{\Gamma(j)}{\Gamma(j+1/2)} \left(1 + \frac{y_1^2}{\nu}\right)^{-j} & \\
\times \left[1 + \operatorname{sgn}(y_2 - \rho y_1) I_{x(\nu, y_2, y_1)} \left(\frac{1}{2}, j\right) \right], &
\end{aligned}
$$

respectively. Here,

$$
\alpha = y_1 + y_2, \qquad \beta = y_1 y_2 + \rho\nu, \qquad \gamma = y_1 y_2 - \nu,
$$

and

$$
\delta = \sqrt{y_1^2 - 2\rho y_1 y_2 + y_2^2 + \nu(1 - \rho^2)}\,.
$$

In the special case $y_1 = y_2 = 0$, both (4) and (5) reduce to the neat expression

$$
P(0, 0; \nu, \rho) = \arctan \frac{\sqrt{1 - \rho^2}}{-\rho}, \tag{6}
$$

which is independent of ν and is therefore identical with the corresponding result for the bivariate normal integral. Since the number of terms in (4) and (5) increases with ν, the usefulness of these expressions is confined to small values of ν. Dunnett and Sobel [**14**] also derived an asymptotic expansion in powers of $1/\nu$, the first few terms of which yield a good approximation to the probability integral even for moderately small values of ν. The method of derivation is essentially the same as that used by Fisher [**17**] to approximate the probability integral of the univariate Student's t distribution: Express the difference $f(x_1, x_2; \nu, \rho) - f(x_1, x_2; \infty, \rho)$ as a power series in $1/\nu$ and then integrate this series term by term over the desired region of integration. Setting

$$
r^2 = \frac{y_1^2 - 2\rho y_1 y_2 + y_2^2}{1 - \rho^2},
$$

Dunnett and Sobel obtained

$$\begin{aligned}\frac{f(y_1,y_2;\nu,\rho)}{f(y_1,y_2;\infty,\rho)} &= 1+\left(\frac{r^2}{4}-r^2\right)\frac{1}{\nu}+\left(\frac{r^8}{32}-\frac{5r^6}{12}+r^4\right)\frac{1}{\nu^2}\\ &\quad+\left(\frac{r^{12}}{384}-\frac{7r^{10}}{96}+\frac{13r^8}{24}-r^8\right)\frac{1}{\nu^3}\\ &\quad+\left(\frac{r^{16}}{6144}-\frac{r^{14}}{128}+\frac{17r^{12}}{144}-\frac{77r^{10}}{120}+r^8\right)\frac{1}{\nu^4}\\ &= 1+D(r),\end{aligned}$$

say. Thus, the desired probability integral is

$$\text{(7)}\quad \begin{aligned}P(y_1,y_2;\nu,\rho) &= \int_{-\infty}^{y_2}\int_{-\infty}^{y_1} f(x_1,x_2;\infty,\rho)\,dx_1\,dx_2\\ &\quad+\int_{-\infty}^{y_2}\int_{-\infty}^{y_1} D(r)f(x_1,x_2;\infty,\rho)\,dx_1\,dx_2.\end{aligned}$$

The first term on the right-hand side of (7) is the integral of the bivariate normal pdf, and it has been tabulated by Pearson [**44**] with a series of correction terms. The second term can be integrated term by term to obtain an asymptotic expansion in powers of $1/\nu$. Dunnett and Sobel gave expressions for the coefficients A_k of the terms $1/\nu^k$ for $k = 1,2,3,4$. The first of these coefficients takes the form

$$\begin{aligned}A_1 &= \frac{ay_2}{4}\phi(a)\phi(y_2)+\frac{by_1}{4}\phi(b)\phi(y_1)\\ &\quad-\frac{y_2(y_2^2+1)}{4}\phi(y_2)\,\Phi(a)-\frac{y_1(y_1^2+1)}{4}\phi(y_1)\,\Phi(b),\end{aligned}$$

where ϕ and Φ are, respectively, the pdf and the cdf of the standard normal distribution, and

$$a=\frac{y_1-\rho y_2}{\sqrt{1-\rho^2}},\qquad b=\frac{y_2-\rho y_1}{\sqrt{1-\rho^2}}.$$

In the special case $y_1 = y_2 = y$, (7) reduces to

$$\text{(8)}\quad \begin{aligned}P(y,y;\nu,\rho) &= \int_{-\infty}^{y}\int_{-\infty}^{y} f(x_1,x_2;\infty,\rho)\,dx_1\,dx_2\\ &\quad+\frac{A_1}{\nu}+\frac{A_2}{\nu^2}+\frac{A_3}{\nu^3}+\frac{A_4}{\nu^4}+\cdots,\end{aligned}$$

with the first two coefficients A_1 and A_2 now taking the forms

$$A_1 = -\frac{y\phi(y)}{2}\left\{\left(y^2+1\right)\Phi(cy) - y\Phi'(cy)\right\}$$

and

$$A_2 = -\frac{y\phi(y)}{48}\Big\{\left(3y^6 - 7y^4 - 5y^2 - 3\right)\Phi(cy) - y\Phi'(cy)\left[3y^4\left(c^4+3c^2+3\right) - y^2\left(c^2+5\right) - 3\right]\Big\},$$

where $c = \sqrt{(1-\rho)/(1+\rho)}$. In this special case, Dunnett and Sobel [**14**] tabulated numerical values of the coefficients A_k for selected values of ρ, y, and ν. The following table gives the values for $\rho = 0.5$

Coefficients of the asymptotic expansion (8) for $\rho = 0.5$

y	ν	A_1	A_2	A_3	A_4
0.25	4	−0.025870	0.003371	0.003816	−0.001050
0.50	4	−0.057784	0.008999	0.006868	−0.002155
0.75	6	−0.100016	0.021983	0.006891	−0.001879
1.00	5	−0.150182	0.047374	−0.006835	0.007991
1.25	6	−0.198378	0.079687	−0.033130	0.036817
1.50	6	−0.231628	0.096254	−0.038696	0.032808
1.75	9	−0.240531	0.067469	0.052274	−0.191482
2.00	12	−0.223682	−0.020268	0.293449	−0.819219
2.25	13	−0.187525	−0.149011	0.623867	−1.618705
2.50	12	−0.142571	−0.276255	0.858993	−1.765249
3.00	18	−0.062685	−0.376815	0.432592	2.236773

These values can be used to construct tables for the probability integral in (8).

3 Gupta and Sobel's probability integrals Gupta and Sobel [**29**] investigated the special case when **X** follows the central p-variate t distribution with degrees of freedom ν and the correlation structure $r_{ij} = \rho = 1/2$, $i \neq j$. If $Y_1, Y_2, \ldots, Y_n, Y$ are independent normal random variables with common mean and common variance σ^2, and if $\nu S^2/\sigma^2$ is a chi-squared random variable with degrees of freedom ν, independent of $Y_1, Y_2, \ldots, Y_n, Y$, then one can rewrite the probability integral as

$$P(d) = \int_{-\infty}^{d} \cdots \int_{-\infty}^{d} f(x_1, \ldots, x_p; \nu, \rho)\, dx_p \cdots dx_1 \tag{9}$$
$$= \Pr\left\{ \frac{\max(Y_1, Y_2, \ldots, Y_p) - Y}{S} \leq \sqrt{2}\, d \right\}$$
$$= \Pr\left(\frac{M_p - Y}{S} < \sqrt{2}\, d \right)$$
$$= \Pr\left(Z < \sqrt{2}\, d \right),$$

where $M_p = \max(Y_1, Y_2, \ldots, Y_p)$ and $Z = (M_p - Y)/S$. Gupta and Sobel [**29**] provided four useful expressions for $P(d)$. These are by now classical results applicable in statistical inference. The first expression is derived by fixing Y and S in (9) and integrating with respect to M_p

$$P(d) = \int_0^{\infty} h(s) \left[\int_{-\infty}^{\infty} \Phi^p(y) \phi\left(y - \sqrt{\frac{2}{\nu}}\, ds \right) dy \right] ds, \tag{10}$$

where ϕ and Φ are, respectively, the pdf and cdf of the standard normal distribution and h is the pdf of the chi-squared distribution with ν degrees of freedom. Based on the fact that the pdf ϕ admits an expansion about $d = 0$, it easy to justify a term-by-term integration of (10) to obtain the second expression

$$P(d) = \frac{1}{p+1} \sum_{k=0}^{\infty} \frac{2^{k/2} d^k}{k!} A_k E \left\{ H_k \left(\frac{\max(X_1, X_2, \ldots, X_{p+1})}{\sigma} \right) \right\},$$

where

$$A_k = E\left\{ \left(\frac{S}{\sigma} \right)^k \right\} = \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\left(\frac{\nu}{2}\right)^{k/2} \Gamma\left(\frac{\nu}{2}\right)} \tag{11}$$

is the kth moment of $\chi_\nu/\sqrt{\nu}$ (provided that $k > -\nu$) and H_k is the kth Hermite polynomial defined by

$$\left(-\frac{d}{dx} \right)^k \exp\left(-\frac{x^2}{2} \right) = H_k(x) \exp\left(-\frac{x^2}{2} \right). \tag{12}$$

A third expression for $P(d)$ is derived by first expanding ϕ about $S=\sigma$ and then integrating term by term, obtaining

$$\begin{aligned}
P(d) &= \int_{-\infty}^{\infty} \Phi^p(y)\Bigg[\phi\left(y-\sqrt{2}\,d\right) - \sqrt{2}\,d\phi^{(1)}\left(y-\sqrt{2}\,d\right) E\left(\frac{S}{\sigma}-1\right) \\
&\qquad + d^2\phi^{(2)}\left(y-\sqrt{2}\,d\right) E\left(\frac{S}{\sigma}-1\right)^2 - \cdots\Bigg] dy \\
&= \int_{-\infty}^{\infty} \Phi^p(y)\phi\left(y-\sqrt{2}\,d\right) dy - \sqrt{2}\,d\left(1-A_1\right) \\
&\qquad \times \int_{-\infty}^{\infty} \left(y-\sqrt{2}\,d\right) \Phi^p(y)\phi\left(y-\sqrt{2}\,d\right) dy \\
&\qquad + 2d^2\left(1-A_1\right) \int_{-\infty}^{\infty} \left\{y^2 - 2\sqrt{2}\,dy + 2d^2 - 1\right\} \\
&\qquad \times \Phi^p(y)\phi\left(y-\sqrt{2}\,d\right) dy + \cdots,
\end{aligned}$$

where A_1 is given by (11). Each of the integrals above can be evaluated by expanding the pdf ϕ about $d=0$, as was done in (10). The fourth and final expression for $P(d)$ given by Gupta and Sobel [**29**] uses the result of Seal [**47**] that the distribution of $D=(M_p-Y)/\sigma$ is asymptotically normal as p tends to infinity. It follows directly from Seal's result that the third and higher central moments of D tend to the corresponding moments of the standard normal distribution. Since the coefficients involving ν in A_{-k} in (11) tend to unity as $\nu\to\infty$, it follows that the third and higher central moments of $Z=(M_p-Y)/S$ tend to the corresponding moments of the standard normal distribution as both ν and p tend to infinity. It is therefore reasonable to approximate the distribution of $W=(Z-E(Z))/\sqrt{Var(Z)}$ by a Gram-Charlier expansion in the Edgeworth form, where

$$E(Z) = A_{-1}a_{p,1}$$

and

$$Var(Z) = A_{-2}\left(a_{p,2}+1\right) - \left(A_{-1}a_{p,1}\right)^2.$$

Here, $a_{p,i}$ denotes the ith moment of the largest of p independent standard normal random variables. Using equation (17.7.3) of Cramér [**12**] and letting $d_s = (\sqrt{2}d - E(Z))/\sqrt{Var(Z)}$, Gupta and Sobel obtained

$$\begin{aligned}
P(d) &= \Pr\left(Z < \sqrt{2}d\right) \\
&= \Pr\left(W < d_s\right) \\
&= \Phi\left(d_s\right) - \frac{\alpha_3}{3!}\phi^{(2)}\left(d_s\right) + \frac{\alpha_4}{4!}\phi^{(3)}\left(d_s\right) + \frac{10\alpha_3^2}{6!}\phi^{(5)}\left(d_s\right) \\
&\qquad - \frac{\alpha_5}{5!}\phi^{(4)}\left(d_s\right) - \frac{35\alpha_3\alpha_4}{7!}\phi^{(6)}\left(d_s\right) - \frac{280\alpha_3^3}{9!}\phi^{(8)}\left(d_s\right) + \cdots,
\end{aligned}$$

where

$$\alpha_k = \frac{\kappa_k}{\sqrt{\kappa_2}} \tag{13}$$

is the kth standardized cumulant of Z obtained from the moments around the origin.

In a related development, Gupta [**27**] studied the above case $\rho = 1/2$ and showed that $P(d) = P(d;\nu)$ satisfies

$$\frac{dP\left(d;\nu\right)}{dd} + \nu\left\{P(d;\nu) - P(d;\nu+2)\right\} = 0, \tag{14}$$

which is Hartley's differential-difference equation for the probability integral of a general class of statistics known as Studentized statistics. Using Hartley's solution (obtained using the theory of characteristics), Gupta obtained an approximation for $P(d;\nu)$ in powers of $1/\nu$ and remarked that it can be computed by using the Gauss-Hermite quadrature. Gupta *et al.* [**28**] extended this result for any $\rho > 0$ and showed that $P(d)$ satisfies (14) in this case too. In this case the approximation for $P(d)$ in powers of $1/\nu$ is

$$P(d) = G(d, \ldots, d) + \sum_{k=1}^{m} L_k(d), \tag{15}$$

where L_k is the kth correction term and G is the joint cdf of a p-variate normal distribution with zero means, common variance σ^2, and the equicorrelation structure $r_{ij} = \rho$, $i \neq j$. Letting $G^{(k)}(d)$ denote the kth-order derivative of $G(d, \ldots, d)$ with respect to d, the first four correction terms can be written as

$$L_1(d) = \frac{1}{d}\left\{\alpha^{(2)} - \alpha^{(1)}\right\},$$

$$L_2(d) = \frac{1}{6\nu^2}\left\{3\alpha^{(4)} - 10\alpha^{(3)} + 9\alpha^{(2)} - 2\alpha^{(1)}\right\},$$

$$L_3(d) = \frac{1}{6\nu^3}\left\{\alpha^{(6)} - 7\alpha^{(5)} + 17\alpha^{(4)} - 17\alpha^{(3)} + 6\alpha^{(2)}\right\},$$

and

$$\begin{aligned} L_4(d) = \frac{1}{360\nu^4}\big\{&15\alpha^{(8)} - 180\alpha^{(7)} + 830\alpha^{(6)} - 1848\alpha^{(5)} \\ &+ 2015\alpha^{(4)} - 900\alpha^{(3)} + 20\alpha^{(2)} + 48\alpha^{(1)}\big\}, \end{aligned}$$

where

$$\alpha^{(k)} = \frac{1}{2^k}\varphi^{(k)}(d), \quad k = 1, 2, \ldots, 8,$$

and the first eight $\varphi^{(k)}(d)$ are

$$\begin{aligned}
\varphi^{(1)}(d) &= dG^{(1)}(d), \\
\varphi^{(2)}(d) &= d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(3)}(d) &= d^3G^{(3)}(d) + 3d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(4)}(d) &= d^4G^{(4)}(d) + 6d^3G^{(3)}(d) + 7d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(5)}(d) &= d^5G^{(5)}(d) + 10d^4G^{(4)}(d) + 25d^3G^{(3)}(d) \\
&\quad + 15d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(6)}(d) &= d^6G^{(6)}(d) + 15d^5G^{(5)}(d) + 65d^4G^{(4)}(d) \\
&\quad + 90d^3G^{(3)}(d) + 31d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(7)}(d) &= d^7G^{(7)}(d) + 21d^6G^{(6)}(d) + 140d^5G^{(5)}(d) \\
&\quad + 350d^4G^{(4)}(d) + 301d^3G^{(3)}(d) \\
&\qquad + 63d^2G^{(2)}(d) + dG^{(1)}(d), \\
\varphi^{(8)}(d) &= d^8G^{(8)}(d) + 28d^7G^{(7)}(d) + 266d^6G^{(6)}(d) \\
&\quad + 1050d^5G^{(5)}(d) + 1701d^4G^{(4)}(d) + 966d^3G^{(3)}(d) \\
&\qquad + 127d^2G^{(2)}(d) + dG^{(1)}(d).
\end{aligned} \tag{16}$$

Thus the evaluation of $P(d)$ in (15) involves that of $G^{(k)}$ for $k = 0, 1, \ldots, 8$.

4 John's probability integrals John [**33**] provided alternative formulas for the evaluation of the probability integral. Although the method is discussed in detail only for the bivariate case, it has wider applicability in the sense that it can be adopted to obtain the probability integral of the multivariate t distribution for any dimension.

Let $\mathbf{X}$ be a p-variate vector having the central t distribution with degrees of freedom ν and correlation matrix $\mathbf{R}$. Using the definition that $\mathbf{X}$ can be represented as $(Z_1/S, Z_2/S, \ldots, Z_p/S)$, where $\mathbf{Z}$ is a p-variate normal random vector with correlation matrix $\mathbf{R}$ and $\nu S^2/\sigma^2$ is an independent chi-squared random variable with degrees of freedom ν, one can show that the characteristic function of $\mathbf{X}$ is

$$
\begin{aligned}
E\left(\exp\left(i\mathbf{t}^T\mathbf{X}\right)\right) &= E\left(E\left(\exp(i\mathbf{t}^T\mathbf{Z}/s \mid S = s)\right)\right) \\
&= \frac{1}{\Gamma(\nu/2)} \int_0^\infty x^{\nu/2-1} \exp\left(-x - \frac{\nu}{4x}\mathbf{t}^T\mathbf{R}^{-1}\mathbf{t}\right) dx.
\end{aligned}
$$

In the case $p = 2$ with the equicorrelation structure $r_{ij} = \rho$, $i \neq j$, the above expression reduces to

$$
\begin{aligned}
E\left(\exp\left(it_1X_1 + it_2X_2\right)\right) &= \frac{1}{\Gamma(\nu/2)} \int_0^\infty x^{\nu/2-1} \\
&\quad \times \left\{ \sum_{i=0}^{\infty} \frac{1}{i!} \left(-\frac{\nu\rho}{2x}\right)^i t_1^i\, t_2^i \right\} \exp\left\{ -x - \frac{\nu\left(t_1^2 + t_2^2\right)}{4x} \right\} dx.
\end{aligned}
$$

By the inversion theorem, John [**33**] derived the corresponding joint pdf as an infinite series of one-dimensional integrals. Integrating the infinite series term by term, the probability integral becomes

$$
P\left(y_1, y_2; \nu, \rho\right) = y_{\nu,0}\left(y_1, y_2\right) + \frac{1}{2\pi} \sum_{i=1}^{\infty} \frac{\rho^i}{i!}\, y_{\nu,i}\left(y_1, y_2\right),
$$

where

$$
y_{\nu,0}\left(y_1, y_2\right) = \frac{1}{\Gamma(\nu/2)} \int_0^\infty x^{\nu/2-1} \exp(-x)\, \Phi\left(\frac{\sqrt{2x}\, y_1}{\sqrt{\nu}}\right) \Phi\left(\frac{\sqrt{2x}\, y_2}{\sqrt{\nu}}\right) dx
$$

and

$$y_{\nu,i}\left(y_1,y_2\right)=\frac{1}{\Gamma(\nu/2)}\int_0^\infty x^{\nu/2-1}\exp\left[-x\left\{1+\frac{y_1^2+y_2^2}{\nu}\right\}\right] \times H_{i-1}\left(\frac{\sqrt{2x}\,y_1}{\sqrt{\nu}}\right)H_{i-1}\left(\frac{\sqrt{2x}\,y_2}{\sqrt{\nu}}\right)dx$$

for $i=1,2,\ldots$. Here, $\Phi(\cdot)$ is the cdf of the standard normal distribution and H_k denotes the Hermite polynomial of order k defined by (12). John provided explicit algebraic expressions for $y_{\nu,i}$ for $i=1,2,\ldots,6$. The first three of them are

$$y_{\nu,1}\left(y_1,y_2\right)=z^{-\nu/2},$$

$$y_{\nu,2}\left(y_1,y_2\right)=y_1\,y_2\,z^{-(\nu/2+1)},$$

and

$$y_{\nu,3}\left(y_1,y_2\right)=\left(1+\frac{2}{\nu}\right)y_1^2\,y_2^2\,z^{-(\nu/2+2)} -\left(y_1^2+y_2^2\right)z^{-(\nu/2+1)}+z^{-\nu/2},$$

where $z=(y_1^2+y_2^2)/\nu+1$. In principle, explicit expressions for $y_{\nu,i}$ can be obtained for any $i\geq 1$. To evaluate $y_{\nu,0}$, the integration has to be done numerically. John tabulated values of this quantity for $\nu=11,12$ using Gauss' formula for a numerical quadrature (Kopal [**36**], p. 371). He also provided several useful recursion relations. For example, values of $y_{\nu,0}(y_1,y_2)$ for y_1 negative or y_2 negative or both negative can be found from the formulas

$$y_{\nu,0}\left(y_1,y_2\right)=T_\nu\left(y_2\right)-y_{\nu,0}\left(-y_1,y_2\right),$$

$$y_{\nu,0}\left(y_1,y_2\right)=T_\nu\left(y_1\right)-y_{\nu,0}\left(y_1,-y_2\right),$$

and

$$y_{\nu,0}\left(y_1,y_2\right)=1+y_{\nu,0}\left(-y_1,-y_2\right)-T_\nu\left(-y_1\right)-T_\nu\left(-y_2\right),$$

where T_ν is the cdf of the Student's t distribution with ν degrees of freedom.

5 Amos and Bulgren's probability integrals In a widely quoted paper, Amos and Bulgren [**4**] derived several representations for (3) in terms of series and simple one-dimensional quadratures, together with efficient computational procedures for the special functions used in their numerical evaluation. One of the quadrature formulas given is

$$\begin{aligned} P = {} & \frac{1}{2\pi(\nu+1)(1+\gamma_1^2+\gamma_2^2)^{\nu/2}} \\ & \qquad \times \int_{\theta_1}^{\theta_2} {}_2F_1\left(1, \frac{\nu}{2}; \frac{\nu+3}{2}; 1 - c^2\cos^2(\theta-\phi)\right) d\theta \\ & \qquad - \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi}\Gamma(\nu/2)(1+\gamma_1^2+\gamma_2^2)^{\nu/2}} \\ & \qquad \times \int_{\theta_1}^{\theta_2} \frac{I\left\{\cos(\theta-\phi)<0\right\}\cos(\theta-\phi)}{\left\{1-c^2\cos^2(\theta-\phi)\right\}^{(\nu+1)/2}}\, d\theta, \end{aligned}$$

where ${}_2F_1$ is the Gauss hypergeometric function, $I\{\}$ is the indicator function,

$$c = \sqrt{\frac{\gamma_1^2+\gamma_2^2}{1+\gamma_1^2+\gamma_2^2}},$$

$$\gamma_1 = (y_2+y_1)\sqrt{\frac{\lambda_1}{2\nu}},$$

$$\gamma_2 = (y_2-y_1)\sqrt{\frac{\lambda_2}{2\nu}},$$

$$\theta_1 = \pi - \arctan\sqrt{\frac{1+\rho}{1-\rho}},$$

$$\theta_2 = \pi + \arctan\sqrt{\frac{1+\rho}{1-\rho}},$$

$$\phi = \begin{cases} \arctan\left(\gamma_2/\gamma_1\right), & \text{if } \gamma_1 > 0, \\ \pi + \arctan\left(\gamma_2/\gamma_1\right), & \text{if } \gamma_1 < 0, \end{cases}$$

$$\lambda_1 = \frac{1}{1+\rho}, \qquad \text{and} \qquad \lambda_2 = \frac{1}{1-\rho}.$$

One of the series formulas given is

$$
\begin{aligned}
(17) \quad P = {} & \frac{1}{2\sqrt{\pi}\Gamma(\nu/2)} \\
& \times \sum_{k=0}^{\infty} \frac{(-c)^k}{(1+\gamma_1^2+\gamma_2^2)^{\nu/2}} \frac{\Gamma((\nu+k)/2)}{\Gamma((1+k)/2)} \int_{\theta_1}^{\theta_2} \cos^k(\theta-\phi)\, d\theta.
\end{aligned}
$$

For the special case $\nu = 1$, P can be reduced to the closed-form expression

$$
P = \frac{1}{\pi} \arctan\left(\frac{2v}{u^2+v^2-1}\right) + I\left\{u^2+v^2<1\right\},
$$

where

$$
u = \frac{2r\sin\phi}{A\left(1+r^2+2r\cos\phi\right)},
$$

$$
v = \frac{1-r^2}{A\left(1+r^2+2r\cos\phi\right)},
$$

$$
r = \frac{\sqrt{\gamma_1^2+\gamma_2^2}}{1+\sqrt{1+\gamma_1^2+\gamma_2^2}},
$$

and

$$
A = \tan\left(\frac{\theta_2-\pi}{2}\right).
$$

If in addition $\rho = 0$, then the expression for P reduces further to

$$
P = \frac{1}{2\pi}\left\{\arctan\left(\frac{y_1y_2}{\sqrt{1+y_1^2+y_2^2}}\right) + \arctan y_1 + \arctan y_2 + \frac{\pi}{2}\right\}.
$$

The advantage of these expressions over the ones given by Dunnett and Sobel [**14**] is that these are easier to compute, especially for large degrees of freedom. For instance, the integral in θ in (17) can be expressed in terms of incomplete beta functions that are extensively tabulated. Amos and Bulgren [**4**] numerically evaluated values of P for all combinations of $\rho = -0.9, -0.5, 0, 0.5, 0.9$ and $\nu = 1, 2, 5, 10, 25, 50$.

6 Steffens' noncentral probabilities Consider the p-variate noncentral t distribution defined by

$$f(\mathbf{x}) = \exp\left\{-\frac{1}{2}\boldsymbol{\xi}^T \mathbf{R} \boldsymbol{\xi}\right\} \frac{\Gamma\left((\nu+p)/2\right)}{(\nu\pi)^{p/2}\Gamma(\nu/2)\left|\mathbf{R}\right|^{1/2}} \times \left\{1 + \frac{1}{\nu}\mathbf{x}^T \mathbf{R}^{-1}\mathbf{x}\right\}^{-(\nu+p)/2} \times \sum_{k=0}^{\infty} \frac{\Gamma\left((\nu+p+k)/2\right)}{k!\Gamma\left((\nu+p)/2\right)} \left\{\frac{\sqrt{2}\,\mathbf{x}^T \mathbf{R}^{-1}\boldsymbol{\xi}}{\sqrt{\nu + \mathbf{x}^T \mathbf{R}^{-1}\mathbf{x}}}\right\}^k, \tag{18}$$

where $\boldsymbol{\xi} = \boldsymbol{\mu}/\sigma$ (this reduces to the standard p-variate t distribution when $\boldsymbol{\mu} = \mathbf{0}$). Motivated by the Studentized maximum and minimum modulus tests, Steffens [**61**] studied the particular case for $p = 2$ and $\mathbf{R} = \mathbf{I}_p$. In this case, the joint pdf (18) reduces to

$$f(x_1, x_2) = \exp\left(-\frac{\xi_1^2 + \xi_2^2}{2}\right) \frac{1}{\pi\Gamma(\nu/2)} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{\Gamma\left((\nu+k+l)/2+1\right)}{k!\,l!\,\nu^{(k+l)/2+1}} \times \left(\sqrt{2}\xi_1 x_1\right)^k \left(\sqrt{2}\xi_2 x_2\right)^l \left(1 + \frac{x_1^2}{\nu} + \frac{x_2^2}{\nu}\right)^{(\nu+k+l+2)/2},$$

where $\xi_j = \mu_j/\sigma$ are the noncentrality parameters and ν denotes the degrees of freedom. The testing procedures involve maximum or minimum values of the components X_1 and X_2 and the computation of the corresponding probabilities. For this reason, Steffens [**61**] derived series representations for probabilities of the form

$$P_1 = \Pr(\mid X_1 \mid \leq A, \mid X_2 \mid \leq A) \quad \text{and} \quad P_2 = \Pr(\mid X_1 \mid > A, \mid X_2 > A).$$

It is seen that

$$P_1 = 2\exp\left(-\frac{\xi_1^2 + \xi_2^2}{2}\right) \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{\left(\xi_1^2/2\right)^k \left(\xi_2^2/2\right)^l}{k!\,l!\,B\left(k+1/2, l+1/2\right)} \times \int_0^{\pi/4} \left(\sin^{2k} v \cos^{2l} v + \sin^{2l} v \cos^{2k} v\right) I_\alpha\left(k+l+1, \frac{\nu}{2}\right) dv$$

and

$$P_2 = 2\exp\left(-\frac{\xi_1^2+\xi_2^2}{2}\right)\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\left(\xi_1^2/2\right)^k\left(\xi_2^2/2\right)^l}{k!\,l!\,B\left(k+1/2,l+1/2\right)}$$
$$\times\int_0^{\pi/4}\left(\sin^{2k}v\cos^{2l}v+\sin^{2l}v\cos^{2k}v\right)\left\{1-I_\beta\left(k+l+1,\frac{\nu}{2}\right)\right\}dv,$$

where I_x denotes the incomplete beta function ratio

$$\alpha = \frac{A^2\sec^2 v}{\nu + A^2\sec^2 v} \quad \text{and} \quad \beta = \frac{A^2\mathrm{cosec}^2 v}{\nu + A^2\mathrm{cosec}^2 v}.$$

Using these representations, Steffens estimated values of the critical points A for all combinations of $\nu = 1, 2, 5, 10, 20, 50, \infty$ and $\xi_1, \xi_2 = 0(1)5$ for the significance level 0.05. In a more recent development, Bohrer *et al.* [**8**] developed a flexible algorithm to compute probabilities of the form $\Pr(c_{11} \leq X_p \leq c_{21}, \ldots, c_{1p} \leq X_p \leq c_{2p})$ associated with the noncentral p-variate distribution (18).

7 Dutt's probability integrals Dutt [**16**] obtained a Fourier transform representation for the probability integral of a central p-variate t distribution with degrees of freedom ν and correlation matrix $\mathbf{R}$

$$P\left(y_1,\ldots,y_p\right) = \int_{-\infty}^{y_1}\cdots\int_{-\infty}^{y_p} f\left(x_1,\ldots,x_p;\nu\right)dx_p\cdots dx_1. \tag{19}$$

Using the definition of multivariate t, one can rewrite (19) as

$$\begin{aligned} P\left(y_1,\ldots,y_p\right) &\\ &= \frac{2}{2^{\nu/2}\Gamma(\nu/2)}\int_0^{\infty} z^{\nu-1}\exp\left(-z^2/2\right)G\left(\widehat{y}_1,\ldots,\widehat{y}_p\right)dz, \end{aligned} \tag{20}$$

where $\widehat{y}_k = y_k z/\sqrt{\nu}$, $k = 1, \ldots, p$ and G is the joint cdf of the multivariate normal distribution with zero means and correlation matrix $\mathbf{R}$. In the case $y_k = 0$, one has P independent of ν and

$$P\left(y_1,\ldots,y_p\right) = G\left(0,\ldots,0\right).$$

Explicit forms of G for $p = 2, 3, 4$ in terms of the D-functions are given in Dutt [**15**]. The D-functions are integral forms over $(-\infty, \infty)$ defined by

$$D_k\left(t_1, \ldots, t_p; \mathbf{R}\right) = \frac{\left|i^k\right|}{(2\pi)^k} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{d_k}{s_1 \cdots s_k} \times \exp\left(i \sum_{l=0}^{k} t_l s_l - \sum_{l=0}^{k} s_l^2 \Big/ 2\right) ds_k \cdots ds_1,$$

where the first five d_k are

$$\begin{aligned}
d_1 &= 1, \\
d_2 &= d_{12}, \\
d_3 &= d_{12+13+23} - \left(d_{12} + d_{13} + d_{23}\right), \\
d_4 &= -d_{12+13+23+14+24+34} + d_{12+13+23} \\
&\quad + d_{12+14+24} + d_{13+14+34} + d_{23+24+34} \\
&\quad - \left(d_{12} + d_{13} + d_{23} + d_{14} + d_{24} + d_{34}\right), \\
d_5 &= -d_{12+13+23+24+34+15+25+35+45} + d_{12+13+23+14+24+34} \\
&\quad + d_{12+13+23+15+25+35} + d_{12+14+24+15+25+45} \\
&\quad + d_{13+14+34+15+35+45} + d_{23+24+34+25+35+45} \\
&\quad - \big(d_{12+13+23} + d_{12+14+24} + d_{12+15+25} + d_{13+14+34} \\
&\quad + d_{13+15+35} + d_{14+15+45} + d_{23+24+34} + d_{23+25+35} \\
&\quad + d_{24+25+45} + d_{34+35+45}\big) + d_{12} + d_{13} + \cdots + d_{45},
\end{aligned}$$

and

$$d_{p_1 q_1 + \cdots + p_m q_m} = 1 - \exp\left\{-\left(r_{p_1 q_1} s_{p_1} s_{q_1} + \cdots + r_{p_m q_m} s_{p_m} s_{q_m}\right)\right\}.$$

Using the notation

$$D_{k:j_1,\ldots,j_k} = D_k\left\{t_{j_1}, \ldots, t_{j_k}; \mathbf{R}\left(t_{j_1}, \ldots, t_{j_k}\right)\right\},$$

where $\mathbf{R}\left(t_{j_1}, \ldots, t_{j_k}\right)$ is the correlation matrix based on the subscripts $j_1, \ldots, j_k$, Dutt [**15**] provided the following explicit forms for G

$$
\begin{aligned}
G\left(t_1, t_2\right) &= \left\{1-\Phi\left(t_1\right)\right\}\left\{1-\Phi\left(t_2\right)\right\}+D_{2:1,2}, \\
G\left(t_1, t_2, t_3\right) &= \left\{1-\Phi\left(t_1\right)\right\}\left\{1-\Phi\left(t_2\right)\right\}\left\{1-\Phi\left(t_3\right)\right\} \\
&\qquad +\left\{1-\Phi\left(t_1\right)\right\} D_{2:2,3}+\left\{1-\Phi\left(t_2\right)\right\} D_{2:1,3} \\
&\qquad +\left\{1-\Phi\left(t_3\right)\right\} D_{2:1,2}+D_{3:1,2,3},
\end{aligned}
$$

and

$$
\begin{aligned}
G\left(t_1, t_2, t_3, t_4\right) = \prod_{k=1}^{4} & \left\{1-\Phi\left(t_k\right)\right\}+\left\{1-\Phi\left(t_1\right)\right\}\left\{1-\Phi\left(t_2\right)\right\} D_{2:3,4} \\
& +\left\{1-\Phi\left(t_1\right)\right\}\left\{1-\Phi\left(t_3\right)\right\} D_{2:2,4} \\
& +\left\{1-\Phi\left(t_2\right)\right\}\left\{1-\Phi\left(t_3\right)\right\} D_{2:1,4} \\
& +\left\{1-\Phi\left(t_1\right)\right\}\left\{1-\Phi\left(t_4\right)\right\} D_{2:2,3} \\
& +\left\{1-\Phi\left(t_2\right)\right\}\left\{1-\Phi\left(t_4\right)\right\} D_{2:1,3} \\
& +\left\{1-\Phi\left(t_3\right)\right\}\left\{1-\Phi\left(t_4\right)\right\} D_{2:1,2} \\
& +\left\{1-\Phi\left(t_1\right)\right\} D_{3:2,3,4}\left\{1-\Phi\left(t_2\right)\right\} D_{3:1,3,4} \\
& +\left\{1-\Phi\left(t_3\right)\right\} D_{3:1,2,4}+\left\{1-\Phi\left(t_4\right)\right\} D_{3:1,2,3} \\
& \quad +D_{4:1,2,3,4}.
\end{aligned}
$$

A much simplified representation for G in terms of the error function, $\operatorname{erf}(\cdot)$, and integral forms over $(0, \infty)$, denoted as the D^*-functions, is given in a later paper by Dutt [**16**]. These D^*-functions are defined by

$$
\begin{aligned}
(21) \quad D_k^*\left(t_1, \ldots, t_p ; \mathbf{R}\right) = \frac{2}{(2\pi)^k} \int_0^{\infty} \cdots \int_0^{\infty} & \frac{d_k^*}{s_1 \cdots s_k} \\
& \times \exp\left(-\sum_{l=0}^{k} s_l^2 \Big/ 2\right) ds_k \cdots ds_1,
\end{aligned}
$$

where for the first few k are

$$
\begin{aligned}
d_1^* &= \sin\left(t_1 s_1\right), \\
d_2^* &= e_{-12} \cos_{1-2} - e_{12} \cos_{1+2}, \\
d_3^* &= e_{12+13+23+14+_24+34} \cos_{1+2+3+4} \\
&\quad + e_{12-13-23-14-24+34} \cos_{-1-2+3+4} \\
&\quad + e_{-12+13-23-14+24-34} \cos_{-1+2-3+4} \\
&\quad + e_{-12-13+23+14-24-34} \cos_{1-2-3+4} \\
&\quad - e_{-12-13+23-14+24+34} \cos_{-1+2+3+4} \\
&\quad - e_{-12+13-23+14-24+34} \cos_{1-2+3+4} \\
&\quad - e_{12-13-23+14+24-34} \cos_{1+2+3+4} \\
&\quad - e_{12+13+23-14-24-34} \cos_{1+2+3-4}
\end{aligned}
$$

and for notation

$$
\begin{aligned}
e_{p_1q_1+\cdots+p_mq_m} &= \exp\left\{-\left(r_{p_1q_1} s_{p_1} s_{q_1} + \cdots + r_{p_mq_m} s_{p_m} s_{q_m}\right)\right\}, \\
\sin_{p_1+\cdots+p_m} &= \sin\left(t_{p_1} s_{p_1} + \cdots + t_{p_m} s_{p_m}\right), \\
\cos_{p_1+\cdots+p_m} &= \cos\left(t_{p_1} s_{p_1} + \cdots + t_{p_m} s_{p_m}\right).
\end{aligned}
$$

(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 $-t_{p_1} s_{p_1}$.) Important special cases of these functions are

$$
\begin{aligned}
D_1^*(y) &= \frac{1}{2}\mathrm{erf}\left(\frac{y}{\sqrt{2}}\right), \\
D_2^*\left(0, 0; \mathbf{R}\right) &= \frac{1}{2\pi} \arcsin\left(r_{12}\right),
\end{aligned}
$$

and

$$D_k^*\left(\mathbf{0}; \mathbf{R}\right) \equiv 0 \quad \text{for } k \text{ odd}.$$

Using the abbreviation that

$$D_{k:j_1,\ldots,j_k}^* = D_k^*\left\{t_{j_1}, \ldots, t_{j_k}; \mathbf{R}\left(t_{j_1}, \ldots, t_{j_k}\right)\right\},$$

Dutt [**16**] provided the following representation for G

$$G(t_1,\ldots,t_p)=\left(\frac{1}{2}\right)^p-\left(\frac{1}{2}\right)^{p-1}\sum_{k=1}^{p}D_{1:k}^{*}+\left(\frac{1}{2}\right)^{p-2}\sum_{k<l=1}^{p}D_{2:kl}^{*}+\left(\frac{1}{2}\right)^{p-3}\sum_{k<l<m=1}^{p}D_{3:klm}^{*}+\cdots+D_{p:1,\ldots,p}.$$

Hence, by (20), the computation of P in (19) can be achieved by successive applications of the Gauss-Hermite quadrature formula using only positive Hermite zeros (Abramowitz and Stegun [**1**], page 924). There are several advantages for this approach. First, it is not necessary to invert the correlation matrix. In addition, (21) permits the use of Gauss quadrature formula that are remarkably effective in estimating the value of an integral from a few points, provided that the integral excluding the weighting function can be accurately approximated by a polynomial. Moreover, often the integrand separates as a product of two functions, one depending only on correlation coefficients and the other on the original limits of integration.

For selected correlation structures and several values of ν and $y=y_k$, $k=1,\ldots,p$, Dutt [**16**] computed values of P accurate up to six decimal places.

8 Amos' probability integral For the equicorrelation structure $r_{ij}=\rho$, $i\neq j$ considered by Gupta and Sobel [**29**] and Gupta [**27**]—but with the common ρ taken to be any positive real number less than 1—Amos [**3**] derived the following simpler expression for the probability integral

$$P(d)=\frac{2^{\nu-3/2}\Gamma\left((\nu+1)/2\right)}{\sqrt{\pi}\left(1+b^2\right)^{\nu/2}}\times\int_{-\infty}^{\infty}\exp\left(-\frac{dx^2}{2}\right)\Phi^p(x)\mathrm{erfc}\left(-\frac{cx}{\sqrt{2}}\right)dx,\tag{22}$$

where erfc$(\cdot)$ is the complementary error function defined by

$$\mathrm{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^{\infty}\exp\left(-z^2\right)dz,$$

and a, b, c and d are constants given by

$$a = \sqrt{\frac{1-\rho}{\rho}}\,, \qquad b = \frac{d}{\sqrt{\rho\nu}}\,,$$
$$c = \frac{ab}{\sqrt{1+b^2}}\,, \qquad d = \frac{a^2}{1+b^2}\,.$$

The reduction to (22) was obtained by means of a relationship between the parabolic cylinder function and the complementary error function. Amos [**3**] suggested computing the integral (22) by locating the x_0 for which the derivative of the integrand is zero and then summing quadratures on intervals of length h to the left and right of x_0 until a limit of integration is reached or the truncation error is small enough. The motivation for this procedure comes from the fact that x_0 can vary widely with extreme parameter values, and h, which estimates the spread of the integrand, can be small or large. Thus, x_0 and h accommodate the parameters, producing meaningful results by preventing quadratures over tails that are negligible or preventing gross misjudgments of the scale of integration. Letting $g(x)$ denote the integrand of (22), Amos [**3**] showed that the derivative of $\log g(x)$ decreases monotonically from ∞ to $-\infty$ as x traverses $(-\infty, \infty)$, guaranteeing a unique root x_0 of $g'(x) = 0$.

9 Fujikoshi's probability integrals Fujikoshi [**19**] provided asymptotic expansions as well as error bounds for the probability integral (19) when the correlation matrix $\mathbf{R} = \mathbf{I}_p$, the $p \times p$ identity matrix. Specifically, letting

$$a_{\delta,j}\,(y_1, \ldots, y_p) = \frac{d^j}{ds^j}\,\left\{\Phi\big(s^{-\delta/2} y_1\big) \cdots \Phi\big(s^{-\delta/2} y_p\big)\right\}\Big|_{s=1}\,,$$

where $\delta = -1$, 1, and Φ denotes the cdf of the standard normal distribution, Fujikoshi established the following approximation for the probability integral

$$\text{(23)} \quad P\,(y_1, \ldots, y_p) = \Phi\,(y_1) \cdots \Phi\,(y_p) + \sum_{j=1}^{k-1} \frac{1}{j!}\, a_{\delta,j}\,(y_1, \ldots, y_p)\, E\left[\left\{\left(\frac{\chi^2_\nu}{\nu}\right)^{\delta} - 1\right\}^j\right],$$

which we shall denote by $A_{\delta,k}(y_1, \ldots, y_p)$. Fujikoshi also derived uniform and nonuniform error bounds for this approximation. Under the assumptions that

$$\bar{a}_{\delta,k} = \sup_{\mathbf{y}} \left| a_{\delta,k} \left(y_1, \ldots, y_p \right) \right| < \infty,$$

and

$$E\left\{ \left(\frac{\chi_\nu^2}{\nu} \right)^k \right\} < \infty, \qquad E\left\{ \left(\frac{\nu}{\chi_\nu^2} \right)^k \right\} < \infty,$$

the uniform bound takes the form

$$\begin{aligned} &\sup_{\mathbf{y}} \left| P\left(y_1, \ldots, y_p \right) - A_{\delta,k}\left(y_1, \ldots, y_p \right) \right| \\ &\qquad\qquad \le \frac{1}{k!} \bar{a}_{\delta,k} E\left[\left\{ \left(\frac{\chi_\nu^2}{\nu} \right) \vee \left(\frac{\nu}{\chi_\nu^2} \right) - 1 \right\}^k \right]. \end{aligned}$$

Under the assumptions that

$$\bar{a}_{\delta,k}(l) = \sup_{\mathbf{y}} \left(1 + \parallel \mathbf{y} \parallel^l \right) \left| a_{\delta,k} \left(y_1, \ldots, y_p \right) \right| < \infty$$

and

$$E\left\{ \left(\frac{\chi_\nu^2}{\nu} \right)^{k+l/2} \right\} < \infty, \qquad E\left\{ \left(\frac{\nu}{\chi_\nu^2} \right)^k \right\} < \infty,$$

the nonuniform bound takes the form

$$\begin{aligned} &\left| P\left(y_1, \ldots, y_p \right) - A_{\delta,k}\left(y_1, \ldots, y_p \right) \right| \\ &\qquad \le \frac{1}{k!} \left(1 + \parallel \mathbf{y} \parallel^l \right)^{-1} \bar{a}_{\delta,k}(l) E\left\{ \left(\frac{\chi_\nu^2}{\nu} \right)^{l/2} \left| \frac{\chi_\nu^2}{\nu} - 1 \right|^k + \left| \frac{\nu}{\chi_\nu^2} \right|^k \right\}. \end{aligned}$$

Clearly the latter bounds are improvements on the uniform bounds in the tail part of the multivariate t distribution. In the case $p = 1$, these results provide useful approximations for the univariate Student's t distribution—see Fujikoshi [**18**] and Fujikoshi and Shimizu [**22**]. The special case of (23) for $y_j = y$ has been investigated more recently by Fujikoshi [**19, 20, 21**], Fujikoshi and Shimizu [**23**], and Shimizu and Fujikoshi [**48**].

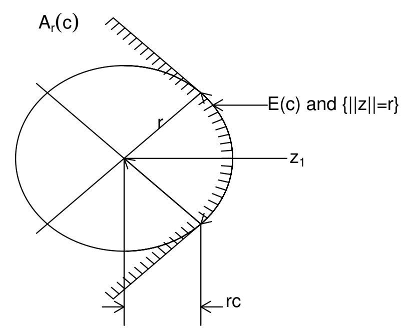

FIGURE 1: The sets $A_r(c)$ and $E(c) \cap \{\|\mathbf{z}\| = r\}$ in two dimensions

10 Probabilities of cone Consider the p-dimensional set

$$A_r(c) = \left\{\mathbf{x} : \mathbf{z}^T\mathbf{x} \le r\|\mathbf{z}\|, \quad \text{all } \mathbf{z} \text{ in } E(c)\right\}, \tag{24}$$

where $E(c) = \{\mathbf{z} : z_1 \ge c\|\mathbf{z}\|\}$, $\|\mathbf{z}\| = \sqrt{\mathbf{z}^T\mathbf{z}}$, and c is a nonnegative constant. The set $E(c)$ is the cone, with vertex at the origin, which intersects origin-centered spheres in spherical caps. This is illustrated in Figure 1 for $p = 2$.

Bohrer [**6**] studied the analytical shape of $A_r(c)$ and the associated probability

$$p(c, r, p, \nu) = \Pr(\mathbf{X} \in A_r(c))$$

when $\mathbf{X}$ has the p-variate t distribution with mean vector $\mathbf{0}$, covariance matrix $\sigma^2\mathbf{I}_p$, and degrees of freedom ν. The evaluation of $p(c, r, p, \nu)$ is of statistical interest and use in the construction of confidence bounds (Wynn and Bloomfield [**68**], Section 3; Bohrer and Francis [**7**], equation (2.3)) and in testing multivariate hypotheses (Kudô [**39**], Theorem 3.1, Section 5; Barlow *et al.* [**5**], pages 136ff, 177).

As regards the shape, Bohrer showed that every two-dimensional section of A_r containing the z_1-axis is exactly the two-dimensional version of A_r illustrated in Figure 1. Thus, A_r is the solid of revolution about the z_1-axis that is swept out by the A_r in Figure 1. To express this more precisely in mathematical terms—for an $p \times 1$ vector $\mathbf{v}$—define polar coordinates $R_{\mathbf{v}}$ and $\boldsymbol{\mu}_{\mathbf{v}} = \{\theta_{\mathbf{v}i}\}$, with $-\pi < \theta_{\mathbf{v}i} \leq \pi$, by

$$
\begin{aligned}
v_1 &= R_{\mathbf{v}} \cos \theta_{\mathbf{v}1}, \\
v_i &= R_{\mathbf{v}} \cos \theta_{\mathbf{v}i} \prod_{j=1}^{i-1} \sin \theta_{\mathbf{v}j}, \quad i = 2, \ldots, p-1,
\end{aligned}
$$

and

$$
v_p = R_{\mathbf{v}} \prod_{j=1}^{i-1} \sin \theta_{\mathbf{v}j}.
$$

Also define

$$
\begin{aligned}
\theta^* &= \arccos c, \\
T_1 &= \left\{ \mathbf{x} : |\theta_{\mathbf{t}1}| \leq \theta^*, \ R_{\mathbf{t}} \leq r \right\}, \\
T_2 &= \left\{ \mathbf{x} : \theta_{\mathbf{t}1} - \theta^* \in \left(0, \frac{\pi}{2} \right], \ R_{\mathbf{t}} \cos \left(\theta_{\mathbf{t}1} - \theta^* \right) \leq r \right\}, \\
T_3 &= \left\{ \mathbf{x} : \theta_{\mathbf{t}1} + \theta^* \in \left[-\frac{\pi}{2}, 0 \right), \ R_{\mathbf{t}} \cos \left(\theta_{\mathbf{t}1} + \theta^* \right) \leq r \right\},
\end{aligned}
$$

and

$$
T_4 = \left\{ \mathbf{x} : |\theta_{\mathbf{t}1}| > \theta^* + \frac{\pi}{2} \right\}.
$$

Then the set A_r is the union of the disjoint sets $T_1, \ldots, T_4$. As regards evaluating the probability $p(c, r, p, \nu)$, Bohrer [**6**] derived the following expression

$$
\begin{aligned}
p(c, r, p, \nu) = {} & \frac{k(\theta^*)}{k(\pi/2)} \Pr \left(F_{p,\nu} \leq \frac{r^2}{\sigma^2} \right) \\
& + \frac{k(\pi/2 - \theta^*)}{k(\pi/2)} \frac{1}{2k(\pi/2)} \sum_{j=0}^{p-2} B \left(\frac{j+1}{2}, \frac{p-j-1}{2} \right)
\end{aligned}
$$

$$\times \binom{p-2}{j} c^j \left(1-c^2\right)^{(p-2-j)/2} \Pr\left(F_{p-2-j,\nu} \le \frac{r^2}{\sigma^2}\right),$$

where $k(\theta)$ is given by

$$k(\theta) = \frac{\left(2^m m!\right)^2}{p!}(1-\cos\theta) - \frac{\cos\theta \sin^{2m}\theta}{p} - \sum_{l=1}^{m-1} \frac{\sin^{2(m-l)}\theta\cos\theta}{p-2l} \prod_{j=1}^{l} \frac{p-2j}{p+1-2j}$$

when $p = 2m+1$ is odd, and by

$$k(\theta) = \frac{(p-1)!\theta}{2^{p-1}(m-1)!m!} - \frac{\sin^{p-1-2m}\theta\cos\theta}{p} - \sum_{l=1}^{m-1} \frac{\sin^{p-1-2l}\theta\cos\theta}{p-2l} \prod_{j=1}^{l} \frac{p+1-2j}{p+2-2j}$$

when $p = 2m$ is even. The statistical questions that motivate this work ask what radius r is required so that $p(c, r, p, \nu) = \alpha$ for preassigned values of α. For $p \le 5$, Bohrer [**6**] provided tables of these percentiles for $\alpha = 0.95$ and 0.99 and for a range of (c, ν) pairs.

11 Probabilities of convex polyhedra It is well known (Nicholson [**41**]; Cadwell [**10**]; Owen [**42**]) that probabilities of polygons under bivariate normal distributions can be evaluated in terms of probabilities of right-angled triangles with vertices $(0,0)$, $(y_1, 0)$, (y_1, y_2), $y_j > 0$, $j = 1, 2$ under bivariate normal distributions with zero correlation. John [**34**] proved an analogous result that probabilities of polygonal and angular regions for a given bivariate t distribution can be expressed in terms of $V_\nu(y_1, y_2)$, the integral of

$$f(x_1, x_2; \nu) = \frac{\Gamma((\nu+2)/2)}{\nu\pi\Gamma(\nu/2)} \left\{1 + \frac{x_1^2 + x_2^2}{\nu}\right\}^{-(\nu+2)/2}$$

over the right-angled triangles with vertices $(0,0)$, $(y_1, 0)$, and (y_1, y_2). John [**34**] also provided several formulas for evaluating $V_\nu(y_1, y_2)$. A formula in terms of the incomplete beta function is

$$V_\nu(y_1, y_2) = \frac{1}{2\pi}\arctan\left(\frac{y_2}{y_1}\right) - \frac{y_1 c^{\nu/2}}{4\pi\sqrt{\nu + y_1^2}}\sum_{k=0}^{\infty} c^k B_u\left(\frac{\nu}{2} + k + \frac{1}{2}, \frac{1}{2}\right), \tag{25}$$

where

$$c = \frac{\nu}{\nu + y_1^2}, \qquad u = \frac{\nu + y_1^2}{\nu + y_1^2 + y_2^2},$$

and

$$B_x(a, b) = \int_x^1 w^{a-1}(1 - w)^{b-1}\, dw$$

is the incomplete beta function. This series converges slowly unless y_1 is large in relation to ν. In the two cases ν odd and ν even, (25) can be reduced considerably. If $\nu = 2m$ for a positive integer m, then

$$V_{2m}(y_1, y_2) = \frac{\sqrt{1 - c}}{4\pi}\sum_{k=0}^{m-1} c^k B_u\left(k + \frac{1}{2}, \frac{1}{2}\right) \tag{26}$$

while if $\nu = 2m + 1$ for a nonnegative integer m, then

$$V_{2m+1}(y_1, y_2) = \frac{1}{2\pi}\arctan\left(\frac{y_2}{y_1}\right) - \frac{1}{4\pi}B_v\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{\sqrt{c(1 - c)}}{4\pi}\sum_{k=0}^{m-1} c^k B_u\left(k + 1, \frac{1}{2}\right), \tag{27}$$

where

$$v = \frac{\nu\left(\nu + y_1^2 + y_2^2\right)}{\nu\left(\nu + y_1^2 + y_2^2\right) + y_1^2 y_2^2}.$$

An attractive feature of (26) and (27) is that, when utilizing them for evaluating V_{2m} and V_{2m+1}, they are already evaluated for lower values of m also. If one performs the summations in the order indicated in the formulas, the addition of each term will yield values of V_{2m} or V_{2m+1} for the next higher value of m. This feature makes it particularly suitable for use in preparing tables.

A second formula for $V_\nu(y_1, y_2)$ given in John [**34**] is an expansion in powers of $1/\nu$

$$V_\nu\left(y_1, y_2\right) = V_\infty\left(y_1, y_2\right) - \frac{\exp\left(-y_1^2/2\right)}{2\pi} \sum_{k=1}^{\infty} \frac{\nu^{-k}}{k!} U_k\left(y_1, y_2\right), \tag{28}$$

where the first three U_k are given by

$$U_1\left(y_1, y_2\right) = \frac{y_1^4}{4} W_1\left(y_1, y_2\right),$$

$$U_2\left(y_1, y_2\right) = y_1^6 \left\{ -\frac{1}{3} W_2\left(y_1, y_2\right) + \frac{y_1^2}{16} W_3\left(y_1, y_2\right) \right\},$$

and

$$U_3\left(y_1, y_2\right) = y_1^8 \left\{ \frac{3}{4} W_3\left(y_1, y_2\right) - \frac{y_1^2}{4} W_4\left(y_1, y_2\right) + \frac{y_1^4}{64} W_5\left(y_1, y_2\right) \right\},$$

where

$$W_\nu\left(y_1, y_2\right) = \int_0^{y_2/y_1} \left(1 + t^2\right)^\nu \exp\left(-\frac{y_1^2 t^2}{2}\right) dt.$$

The term V_∞ in (28) is the integral of $\exp\{-(y_1^2 + y_2^2)\}/(2\pi)$ over the right-angled triangle with vertices $(0, 0)$, $(y_1, 0)$, and $(0, y_2)$. The method of derivation for (28) is similar to the classical method employed by Fisher [**17**] for expanding the probability integral of Student's t. Despite the complexity of (28) over (25), (28) should be preferred if ν is sufficiently large. The first two or three terms of (28) then can be expected to provide fairly accurate values of V_ν.

John [**34**] also provided a recurrence relation and an approximation for $V_\nu(y_1, y_2)$; the latter proved to be satisfactory only when either ν is too small or y_2/y_1 is too large. In a subsequent paper, John [**35**] extended this result to higher dimensions, by showing that the probabilities of the p-dimensional convex polyhedra with vertices (0, 0, 0, 0, …, 0), (y_1, 0, 0, 0, …, 0), (y_1, y_2, 0, 0, …, 0), …, (y_1, y_2, y_3, y_4, …, y_p), $h_j > 0$, $j = 1, 2, \ldots, p$ under a p-variate t distribution with ν degrees of freedom can be expressed in terms of the function $V_\nu(y_1, y_2, \ldots, y_p)$, the integral of the p-variate t pdf

$$f\left(x_1, x_2, \ldots, x_p; \nu\right) = \frac{\Gamma((\nu + p)/2)}{(\nu\pi)^{p/2}\Gamma(\nu/2)} \left\{ 1 + \frac{x_1^2 + x_2^2 + \cdots + x_p^2}{\nu} \right\}^{-\frac{\nu+p}{2}}$$

over the same p-dimensional convex polyhedra. John also provided an important asymptotic expansion in powers of $1/\nu$ connecting $V_\nu(y_1,\, y_2,\, \ldots,\, y_p)$ with $V(y_1, y_2, \ldots, y_p)$, the integral of the p-variate normal pdf

$$f\left(x_1, x_2, \ldots, x_p; \infty\right) = (2\pi)^{-p/2} \exp\left\{-\left(x_1^2 + x_2^2 + \cdots + x_p^2\right)/2\right\}$$

over the same polyhedra discussed above. Up to the order of the term $O(1/\nu^2)$, the expansion is

$$\begin{aligned}
&V_\nu\left(y_1, y_2, \ldots, y_p\right) \\
&\quad = V\left(y_1, y_2, \ldots, y_p\right) + \frac{1}{4\nu}\left\{y_1 y_2 f\left(y_1, y_2\right) V\left(y_3, y_4, \ldots, y_p\right)\right. \\
&\quad\quad \left. - y_1\left(1 + y_1^2\right) f\left(y_1\right) V\left(y_2, y_3, \ldots, y_p\right)\right\} \\
&\quad\quad + \frac{1}{96\nu^2}\left\{3 y_1 y_2 y_3 y_4 f\left(y_1, y_2, y_3, y_4\right) V\left(y_5, y_6, \ldots, y_p\right)\right. \\
&\quad\quad - y_1 y_2 y_3\left(2 + 9y_1^2 + 6y_2^2 + 3y_3^2\right) f\left(y_1, y_2, y_3\right) V\left(y_4, \ldots, y_p\right) \\
&\quad\quad - y_1 y_2\left(3 + 5y_1^2 + y_2^2 - 9y_1^4 - 9y_1^2 y_2^2 - 3y_2^4\right) V\left(y_3, \ldots, y_p\right) f\left(y_1, y_2\right) \\
&\quad\quad \left. + y_1\left(3 + 5y_1^2 + 7y_1^4 - 3y_1^6\right) f\left(y_1\right) V\left(y_2, \ldots, y_p\right)\right\} + o\left(\frac{1}{\nu^2}\right).
\end{aligned}$$

In this formula, $V(y_m, y_{m+1}, \ldots, y_p)$ is to be replaced by 0 if $m \geq p+2$ and by 1 if $m = p+1$. In principle, there is no difficulty in determining further terms of this expansion, but the coefficients of higher powers of $(1/\nu)$ have rather complicated expressions. Other useful results given by John [**35**] include recursion formulas connecting $V_\nu(y_1, y_2, \ldots, y_p)$ with $V_{\nu\pm 2}(y_1, y_2, \ldots, y_p)$.

More recently, several authors have looked into the problem of computing multivariate t probabilities of the form

$$P = \int_A f\left(\mathbf{x}; \nu\right) d\mathbf{x}, \tag{29}$$

where $\mathbf{X}$ has the central multivariate t distribution with correlation matrix $\mathbf{R}$ and A is any convex region. Somerville [**50, 51, 52, 53**] developed the first known procedures for evaluating P in (29). Let $\mathbf{M}\mathbf{M}^T$ be the Cholesky decomposition of $\mathbf{R}$ (where $\mathbf{M}$ is a lower triangular matrix) and set $\mathbf{X} = \mathbf{M}\mathbf{W}$. Then $\mathbf{W}$ is multivariate t with correlation matrix $\mathbf{I}_p$. If one further sets $r^2 = \mathbf{W}^T\mathbf{W}$, then $F = r^2/p$ has the well known

F distribution with degrees of freedom p and ν. Let A be the region bounded by p hyperplanes and described by

$$\mathbf{GW} \le \mathbf{d},$$

where $\mathbf{G} = (\mathbf{g}_1, \ldots, \mathbf{g}_p)$ and the jth hyperplane is $\mathbf{g}_j^T \mathbf{W} = d_j$. For a random direction $\mathbf{c}$, let r be the distance from the origin to the boundary of A, that is, the smallest positive distance from the origin to the jth plane, $j = 1, \ldots, p$. Then an unbiased estimate of the integral P in (29) is

$$\Pr\left(F \le r^2/p\right). \tag{30}$$

To implement the procedure, Somerville chose successive random directions $\mathbf{c}$ and obtained corresponding estimates of (30). The value of P was then taken as the arithmetic mean of the individual estimates.

Somerville [**54, 56**] provided the following modification of the above procedure. Let r^* be the minimum distance from the origin to the boundary of A, that is, the smallest of the r for all random directions $\mathbf{c}$. Divide A into two regions, the portion inside the hypersphere of radius r^* and centered at the origin, and the region outside. The probability content of the hypersphere is

$$P_1 = \Pr\left(F \le {r^*}^2/p\right),$$

and this can be estimated as in Somerville [**50, 51, 52, 53**]. If $E(v)$ and $e(v)$, respectively, denote the cdf and the pdf of $v = 1/r$ (the reciprocal distance from the origin from and to the boundary of A), then the probability content of the outer region is

$$P_2 = \int_0^{1/r^*} E(v)e(v)\, dv.$$

Since $F = r^2/p$, the pdf of v is

$$e(v) = \frac{2\nu^{\nu/2}\Gamma\left((\nu+p)/2\right)}{\Gamma\left(\nu/2\right)\Gamma\left(p/2\right)} \cdot \frac{v^{\nu-1}}{\left(1+\nu v^2\right)^{(\nu+p)/2}}.$$

The strategy is to use some numerical method to estimate $E(v)$ and then evaluate the integral P_2 using the Gauss-Legendre quadrature. The approaches of Somerville [**54, 56**] differ in that Somerville [**54**] applied

Monte Carlo techniques to estimate $E(v)$ while Somerville [**56**] used a binning procedure. It should be noted, however, that an approach similar to these had been introduced earlier by Deak [**13**].

Somerville [**57**] provided an extension of the above methodologies to evaluate P in (29) when A is an ellipsoidal region. This has potential applications in the field of reliability (in particular relating to the computation of the tolerance factor for multivariate normal populations) and to the calculation of probabilities for linear combinations of central and noncentral chi-squared and F. In the coordinate system of the transformed variables $\mathbf{W}$, assume, without loss of generality, that the axes of the ellipsoid are parallel to the coordinate axes and the ellipsoid has the equation $(\mathbf{w}-\mathbf{u})^T\mathbf{B}^{-1}(\mathbf{w}-\mathbf{u}) = 1$, where $\mathbf{B}$ is a diagonal matrix with the ith element given by b_i. If the ellipsoid contains the origin, then for each random direction $\mathbf{c}$ there is a unique distance r to the boundary. An unbiased estimate of P is then given by

$$\Pr\left(F \le r^2/p\right).$$

If the ellipsoid does not contain the origin, then, for a random direction, a line from the origin in that direction either intersects the boundary of the ellipsoid at two points (say $r \ge r_*$) or does not intersect it at all. If the line intersects the boundary, then an unbiased estimate of P is given by the difference

$$\Pr\left(F \le r^2/p\right) - \Pr\left(F \le r_*^2/p\right).$$

If the line does not intersect the ellipsoid, an unbiased estimate is 0. As in the first procedure described above, this is repeated for successive random directions $\mathbf{c}$, each providing an unbiased estimate. The value of P is then taken as the arithmetic average. A modification of this procedure along the lines of Somerville [**54, 56**] is described in Somerville [**57**].

Somerville [**58**] provided an application of his methods for multiple testing and comparisons by taking A in (29) to be

$$A = \left\{\mathbf{x} \in \Re^p : \max \mathbf{c}^T\mathbf{x} \le \frac{q}{\sqrt{2}}\right\}, \quad \mathbf{c} \in B,$$

where B is the set of contrasts corresponding to the different hypotheses and $q > 0$. The purpose is to calculate the value of q for arbitrary $\mathbf{R}$ and ν and arbitrary sets B such that the probability content of A has a preassigned value γ. Somerville and Bretz [**60**] have written two *Fortran 90* programs (QBATCH4.FOR and QINTER4.FOR) and two *SAS-IML* programs (QBATCH4.SAS and QINTER4.SAS) for this purpose.

QINTER4.FOR and QINTER4.SAS are interactive programs, while the other two are batch programs. A compiled version of the *Fortran 90* programs that should run on any PC with Windows 95 or later can be found at

`http://pegasus.cc.ucf.edu/~somervil/home.html`

These programs implement the methodology described above to evaluate the probability content of A (A *Fortran 90* programs MVI3.FOR used to evaluate multivariate t integrals over any convex region is described in Somerville [**55**]. An extended *Fortran 90* programs MVELPS.FOR to evaluate multivariate t integrals over any ellipsoidal regions is described in Somerville [**59**]. The average running times for the latter program range from 0.075 and 0.109 second for $p = 2$ and 3, respectively, to 0.379 and 0.843 second for $p = 10$ and 20, respectively.). The so-called "Brent's method," an interactive procedure described in Press [**45**], is used to solve for the value of q. The time to estimate the q values (with a standard error of 0.01) using QINTER4 or QBATCH4 range from 10 seconds for Dunnett's multiple comparisons procedure to 52 seconds for Tukey's procedure, using a 486-33 processor.

A problem that frequently arises in statistical analysis is to compute (29) when A is a rectangular region, that is,

$$P = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_p}^{b_p} f\left(x_1, x_2, \ldots, x_p\right) dx_p \cdots dx_2\, dx_1. \tag{31}$$

Wang and Kennedy [**67**] employed numerical interval analysis to compute P. The method is similar to the approaches of Corliss and Rall [**11**] for univariate normal probabilities and Wang and Kennedy [**66**] for bivariate normal probabilities. The basic idea is to apply the multivariate Taylor expansion to the joint pdf f. Letting $c_j = (a_j + b_j)/2$, the Taylor expansion of f at the mid point $(c_1, c_2, \ldots, c_p)$ is

$$\begin{aligned} & f\left(x_1, x_2, \ldots, x_p\right) \\ &= \sum_{k=0}^{m-1} \left\{ \sum_{]k[} \frac{1}{k_1! \cdots k_p!} \cdot \frac{\partial^k f\left(c_1, c_2, \ldots, c_p\right)}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_p^{k_p}} \prod_{j=1}^{p} \left(x_j - c_j\right)^{k_j} \right\} \\ &\qquad + \sum_{]m[} \frac{1}{m_1! \cdots m_p!} \cdot \frac{\partial^m f\left(\xi_1, \xi_2, \ldots, \xi_p\right)}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_p^{m_p}} \prod_{j=1}^{p} \left(x_j - c_j\right)^{m_j}, \end{aligned} \tag{32}$$

where ξ_j is contained in the integration region $[a_j, b_j]$ and $]k[$ denotes all possible partitions of k into p parts. For example, in the case $p = 3$, $]2[$ will result in 6 possible partitions of '2' into $\{k_1, k_2, k_3\}$: $\{0, 0, 2\}$, $\{0, 1, 1\}$, $\{0, 2, 0\}$, $\{1, 0, 1\}$, $\{1, 1, 0\}$, and $\{2, 0, 0\}$. The main problem with computing (32) is the presence of high-order partial derivatives of f. Defining

$$(f)_{k_1 k_2 \cdots k_p} = \frac{1}{k_1! k_2! \cdots k_p!} \cdot \frac{\partial^{k_1 + k_2 + \cdots + k_p} f}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_p^{k_p}}, \tag{33}$$

Wang and Kennedy derived the following recursive formula

$$\begin{aligned}(f)_{k_1 k_2 \cdots k_p} = -\frac{1}{k_1} \left(1 + \frac{\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}}{\nu}\right)^{-1} & \\ \times \sum_{l_1=0}^{k_1} \sum_{l_2=0}^{k_2} \cdots \sum_{l_p=0}^{k_p} & \left\{\frac{p+\nu}{2} (k_1 - l_1) + l_1\right\} (f)_{l_1 l_2 \cdots l_p} \\ \times & \left(1 + \frac{\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}}{\nu}\right)_{k_1 - l_1, k_2 - l_2, \ldots, k_p - l_p}.\end{aligned}$$

With regard to the last quadratic term, it should be noted that higher than second-order partial derivatives are all zero. To carry out the computation of (33) for a given $(k_1, k_2, \ldots, k_p)$, one can

- first let one l_j be $k_j - 1$ (if this $k_j \neq 1$) and all the other l_j's be their corresponding k_j's;
- next let l_r and l_s be $k_r - 1$ and $k_s - 1$, respectively (if $k_r \neq 1$ and $k_s \neq 1$), while all the other l_j's take their corresponding k_j's;
- finally, let some l_j be $k_j - 2$ (if $k_j \geq 2$) and all other l_j's be the corresponding k_j's.

The total number of terms that contribute to computing $(f)_{k_1 k_2 \cdots k_p}$ is at most $p(p+3)/2$. Compared to the multivariate normal distribution, this number is larger (Wang and Kennedy [**66**]). The following table gives the running times and the accuracy for computing (31) with $\nu = 10$.

Running time and accuracy for computing P in (31)

p	Running time (min)	$a_j = -0.5$ $b_j = 0.5$	$a_j = -0.4$ $b_j = 0.4$	$a_j = -0.3$ $b_j = 0.3$	$a_j = -0.2$ $b_j = 0.2$
10	80			2 sig	4 sig
9	70			3 sig	7 sig
8	85		0 sig	5 sig	10 sig
7	90		3 sig	8 sig	
6	110	3 sig	8 sig		
5	180	10 sig			

Another point to note about Wang and Kennedy's method is that when the integration region is near the origin it works better for larger ν, while when the integration region is off the origin it works better for smaller ν.

The main problem with Wang and Kennedy's [**67**] method is that the calculation times required are too large even for low accuracy results (see the table above). Genz and Bretz [**25**] proposed a new method for computing (31) by transforming the p-variate integrand into a product of univariate integrands. The method is similar to the one used by Genz [**24**] for the multivariate normal integral.

Letting $\mathbf{M}\mathbf{M}^T$ be the Cholesky decomposition of $\mathbf{R}$, define the following transformations

$$X_j = \sum_{k=1}^{p} M_{j,k} Y_k,$$

$$Y_j = U_j \sqrt{\frac{\nu + \sum_{k=1}^{j-1} Y_k^2}{\nu + j - 1}},$$

$$U_j = T_{\nu+j-1}\left(Z_j\right),$$

and

$$Z_j = d_j + W_j\left(e_j - d_j\right),$$

where T_τ denotes the cdf of the univariate Student's t distribution with degrees of freedom τ,

$$d_j = T_{\nu+j-1}\left(\hat{a}_j\right),$$

$$e_j = T_{\nu+j-1}\left(\hat{b}_j\right),$$

$$\hat{a}_j = a'_j \sqrt{\frac{\nu + j - 1}{\nu + \sum_{k=1}^{j-1} y_k^2}},$$

$$\hat{b}_j = b'_j \sqrt{\frac{\nu + j - 1}{\nu + \sum_{k=1}^{j-1} y_k^2}},$$

$$a'_j = \frac{a_j - \sum_{k=1}^{j-1} m_{j,k} Y_k}{m_{j,j}},$$

and

$$b'_j = \frac{b_j - \sum_{k=1}^{j-1} m_{j,k} Y_k}{m_{j,j}}.$$

Applying the above transformations successively, Genz and Bretz reduced (31) to

$$P = (e_1 - d_1) \int_0^1 (e_2 - d_2) \cdots \int_0^1 (e_p - d_p) \int_0^1 d\mathbf{w} \tag{34}$$

$$= \int_0^1 \int_0^1 \cdots \int_0^1 f(\mathbf{w})\, d\mathbf{w}. \tag{35}$$

The transformation has the effect of flattening the surface of the original function, and P becomes an integral of $f(\mathbf{w}) = (e_1 - d_1) \cdots (e_p - d_p)$ over the $(p-1)$-dimensional unit hypercube. Hence, one has improved numerical tractability and (35) can be evaluated with different multidimensional numerical computation methods. Genz and Bretz considered three numerical algorithms for this: an acceptance-rejection sampling algorithm, a crude Monte Carlo algorithm, and a lattice rule algorithm.

- Acceptance-rejection sampling algorithm: Generate p-dimensional uniform random vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N$ and estimate P by

  $$\widehat{P} = \frac{1}{N} \sum_{l=1}^{N} h(\mathbf{M}\mathbf{y}_l),$$

  where

  $$h(\mathbf{x}) = \begin{cases} 1, & \text{if } a_j \le x_j \le b_j,\ j = 1, 2, \ldots, p, \\ 0, & \text{otherwise} \end{cases}$$

and

$$y_{l,j} = T^{-1}_{\nu+j-1}\left(w_{l,j}\right)\sqrt{\frac{\nu+\sum_{k=1}^{j-1} y_k^2}{\nu+j-1}},$$

$$j = 1, 2, \ldots, p, \quad l = 1, 2, \ldots, N.$$

- A crude Monte Carlo algorithm: Generate $(p-1)$-dimensional uniform random vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N$ and estimate P by

$$\widehat{P} = \frac{1}{N}\sum_{l=1}^{N} f\left(\mathbf{w}_l\right),$$

an unbiased estimator of the integral (35).
- A lattice rule algorithm (Joe [**32**]; Sloan and Joe [**49**]): Generate $(p-1)$-dimensional uniform random vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N$ and estimate P by

$$\widehat{P} = \frac{1}{Nq}\sum_{l=1}^{N}\sum_{j=1}^{q} f\left(\left|2\left\{\frac{j}{q}\mathbf{z} + \mathbf{w}_l\right\} - \mathbf{1}_p\right|\right).$$

Here N is the simulation size, usually very small, q corresponds to the fineness of the lattice, and $\mathbf{z} \in \Re^{p-1}$ denotes a strategically chosen lattice vector. Braces around vectors indicate that each component has to be replaced by its fractional part. One possible choice of $\mathbf{z}$ follows the good lattice points; see, for example, Sloan and Joe [**49**].

For all three algorithms—to control the simulated error—one may use the usual error estimate of the means. Perhaps the most intuitive one of the three is the acceptance-rejection method. However, Deak [**13**] showed that, among various methods, it is the one with the worst efficiency. Genz and Bretz [**26**] proposed the use of the lattice rule algorithm. Bretz *et al.* [**9**] provided an application of this algorithm for multiple comparison procedures.

The method of Genz and Bretz [**25**] described above also includes an efficient evaluation of probabilities of the form

$$P = \int_{\mathbf{a}}^{\mathbf{b}} g(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x},$$

where $g(\mathbf{x})$ is some nuisance function. *Fortran* and *SAS-IML* codes to implement the method for $p \leq 100$ are available from the Web sites with URLs http://www.bioinf.uni-hannover.de/~betz/

and

http://www.sci.wsu.edu/math/faculty/genz/homepage.

12 Probabilities of linear inequalities Let X be a random variable characterizing the "load," and let Y be a random variable determining the "strength" of a component. Then the probability that a system is "trouble-free" is $\Pr(Y > X)$. In a more complicated situation, the operation of the system may depend on a linear combination of random vectors, say $\mathbf{a}_1^T \mathbf{X}_1 + \mathbf{a}_2^T \mathbf{X}_2 + b$, and the probability of a trouble-free operation will be

$$\Pr\left(\mathbf{a}_1^T \mathbf{X}_1 + \mathbf{a}_2^T \mathbf{X}_2 + b > 0\right), \tag{36}$$

where $\mathbf{X}_j$ are independent k_j-dimensional random vectors, $\mathbf{a}_j$ are k_j-dimensional constant vectors, and b is a scalar constant. Abusev and Kolegova [**2**] studied the problem of constructing unbiased, maximum likelihood, and Bayesian estimators of the probability (36) when $\mathbf{X}_j$ is assumed to have the multivariate t distribution with mean vector $\boldsymbol{\mu}_j$ and correlation matrix $\mathbf{R}_j$. If $\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1}$ and $\mathbf{x}_{21}, \ldots, \mathbf{x}_{2n_2}$ are iid samples from the two multivariate t distributions, then – in the where case both $\boldsymbol{\mu}_j$ and $\mathbf{R}_j$ are unknown—it was established that the unbiased and the maximum likelihood estimators are

$$\widehat{\Pr}\left(\mathbf{a}_1^T \mathbf{X}_1 + \mathbf{a}_2^T \mathbf{X}_2 + b > 0\right) = \frac{\Gamma\left(n_1/2\right)\Gamma\left(n_2/2\right)}{\pi\Gamma\left((n_1-1)/2\right)\Gamma\left((n_2-1)/2\right)} \times \int_{\Omega_1} \prod_{j=1}^{2} \left(1-\nu_j^2\right)^{(n_j-3)/2} d\nu_1 d\nu_2$$

and

$$\breve{\Pr}\left(\mathbf{a}_1^T \mathbf{X}_1 + \mathbf{a}_2^T \mathbf{X}_2 + b > 0\right) = \Phi\left(\frac{\mathbf{a}_1^T \overline{\mathbf{x}}_{n_1} + \mathbf{a}_2^T \overline{\mathbf{x}}_{n_2} + b}{\sqrt{\mathbf{a}_1^T \mathbf{S}_{n_1+1} \mathbf{a}_1 + \mathbf{a}_2^T \mathbf{S}_{n_2+1} \mathbf{a}_2}}\right),$$

respectively, where

$$\Omega_1 = \left\{\nu_j^2 < 1,\ j=1,2,\ \sum_{j=1}^{2} \nu_j \sqrt{n_j \mathbf{a}_j^T \mathbf{S}_{n_j+1} \mathbf{a}_j} + \sum_{j=1}^{2} \mathbf{a}_j^T \overline{\mathbf{x}}_j + b > 0\right\},$$

$$\overline{\mathbf{x}}_{n_j} = \frac{1}{n_j}\sum_{m=1}^{n_j} \mathbf{x}_m, \qquad (n_j+1)\,\overline{\mathbf{x}}_j = \sum_{m=1}^{n_j+1} \mathbf{x}_{jm},$$

$$(n_j+1)\,\mathbf{S}_{n_j+1} = \sum_{m=1}^{n_j+1} (\mathbf{x}_{jm}-\overline{\mathbf{x}}_j)\,(\mathbf{x}_{jm}-\overline{\mathbf{x}}_j)^T,$$

and $\mathbf{x}_{n_j+1} = \mathbf{x}$. A Bayesian estimator of (36) with unknown parameters $\boldsymbol{\mu}_j$ and $\mathbf{R}_j$ and the Lebesgue measure $p(\boldsymbol{\theta})\,d\boldsymbol{\theta} = d\boldsymbol{\mu}\,d\mathbf{R}$ was calculated to be

$$\begin{aligned}
\mathrm{Pr}_B &\left(\mathbf{a}_1^T\mathbf{X}_1 + \mathbf{a}_2^T\mathbf{X}_2 + b > 0\right) \\
&= \prod_{j=1}^{2} \frac{\Gamma\left((n_j-k_j)/2\right) n_j^{k_j}}{\pi\Gamma\left((n_j-1)/2\right)(n_j+1)^{k_j}} \prod_{j=1}^{2} \frac{\Gamma\left((n_j-k_j-1)/2\right)}{\Gamma\left((n_j-2k_j-1)/2\right)} \\
&\qquad\times \int_{\Omega_2} \prod_{j=1}^{2} \left(1-z_j^2\right)^{\frac{n_j-3}{2}} dz_1\,dz_2,
\end{aligned}$$

where

$$\Omega_2 = \left\{ z_j^2 < 1,\ j=1,2,\ \sum_{j=1}^{2} z_j \sqrt{n_j \mathbf{a}_j^T \mathbf{S}_{n_j+1}\mathbf{a}_j} + \sum_{j=1}^{2} \mathbf{a}_j^T \overline{\mathbf{x}}_j + b > 0 \right\}.$$

This Bayesian estimator is biased and is related to the unbiased estimator via the relation

$$\mathrm{Pr}_B\left(\mathbf{a}_1^T\mathbf{X}_1 + \mathbf{a}_2^T\mathbf{X}_2 + b > 0\right) = A\widehat{\mathrm{Pr}}\left(\mathbf{a}_1^T\mathbf{X}_1 + \mathbf{a}_2^T\mathbf{X}_2 + b > 0\right),$$

where

$$A = \prod_{j=1}^{2} \frac{\Gamma\left((n_j-k_j-1)/2\right)\Gamma\left((n_j-k)/2\right)}{\Gamma\left(n_j-2k_j-1)/2\right)\Gamma\left(n_j/2\right)} \cdot \frac{(n_j+1)^{k_j}}{n_j^{k_j}}.$$

The coefficient A can be expanded as

$$A = 1 + \frac{k}{n} - \frac{k}{n-k} + O\left(\frac{1}{n^2}\right),$$

where $n = \max(n_1, n_2)$ and $k = \max(k_1, k_2)$. Therefore, the Bayesian estimator is asymptotically unbiased as $n \to \infty$.

Substantial literature is now available on problems concerning probabilities of the form (36) for various distributions. For a comprehensive and up-to-date summary, the reader is referred to Kotz *et al.* [**38**].

13 Maximum probability content Let $\mathbf{X}$ be a bivariate random vector with the joint pdf of the form

$$f(\mathbf{x}) = g\left((\mathbf{x}-\boldsymbol{\mu})^T \mathbf{R}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right), \tag{37}$$

which, of course, includes the bivariate t pdf. Consider the class of rectangles

$$R(a) = \left\{ (x_1, x_2) : |x_1| \leq a, \ |x_2| \leq \frac{\lambda}{4a} \right\}$$

with the area equal to λ. Kunte and Rattihalli [**40**] studied the problem of characterizing the region R in this class for which the probability $P(R(a)) = \Pr(\mathbf{X} \in R(a))$ is maximum. As noted in Rattihalli [**46**], the characterizations of such regions is useful for obtaining Bayes regional estimators when (i) the decision space is the class of rectangular regions and (ii) the loss function is a linear combination of the area of the region and the indicator of the noncoverage of the region. It was shown that, for any fixed $\lambda > 0$, the maximal set is

$$\left\{ (x_1, x_2) : |x_1 - \mu_1| \leq c, \ |x_2 - \mu_2| \leq \frac{\lambda}{4c} \right\},$$

where c is given by

$$c = \sqrt{\frac{\lambda}{4} \sqrt{\frac{r^{22}}{r^{11}}}}.$$

Here, r^{ij} denotes the (i,j)th element of the inverse of $\mathbf{R}$. In particular, if $\boldsymbol{\mu} = \mathbf{0}$, $r^{12} = r^{21} = \rho$ and $|\ \rho\ | < 1$ in (37), then $P(R(a))$ is increasing for $a < \sqrt{\lambda}/2$ and is decreasing for $a > \sqrt{\lambda}/2$.

14 Monte Carlo evaluation Let $\mathbf{X}$ be a central p-variate t random vector with correlation matrix $\mathbf{R}$ and degrees of freedom ν. Vijverberg [**63**] developed a family of simulators of the multivariate t probability $p = \Pr(\mathbf{X} \leq \mathbf{X}_0)$ based on Monte Carlo simulation and recursive importance sampling. We shall provide the basic steps of this rather complicated but powerful procedure.

Define $\mathbf{Z} = \mathbf{A}\mathbf{X}$, where $\mathbf{A}$ is an upper triangular matrix such that $\mathbf{A}^T\mathbf{A} = \mathbf{R}$. Then it is well known that the pdf of $\mathbf{Z}$ can be expressed as a product of univariate Student's t pdfs

$$f(\mathbf{z}) = \prod_{k=1}^{p} f_1\left(z_k; \sigma_k^2, \nu_k\right),$$

where

$$\nu_k = \nu + p - k, \quad k = 1, 2, \ldots, p,$$

$$\sigma_k^2 = \frac{\nu + y_{k+1}^2 + \cdots + y_p^2}{\nu_k}, \quad k = 1, 2, \ldots, p-1,$$

$$\sigma_n^2 = 1,$$

and

$$f_1\left(x; \sigma^2, \nu\right) = \frac{\Gamma\left((\nu+1)/2\right)}{\sqrt{\pi}\,\sigma\Gamma\left(\nu/2\right)} \left[1 + \frac{1}{\nu}\frac{x^2}{\sigma^2}\right]^{-(\nu+1)/2}.$$

We shall denote by $F_1(x; \sigma^2, \nu)$ the cdf corresponding to f_1. For convenience denote $\mathbf{A}^{-1} = \mathbf{B} = b_{ij}$, where $\mathbf{B}$ is an upper triangular matrix with $b_{pp} = 1$ and $b_{jj} > 0$ for all j. Then, since the integral over $\mathbf{X}$ covers the region $\mathbf{X} \le \mathbf{X}_0$, the integral over $\mathbf{Z}$ is determined by the inequality $\mathbf{BZ} < \mathbf{X}_0$, and the bounds can be written as

$$z_p < x_{p0} \equiv z_{n0}$$

and

$$z_k < b_{kk}^{-1}\left(x_{k0} - \sum_{i=k+1}^{p} b_{ki} z_i\right) \equiv z_{k0}\left(x_{k0}, z_{k+1}, \ldots, z_p\right)$$

for $k = 1, 2, \ldots, p-1$. Utilizing this transformation, the probability p can be written as $p = J_p$, where

$$\begin{aligned} J_k &= \int_{-\infty}^{y_{k0}} f_1\left(z_k; \sigma_k^2, \nu_k\right) J_{k-1} dz_k \\ &= \int_{-\infty}^{z_{k0}} F_1\left(z_{k_0}; \sigma_k^2, \nu_k\right) J_{k-1} f_1^c\left(z_k; \sigma_k^2, \nu_k\right) dz_k \\ &= E_{f_1^c}\left[F_1\left(z_{k_0}; \sigma_k^2, \nu_k\right) J_{k-1}\right], \quad k = 2, 3, \ldots, p, \end{aligned} \tag{38}$$

where

$$f_1^c\left(z_k; \sigma_k^2, \nu_k\right) = \frac{f_1\left(z_k; \sigma_k^2, \nu_k\right)}{F_1\left(z_{k0}; \sigma_k^2, \nu_k\right)}$$

is the univariate unconditional t pdf for $z_k \le z_{k0}$ and $J_1 = F_1(z_{10}; \sigma_1^2, \nu_1)$. Hence, J_k is the probability over the range of $(z_1, \ldots, z_k)$ conditional on the values for $(z_{k+1}, \ldots, z_p)$.

The Monte Carlo simulation starts off by drawing random values of z_p from the distribution $f_1^c(\cdot; z_{p0}, \sigma_p^2, \nu_p)$, which we shall denote by $\tilde{z}_{p,r}$, $r = 1, \ldots, R$. Each of these yields a different bound $\tilde{z}_{p-1,0,r}$ and parameter value $\tilde{\sigma}_{p-1,r}^2$ for each draw of z_{p-1}; $\tilde{z}_{p-1,r}$ is then drawn from the distribution $f_1^c(\cdot; \tilde{z}_{p-1,0,r}, \tilde{\sigma}_{p-1,r}^2, \nu_{p-1})$. This process continues until $\tilde{z}_{2,r}$ is drawn and $\hat{J}_1 = F_1(\tilde{z}_{10,r}; \tilde{\sigma}_{1,r}^2, \nu_1)$ is computed with a commonly available approximation routine for the univariate Student's t cdf. The simulated estimate $\widehat{p}$ of p is then found as the sample average of the $\widehat{J}_p$ values across the simulated sample of R elements

$$\widehat{p} \equiv \widehat{J}_p = \frac{1}{R} \sum_{r=1}^{R} F_1 \left(\tilde{z}_{p0,r}; \tilde{\sigma}_{p,r}^2, \nu_p \right) \hat{J}_{p-1},$$

where $\widehat{J}_k = F_1(\tilde{z}_{k0,r}; \tilde{\sigma}_{k,r}^2, \nu_k) \hat{J}_{k-1}$ for $k = 2, \ldots, p-1$. It is more efficient to estimate J_p by averaging over a large number of elements than to obtain close approximations of its components J_k for $k < p$. Therefore, a better estimate for p is

$$\widehat{p} = \frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{k=1}^{p} F_1 \left(\tilde{z}_{k0,r}; \tilde{\sigma}_{k,r}^2, \nu_k \right) \right\}.$$

The right-hand side of (38) remains unchanged if the integrand is divided or multiplied by any nonzero function of z. Let g_p be a p-dimensional pdf such that

$$g_p(z; \nu) = \prod_{k=1}^{p} g_1 \left(z_k; \tau_k^2 \right),$$

where g_1 is a univariate pdf of a type to be mentioned below with $Var(z_k) = \tau_k^2 = \sigma_k^2 \nu_k / (\nu_k - 2)$, and σ_k^2 and ν_k are as defined above. Let $G_1(z_k; \tau_k^2)$ be the associated cdf, and let

$$g_1^c \left(z_k; z_{k0}, \tau_k^2 \right) = \frac{g_1 \left(z_k; \tau_k^2 \right)}{G_1 \left(z_{k0}; \tau_k^2 \right)}$$

be the conditional pdf. Finally, let

$$g_p^c \left(z; z_0, \nu \right) = \prod_{k=1}^{p} g_1^c \left(z_k; z_{k0}, \tau_k^2 \right).$$

With these definitions, one can write $p = J_p$ in terms of

$$J_k = \int_{-\infty}^{z_{k0}} \frac{f_1\left(z_k; \sigma_k^2, \nu_k\right)}{g_1^c\left(z_k; z_{k0}, \tau_k^2\right)} J_{k-1}\, g_1^c\left(z_k; z_{k0}, \tau_k^2\right) dz_k$$
$$= E_{g_1^c}\left[\frac{f_1\left(z_k; \sigma_k^2, \nu_k\right)}{g_1^c\left(z_k; z_{k0}, \tau_k^2\right)} J_{k-1}\right], \quad k = 2, \ldots, p,$$

and $J_1 = F(z_{10}; \sigma_1^2, \nu_1)$. Clearly, g_p^c, and, more particularly, g_1^c, is an important sampling density (see, for example, Hammersley and Handscomb [**30**]). To evaluate p, the procedure is as follows: Generate random drawings $\tilde{z}_{p,r}$ for $r = 1, \ldots, R$ from the distribution $g_1^c(\cdot; z_{n0}, \tau_n^2)$; compute the implied values $\tilde{z}_{p-1,0,r}$ and $\tilde{\tau}_{p-1,r}^2$ for each drawing of z_{p-1}; draw $\tilde{z}_{p-1,r}$ from the distribution $g_1^c(\cdot; \tilde{z}_{p-1,0,r}, \tilde{\tau}_{p-1,r}^2, \nu_p)$; and continue on until $\tilde{z}_{2,0,r}$ is drawn and $\widehat{J}_1$ is computed. Based on this procedure, $\widehat{p}$ may be written in the form

$$\widehat{p} = \frac{1}{R} \sum_{r=1}^{R} \left(F_1\left(\tilde{z}_{10,r}; \tilde{\tau}_{1,r}^2\right) \prod_{k=1}^{p} \frac{f_1\left(\tilde{z}_{k,r}; \tilde{z}_{k0,r}, \tilde{\sigma}_{k,r}^2, \nu_k\right)}{g_1^c\left(\tilde{z}_{k,r}; \tilde{z}_{k0,r}, \tilde{\tau}_{k,r}^2\right)} \right).$$

Three suitable choices for the importance density function are

- the logit with
$$g_1(x) = \frac{\lambda}{\tau} q(1-q),$$
where
$$q = \left[1 + \exp\left(-\lambda x/\tau\right)\right]^{-1}$$
and $\lambda = \pi/\sqrt{3}$;
- transformed beta (2, 2) density (Vijverberg [**62**]) with
$$g_1(x) = 6z^2(1-z)^2,$$
where
$$z = \frac{\exp\left(x/\sigma\right)}{1 + \exp\left(x/\sigma\right)};$$
- the normal $N(0, \sigma^2)$ density.

Vijverberg [**64, 65**] has developed a new family of simulators that extends the above research on the simulation of high-order probabilities. For instance, Vijverberg [**65**] has reported that the gain in precision using the new family translates into a 40% savings in computational time.

REFERENCES

1. M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.
2. R. A. Abusev and N.V. Kolegova, *On estimation of probabilities of linear inequalities for multivariate t distributions,* Journal of Mathematical Sciences **103** (2001), 542–546.
3. D. E. Amos, *Evaluation of some cumulative distribution functions by numerical evaluation,* SIAM Review **20** (1978), 778–800.
4. D. E. Amos and W. G. Bulgren, *On the computation of a bivariate t distribution*, Mathematics and Computation **23** (1969), 319–333.
5. R. E. Barlow, D. J. Bartholomew, J. M. Bremner and H. D. Brunk, *Statistical Inference under Order Restrictions*, John Wiley and Sons, Chichester, 1972.
6. R. Bohrer, *A multivariate t probability integral*, Biometrika **60** (1973), 647–654.
7. R. Bohrer and G. K. Francis, *Sharp one-sided confidence bounds over positive regions,* Annals of Mathematical Statistics **43** (1972), 1541–1548.
8. R. Bohrer, M. Schervish and J. Sheft, *Algorithm AS 184: Non-central studentized maximum and related multiple-t probabilities,* Applied Statistics **31** (1982), 309–317.
9. F. Bretz, A. Genz and L. A. Hothorn, *On the numerical availability of multiple comparison procedures,* Biometrical Journal **43** (2001), 645–656.
10. J. H. Cadwell, *The bivariate normal integral,* Biometrika **38** (1951), 475–481.
11. G. F. Corliss and L. B. Rall, *Adaptive, self-validating numerical quadrature,* SIAM Journal on Scientific and Statistical Computing **8** (1987), 831–847.
12. H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1951.
13. I. Deak, *Random Number Generators and Simulation*, Akademiai Kiado, Budapest, 1990.
14. C. W. Dunnett and M. Sobel, *A bivariate generalization of Student's t-distribution with tables for certain special cases,* Biometrika **41** (1954), 153–169.
15. J. E. Dutt, *A representation of multivariate normal probability integrals by integral transforms,* Biometrika **60** (1973), 637–645.
16. J. E. Dutt, *On computing the probability integral of a general multivariate t,* Biometrika **62** (1975), 201–205.
17. R. A. Fisher, *Expansion of Student's integral in power of* n^{-1}, Metron **5** (1925), 109.
18. Y. Fujikoshi, *Error bounds for asymptotic expansions of scale mixtures of distributions,* Hiroshima Mathematics Journal **17** (1987), 309–324.
19. Y. Fujikoshi, *Non-uniform error bounds for asymptotic expansions of scale mixtures of distributions,* Journal of Multivariate Analysis **27** (1988), 194–205.
20. Y. Fujikoshi, *Error bounds for asymptotic expansions of the maximums of the multivariate t- and F-variables with common denominator,* Hiroshima Mathematics Journal **19** (1989), 319–327.
21. Y. Fujikoshi, *Error bounds for asymptotic approximations of some distribution functions,* in Multivariate Analysis: Future Directions, (C. R. Rao, ed.), North-Holland, Amsterdam, 1993, 181–208.
22. Y. Fujikoshi and R. Shimizu, *Error bounds for asymptotic expansions of scale mixtures of univariate and multivariate distributions,* Journal of Multivariate Analysis **30** (1989), 279–291.
23. Y. Fujikoshi and R. Shimizu, *Asymptotic expansions of some distributions and their error bounds–the distributions of sums of independent random variables and scale mixtures,* Sugaku Expositions **3** (1990), 75–96.

24. A. Genz, *Numerical computation of the multivariate normal probabilities,* Journal of Computational and Graphical Statistics **1** (1992), 141–150.
25. A. Genz and F. Bretz, *Numerical computation of multivariate t probabilities with application to power calculation of multiple contrasts,* Journal of Statistical Computation and Simulation **63** (1999), 361–378.
26. A. Genz and F. Bretz, *Methods for the computation of multivariate t-probabilities,* Journal of Computational and Graphical Statistics (2001).
27. S. S. Gupta, *Probability integrals of multivariate normal and multivariate t,* Annals of Mathematical Statistics **34** (1963), 792–828.
28. S. S. Gupta, S. Panchapakesan and J. K. Sohn, *On the distribution of the studentized maximum of equally correlated normal random variables,* Communications in Statistics—Simulation and Computation **14** (1985), 103–135.
29. S. S. Gupta and M. Sobel, *On a statistic which arises in selection and ranking problems,* Annals of Mathematical Statistics **28** (1957), 957–967.
30. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods,* Methuen & Co. Ltd, London, 1964.
31. D. R. Jensen, *Closure of multivariate t and related distributions,* Statistics and Probability Letters **20** (1994), 307–312.
32. S. Joe, *Randomization of lattice rules for numerical multiple integration,* Journal of Computational and Applied Mathematics **31** (1990), 299–304.
33. S. John, *On the evaluation of the probability integral of the multivariate t distribution,* Biometrika **48** (1961), 409–417.
34. S. John, *Methods for the evaluation of probabilities of polygonal and angular regions when the distribution is bivariate t,* Sankhyā A **26** (1964), 47–54.
35. S. John, *On the evaluation of probabilities of convex polyhedra under multivariate normal and t distributions,* Journal of the Royal Statistical Society B **28** (1966), 366–369.
36. Z. Kopal, *Numerical Analysis*, Chapman and Hall, London, 1955.
37. S. Kotz, N. Balakrishnan and N. L. Johnson, *Continuous Multivariate Distributions*, *Volume 1: Models and Applications*, second edition John Wiley and Sons, New York, 2000.
38. S. Kotz, Y. Lumelskii and M. Pensky, *The Stress-Strength Model and Its Generalizations: Theory and Applications*, World Scientific Publishing Co., River Edge, New Jersey, 2003.
39. A. Kudo, *A multivariate analogue of the one-sided test,* Biometrika **50** (1963), 403–418.
40. S. Kunte and R. N. Rattihalli, *Rectangular regions of maximum probability content,* Annals of Statistics **12** (1984), 1106–1108.
41. C. Nicholson, *The probability integral for two variables,* Biometrika **33** (1943), 59–72.
42. D. B. Owen, *Tables for computing bivariate normal probabilities,* Annals of Mathematical Statistics **27** (1956), 1075–1090.
43. K. Pearson, *On non-skew frequency surfaces,* Biometrika **15** (1923), 231.
44. K. Pearson, *Tables for Statisticians and Biometricians*, Part II, Cambridge University Press for the Biometrika Trust, London, 1931.
45. S. J. Press and J. E. Rolph, *Empirical Bayes estimation of the mean in a multivariate normal distribution,* Communications in Statist. Theory and Methods **15** (1986), 2201–2228.
46. R. N. Rattihalli, *Regions of maximum probability content and their applications,* Ph.D. Thesis, University of Poona, India, 1981.
47. K. C. Seal, *On a class of decision procedures for ranking means,* Institute of Statistics Mimeograph Series No. 109, University of North Carolina at Chapel Hill, 1954.

48. R. Shimizu and Y. Fujikoshi, *Sharp error bounds for asymptotic expansions of the distribution functions for scale mixtures,* Annals of the Institute of Statistical Mathematics **49** (1997), 285–297.
49. I. H. Sloan and S. Joe, *Lattice Methods for Multiple Integration,* Clarendon Press, Oxford, 1994.
50. P. N. Somerville, *Simultaneous confidence intervals (General linear model),* Bulletin of the International Statistical Institute **2** (1993), 427–428.
51. P. N. Somerville, *Exact all-pairwise multiple comparisons for the general linear model,* in Proceedings of the 25th Symposium on the Interface, Computing Science and Statistics (1993), (Interface Foundation, Virginia), 352–356.
52. P. N. Somerville, *Simultaneous multiple orderings,* Technical Report TR-93-1 (1993), Department of Statistics, University of Central Florida, Orlando.
53. P. N. Somerville, *Multiple comparisons,* Technical Report TR-94-1 (1994), Department of Statistics, University of Central Florida, Orlando.
54. P. N. Somerville, *Multiple testing and simultaneous confidence intervals: Calculation of constants,* Computational Statistics and Data Analysis **25** (1997), 217–233.
55. P. N. Somerville, *A Fortran 90 program for evaluation of multivariate normal and multivariate-t integrals over convex regions,* Journal of Statistical Software (1998), http://www.stat.ucla.edu/journals/jss/v03/i04.
56. P. N. Somerville, *Numerical computation of multivariate normal and multivariate t probabilities over convex regions,* Journal of Computational and Graphical Statistics **7** (1998), 529–544.
57. P. N. Somerville, *Numerical evaluation of multivariate integrals over ellipsoidal regions,* Bulletin of the International Statistical Institute (1999).
58. P. N. Somerville, *Critical values for multiple testing and comparisons: one step and step down procedures,* Journal of Statistical Planning and Inference **82** (1999), 129–138.
59. P. N. Somerville, *Numerical computation of multivariate normal and multivariate t probabilities over ellipsoidal regions,* Journal of Statistical Software (2001), http://www.stat.ucla.edu/www.jstatsoft.org/v06/i08.
60. P. N. Somerville and F. Bretz, *Fortran 90 and SAS-IML programs for computation of critical values for multiple testing and simultaneous confidence intervals,* Journal of Statistical Software (2001), http://www.stat.ucla.edu/www.jstatsoft.org/v06/i05.
61. F. E. Steffens, *Power of bivariate studentized maximum and minimum modulus tests,* Journal of the American Statistical Association **65** (1970), 1639–1644.
62. W. P. M. Vijverberg, *Monte Carlo evaluation of multivariate normal probabilities,* Journal of Econometrics (1995).
63. W. P. M. Vijverberg, *Monte Carlo evaluation of multivariate Student's t probabilities,* Economics Letters **52** (1996), 1–6.
64. W. P. M. Vijverberg, *Monte Carlo evaluation of multivariate normal probabilities*, Journal of Econometrics **76** (1997), 281–307.
65. W. P. M. Vijverberg, *Rectangular and wedge-shaped multivariate normal probabilities,* Economics Letters **68** (2000), 13–20.
66. O. Wang and W. J. Kennedy, *Comparison of algorithms for bivariate normal probability over a rectangle based on self-validating results from interval analysis,* Journal of Statistical Computation and Simulation **37** (1990), 13–25.
67. O. Wang and W. J. Kennedy, *Application of numerical interval analysis to obtain self-validating results for multivariate probabilities in a massively parallel environment,* Statistics and Computing **7** (1997), 163–171.
68. H. P. Wynn and P. Bloomfield, *Simultaneous confidence bands for regression analysis* (with discussion), Journal of the Royal Statistical Society B **33** (1997), 202–217.

Department of Statistics, University of Nebraska, Lincoln, NE 68583
E-mail address: snadaraj@unlserve.unl.edu

Department of Engineering Management and Systems Engineering,
The George Washington University, Washington, D.C. 20052