NUMERICAL COMPUTATION OF MULTIVARIATE T-PROBABILITIES WITH APPLICATION TO POWER CALCULATION OF MULTIPLE CONTRASTS

ALAN GENZ[a] and FRANK BRETZ[b]

[a] *Department of Mathematics, Washington State University, Pullman, WA 99164-3113, USA;*
[b] *Department of Bioinformatics, University of Hanover, 30419 Hanover, Germany*

A new method to calculate the multivariate t-distribution is introduced. We provide a series of substitutions, which transform the starting q-variate integral into one over the $(q-1)$-dimensional hypercube. In this situation standard numerical integration methods can be applied. Three algorithms are discussed in detail. As an application we derive an expression to calculate the power of multiple contrast tests assuming normally distributed data.

Keywords: Multivariate t-distribution; Monte Carlo methods; lattice rule algorithm; calculation of power; multiple contrast test

1. INTRODUCTION

A problem, which frequently arises in statistical analysis, is to compute the standardised q-variate central t-distribution function

$$T(\boldsymbol{a},\boldsymbol{b}) = \frac{K_\nu^{(q)}}{\sqrt{|\boldsymbol{R}|}} \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_q}^{b_q}\left(1+\frac{\boldsymbol{x}^t\boldsymbol{R}^{-1}\boldsymbol{x}}{\nu}\right)^{-\frac{\nu+q}{2}} d\boldsymbol{x}. \qquad (1)$$

Here $K_\nu^{(q)} = \frac{\Gamma\left(\frac{\nu+q}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu\pi)^q}}$, $\boldsymbol{R}$ is a $q \times q$ positive definite correlation matrix, and $\boldsymbol{x} = \left(\boldsymbol{x_1}, \ldots, x_q\right)^t$. There are many applications. One example is the area of multiple comparisons, where selection and ranking procedures or many-to-one comparisons require the evaluation of $T(\boldsymbol{a},\boldsymbol{b})$ (see Hochberg and Tamhane (1987) for a comprehensive overview).

Until recently a direct evaluation of the multiple integral for arbitrary correlation matrices was considered computationally infeasible. Wang and Kennedy (1997) used interval analysis by applying the multivariate Taylor expansion to the density. But even low accuracy results required calculation times that were too large for practical purposes. Two different methodologies were developed by Somerville (1997, 1998). The first approach applies Monte Carlo techniques to a transformed multivariate t-integral. The second approach uses a binning procedure combined with an univariate quadrature method to estimate the integral. Dunnett (1955) reduced $T(\boldsymbol{a},\boldsymbol{b})$ to an inner q-variate normal integral combined with an outer chi integral. However, this method requires the evaluation of an additional integral. The advantage of this approach lies in the larger number of articles published about the multinormal integral (see, for example, Schervish, 1984, Solow, 1990, Genz, 1992, Joe, 1995, or Hajivassiliou et al., 1996). If $\boldsymbol{R}$ has a special structure (for example the product correlation structure $\rho_{i,j} = \lambda_i\lambda_j$, $-1 < \lambda_i, \lambda_j < 1$), the multivariate normal distribution reduces to a single integral over q multiplied univariate normal integrals. But since this approach is limited to problems with correlation matrices that have a special structure, we cannot use this approach

for the more general problems considered in this paper. In this context Hsu (1992) proposed to approximate a given arbitrary correlation matrix $\boldsymbol{R}$ by the 'closest' $\hat{\boldsymbol{R}}$, which satisfies the structural condition. This can be done by the use of factor analytic methods (Hsu, 1992) or by linear programming techniques (Hsu and Nelson, 1998). Finally, Edwards and Berry (1987) introduced crude Monte Carlo for equation (1) in order to calculate critical constants for different multiple comparison procedures.

Note that we could have defined $T(\boldsymbol{a}, \boldsymbol{b})$ by introducing the noncentral multivariate t-distribution. But we leave this generalisation for Section 5 and focus in Sections 2 - 4 on the alternate definition in central form without the noncentrality vector δ. In the following section a new method for evaluating $T(\boldsymbol{a},\boldsymbol{b})$ will be introduced. With this method, we transform the q-variate integrand into a product of univariate integrands. In Section 3, we describe algorithms for evaluating the transformed integral. An illustration of the use of this transformation by computing simultaneous confidence intervals is given in Section 4. In Section 5, we apply these methods to integrals required for power calculation of multiple contrast tests. We provide an expression in closed form and illustrate the numerical computation by an example.

2. TRANSFORMATION OF THE MULTIVARIATE T-INTEGRAL

We assume that in equation (1) for each i at least a_i or b_i is finite, because otherwise an appropriate transformation reduces the dimension of the problem by integrating the $i^{\underline{\text{th}}}$ variable explicitly. In a series of substitutions based on a talk held by Genz (1993) we transform $T(\boldsymbol{a}, \boldsymbol{b})$ to an integral over the (q - 1)-dimensional unit hypercube. The method is similar to the one used by Genz (1992) for the multivariate normal integral. In this form several numerical procedures can be used for the final evaluation, and some of these procedures are presented as algorithms in the next section.

Let $\boldsymbol{C}\boldsymbol{C}^{\mathrm{t}} = \boldsymbol{R}$ be the Cholesky decomposition for $\boldsymbol{R}$, where $\boldsymbol{C}$ is a lower triangular matrix. If we let $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, then $\boldsymbol{x}^{\mathrm{t}}\boldsymbol{R}^{-1}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{t}}\boldsymbol{C}^{\mathrm{t}}\boldsymbol{C}^{-\mathrm{t}}\boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{t}}\boldsymbol{y}$ and $\mathrm{d}\boldsymbol{x} = |\boldsymbol{C}|\mathrm{d}\boldsymbol{y} = |\boldsymbol{R}|^{½}\mathrm{d}\boldsymbol{y}$ and so

$$T(\boldsymbol{a},\boldsymbol{b}) = K_\nu^{(q)} \int_{a_1'}^{b_1'}\int_{a_2'}^{b_2'} \ldots \int_{a_q'}^{b_q'} \left(1+\frac{\boldsymbol{y}^t\boldsymbol{y}}{\nu}\right)^{-\frac{\nu+q}{2}} d\boldsymbol{y},$$

with $a_i' = \left(a_i - \sum_{j=1}^{i-1} c_{i,j}y_j\right)\Big/ c_{i,i}$ and $b_i' = \left(b_i - \sum_{j=1}^{i-1} c_{i,j}y_j\right)\Big/ c_{i,i}$. We now split the integrand into a product of q factors by use of

$$\left(1+\frac{y_1^2+y_2^2+\ldots+y_q^2}{\nu}\right)^{-\frac{\nu+q}{2}} = \left(1+\frac{y_1^2}{\nu}\right)^{-\frac{\nu+q}{2}} \cdot \left(1+\frac{y_2^2}{\nu+y_1^2}\right)^{-\frac{\nu+q}{2}} \cdot \ldots \cdot \left(1+\frac{y_q^2}{\nu+y_1^2+\ldots+y_{q-1}^2}\right)^{-\frac{\nu+q}{2}}.$$

When applied to T($\boldsymbol{a}$, $\boldsymbol{b}$) this transformation yields

$$T(\boldsymbol{a},\boldsymbol{b}) = K_\nu^{(q)} \int_{a_1'}^{b_1'} \left(1+\frac{y_1^2}{\nu}\right)^{-\frac{\nu+q}{2}} \ldots \int_{a_q'}^{b_q'} \left(1+\frac{y_q^2}{\nu+y_1^2+\ldots+y_{q-1}^2}\right)^{-\frac{\nu+q}{2}} d\boldsymbol{y}.$$

At this stage each of the y_i's can be substituted using $y_i = u_i\sqrt{\frac{\nu+\sum_{j=1}^{i-1} y_j^2}{\nu+i-1}}$. Beginning with $i = q$ the q-fold substitution results in

$$T(\boldsymbol{a},\boldsymbol{b}) = K_\nu^{(q)}\sqrt{\frac{\nu}{\nu+1}\cdot\frac{\nu}{\nu+2}\cdot\ldots\cdot\frac{\nu}{\nu+q-1}}\int_{\hat{a}_1}^{\hat{b}_1}\left(1+\frac{u_1^2}{\nu}\right)^{-\frac{\nu+1}{2}}\ldots\int_{\hat{a}_q}^{\hat{b}_q}\left(1+\frac{u_q^2}{\nu+q-1}\right)^{-\frac{\nu+q}{2}}d\boldsymbol{u}$$

$$= K_\nu^{(1)}\int_{\hat{a}_1}^{\hat{b}_1}\left(1+\frac{u_1^2}{\nu}\right)^{-\frac{\nu+1}{2}}\ldots K_{\nu+q-1}^{(1)}\int_{\hat{a}_q}^{\hat{b}_q}\left(1+\frac{u_q^2}{\nu+q-1}\right)^{-\frac{\nu+q}{2}}d\boldsymbol{u},$$

where $K_\nu^{(q)}\sqrt{\prod_{j=0}^{q-1}\frac{\nu}{\nu+j}} = \prod_{j=0}^{q-1}K_j^{(1)}$ follows from induction, and $\hat{a}_i = a_i'\sqrt{\frac{\nu+i-1}{\nu+\sum_{j=1}^{i-1}y_j^2}}$ and $\hat{b}_i = b_i'\sqrt{\frac{\nu+i-1}{\nu+\sum_{j=1}^{i-1}y_j^2}}$ are the new integration limits. For the last steps we set $u_i = t_{\nu+i-1}^{-1}(z_i)$, where $t_\nu(u) = K_\nu^{(1)}\int_{-\infty}^{u}\left(1+\frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}}ds$ denotes the univariate t-distribution with ν degrees of freedom. Using $dz_i = K_{\nu+i-1}^{(1)}\left(1+\frac{u_i^2}{\nu+i-1}\right)^{-\frac{\nu+i}{2}}du_i$ equation (1) becomes

$$\begin{aligned} T(\boldsymbol{a},\boldsymbol{b}) &= \int_{d_1}^{e_1}\int_{d_2}^{e_2}\ldots\int_{d_q}^{e_q}d\boldsymbol{z} \\ &= (e_1-d_1)\int_0^1(e_2-d_2)\ldots\int_0^1(e_q-d_q)\int_0^1 d\boldsymbol{w} \\ &= \underbrace{\int_0^1\int_0^1\ldots\int_0^1}_{q-1\text{ integrals}} f(\boldsymbol{w})\,d\boldsymbol{w}, \end{aligned} \qquad (2)$$

where $d_i = t_{\nu+i-1}(\hat{a}_i)$, $e_i = t_{\nu+i-1}(\hat{b}_i)$ and $z_i = d_i + w_i(e_i - d_i)$. With the sequence of transformations described here $T(\boldsymbol{a},\boldsymbol{b})$ becomes now an integral of $f(\boldsymbol{w}) = (e_1 - d_1)\cdot\ldots\cdot(e_q - d_q)$ over the (q - 1)-dimensional unit hypercube, which can be evaluated with different multidimensional numerical computation methods.

3. ALGORITHMS

In this section we present three numerical algorithms that use equation (2) to estimate $T(\boldsymbol{a},\boldsymbol{b})$ for a given error requirement ε. For the last algorithm, a lattice rule algorithm, the theoretical research is still ongoing. The other two algorithms are an acceptance-rejection sampling and a crude Monte Carlo algorithm. These methods are simpler and known to be reliable. Other Monte Carlo techniques could be applied but will not be analysed here. Deak (1990) gives a good overview with application to the multinormal case.

Acceptance-Rejection Method

This procedure generates q-dimensional uniform random vectors $\boldsymbol{w}_1$, ..., $\boldsymbol{w}_N$ and estimates T($\boldsymbol{a}$, $\boldsymbol{b}$) using

$$\overline{T}_{AR}(\boldsymbol{a},\boldsymbol{b}) = \frac{1}{N}\sum_{l=1}^{N} g(\boldsymbol{C}\boldsymbol{y}_l),$$

with $g(\boldsymbol{x}) = \begin{cases} 1, & a_i \le x_i \le b_i \\ 0, & \text{otherwise} \end{cases}$ and $y_{l,i} = t^{-1}_{\nu+i-1}(w_{l,i})\sqrt{\frac{\nu+\sum_{j=1}^{i-1} y_j^2}{\nu+i-1}}$, i = 1, ..., q, l = 1, ..., N. To control the simulated error we make use of the usual error estimate of the means

$$s_{\overline{T}_{AR}} = \frac{s}{\sqrt{N}} = \sqrt{\frac{\sum\left(g - \overline{T}_{AR}\right)^2}{N(N-1)}}.$$

Further we denote by γ the Monte Carlo confidence factor for the standard error. If, for example, γ = 3, we then expect the actual error of $\overline{T}_{AR}$ to be less than the error bound ε in 99.7% of the cases.

1. **INPUT** q, γ, ν, $\boldsymbol{R}$, $\boldsymbol{a}$, $\boldsymbol{b}$, ε.
2. Compute lower triangular Cholesky factor $\boldsymbol{C}$ for $\boldsymbol{R}$.
3. Initialise N = 0, *Intval* = 0, *Varsum* = 0.
4. **REPEAT**
   a) Generate uniform random w_1, ..., $w_q \in$ [0, 1].
   b) Set f = 1.
   c) **FOR** i = 1, 2, ..., q
      Set $y_i = t^{-1}_{\nu+i-1}(w_i)\sqrt{\frac{\nu+\sum_{j=1}^{i-1} y_j^2}{\nu+i-1}}$.
      IF $\sum_{j=1}^{i} c_{i,j} y_j < a_i$ **OR** $\sum_{j=1}^{i} c_{i,j} y_j > b_i$ **THEN** go to step d).
   END FOR
   Go to step e).
   d) Set f = 0.
   e) Set $N = N + 1$,
   Varsum = *Varsum* + (N - 1)(f - *Intval*)^2/N,
   Intval = *Intval* + (f - *Intval*)/N,
   $ErrEst = \gamma\sqrt{Varsum/(N(N-1))}$.
   UNTIL *ErrEst* < ε.
5. **OUTPUT** *Intval*, *ErrEst*, N.

The acceptance-rejection method is widely used and may be the most intuitive way to deal with equation (2). Deak (1990), however, showed that among the various Monte Carlo methods it is the one with the worst efficiency and therefore other approaches to evaluate T($\boldsymbol{a}$, $\boldsymbol{b}$) are needed.

Monte Carlo Method

For estimating (2) by the crude Monte Carlo method we let $\boldsymbol{w}_1$, ..., $\boldsymbol{w}_N$ be uniformly and independently distributed on [***0, 1***]$^{q-1}$. Then the random variables $f(\boldsymbol{w}_l)$, l = 1, ..., N, are independent and their expected value is $E[f(\boldsymbol{w})] = \int_0^1 \dots \int_0^1 f(\boldsymbol{w})d\boldsymbol{w} = T(\boldsymbol{a},\boldsymbol{b})$. Consequently the arithmetic average

$$\bar{T}_{MC}(\boldsymbol{a},\boldsymbol{b}) = \frac{1}{N}\sum_{l=1}^{N} f(\boldsymbol{w}_l) \tag{3}$$

is an unbiased estimator of the integral $T(\boldsymbol{a},\boldsymbol{b})$.

The Monte Carlo and Lattice rule algorithms that we describe in this section both require the evaluation of $f(\boldsymbol{w})$ for particular values of $\boldsymbol{w}$, so we give here an algorithm for $f(\boldsymbol{w})$ that will be used by both of the numerical integration algorithms. In the following algorithm note that the initializations of d_1 and e_1 are required only for the first evaluation of $f(\boldsymbol{w})$ (assuming $\boldsymbol{a}$, $\boldsymbol{b}$, ν, and $\boldsymbol{C}$ are fixed for a particular integral), and we would avoid wasteful computation of t-values be setting $d_i = 0$ or $e_i = 1$ if $a_i = -\infty$ or $b_i = \infty$, respectively.

1. **INPUT** $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{w}, \nu, \boldsymbol{C}$.
2. Initialise $d_1 = t_\nu(a_1/c_{1,1})$, $e_1 = t_\nu(b_1/c_{1,1})$, $f_1 = e_1 - d_1$.
3. **FOR** $i = 1, 2, ..., q - 1$

   Set $y_i = t_{\nu+i-1}^{-1}\left(d_i + w_i(e_i - d_i)\right)\sqrt{\frac{\nu+\sum_{j=1}^{i-1} y_j^2}{\nu+i-1}}$,

   $d_{i+1} = t_{\nu+i}\left(\left(a_{i+1} - \sum_{j=1}^{i} c_{i+1,j} y_j\right)\sqrt{\frac{\nu+i}{\nu+\sum_{j=1}^{i} y_j^2}} \Big/ c_{i+1,i+1}\right)$,

   $e_{i+1} = t_{\nu+i}\left(\left(b_{i+1} - \sum_{j=1}^{i} c_{i+1,j} y_j\right)\sqrt{\frac{\nu+i}{\nu+\sum_{j=1}^{i} y_j^2}} \Big/ c_{i+1,i+1}\right)$,

   $f_{i+1} = (e_{i+1} - d_{i+1}) f_i$.

   END FOR
4. **OUTPUT** $f(\boldsymbol{w}) = f_q$.

In the following we give the algorithm when applying crude Monte Carlo on equation (2).

1. **INPUT** q, γ, ν, $\boldsymbol{R}$, $\boldsymbol{a}$, $\boldsymbol{b}$, ε.
2. Compute lower triangular Cholesky factor $\boldsymbol{C}$ for $\boldsymbol{R}$.
3. Initialise $N = 0$, $Intval = 0$, $Varsum = 0$.
4. **REPEAT**
   a) Generate uniform random $w_1, ..., w_{q-1} \in [0, 1]$.
   b) Evaluate $f_q = f(\boldsymbol{w})$.
   c) Set $N = N + 1$,
   $Varsum = Varsum + (N - 1)(f_q - Intval)^2/N$,
   $Intval = Intval + (f_q - Intval)/N$,
   $ErrEst = \gamma\sqrt{Varsum/(N(N-1))}$.

   UNTIL $ErrEst < \varepsilon$.
5. **OUTPUT** *Intval*, *ErrEst*, *N*.

Randomised Lattice Rule Method

As seen from the central limit theorem, the Monte Carlo integration yields a probabilistic error bound in $O(N^{-½})$ (note that the order of magnitude does not depend on q). This means halving the error requires quadrupling the number of sample points. However, Quasi-Monte Carlo integration methods use sequences of nodes that are designed to be more uniform than random, while still using an integration formula with equal weights similar to (3). Niederreiter (1992) and others showed that under suitable conditions a deterministic error bound is given by $O(N^{-1}\log^{q-1}N)$. In the sequel we will make use of one approach out of this broad class, the so-called randomised lattice rule (Joe, 1990; Sloan and Joe, 1994, p.170)

$$\overline{T}_L = \frac{1}{N}\sum_{l=1}^{N} T_{L,l} = \frac{1}{Np}\sum_{l=1}^{N}\sum_{j=1}^{p} f\left(\left|2\left\{\frac{j}{p}\boldsymbol{z} + \boldsymbol{w}_l\right\} - 1\right|\right). \tag{4}$$

Here N is the simulation size, usually being very small (e.g. 10 - 20), p corresponds to the fineness of the lattice and $\boldsymbol{z} \in \mathbb{R}^{q-1}$ denotes the strategically chosen lattice vector. Braces around vectors indicate that each component has to be replaced by its fractional part. Finally, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ denote again $[\boldsymbol{0}, \boldsymbol{1}]^{q-1}$-uniform random vectors. The error estimate for $\overline{T}_L$ is

$$s_{\overline{T}_L} = \sqrt{\frac{\sum\left(T_{L,l} - \overline{T}_L\right)^2}{N(N-1)}}.$$

For more details concerning lattice rules in general and the approach used here the reader is referred to the book of Sloan and Joe (1994).

The algorithm listed below consists basically of two loops. The *l*-loop stands for the outer sum in equation (4), and *j* is the variable of the inner sum. This is repeated several times until a pre-assigned error level ε is reached. At each step the counting variable *n* is incremented by 1 and the number of lattice points $p = p_n$ is increased in dependence on *n*.

1. **INPUT** q, γ, ν, $\boldsymbol{R}$, $\boldsymbol{a}$, $\boldsymbol{b}$, ε.
2. Compute lower triangular Cholesky factor $\boldsymbol{C}$ for $\boldsymbol{R}$.
3. Initialise $N = 10$, $n = 0$.
4. **REPEAT**
   a) Set $n = n + 1$, *Intval* = 0, *Varsum* = 0.
   b) **FOR** $l = 1, 2, \ldots, N$
      i) Set *Latsum* = 0.
      ii) Generate uniform random $w_1, \ldots, w_{q-1} \in [0, 1]$.
      iii) **FOR** $j = 1, 2, \ldots, p_n$
      - Set $\boldsymbol{w} = \left|2\left\{\boldsymbol{w} + \frac{j}{p_n}\boldsymbol{z}\right\} - 1\right|$.
      - Evaluate $f_q = f(\boldsymbol{w})$.
      - Set *Latsum* = *Latsum* + (f_q - *Latsum*)/*j*.

      END FOR
      iv) Set *Varsum* = *Varsum* + (*l* - 1)(*Latsum* - *Intval*)^2/*l*,
      Intval = *Intval* + (*Latsum* - *Intval*)/*l*.

   END FOR
   c) Set $ErrEst = \gamma\sqrt{Varsum/(N(N-1))}$.

   UNTIL *ErrEst* < ε.
5. **OUTPUT** *Intval*, *ErrEst*, Np_n.

The best choice of $\boldsymbol{z}$ is still an open research question. Several proposals have been published in the literature, the most common of which is to choose $\boldsymbol{z}$ of the form $\boldsymbol{z}(h) = (1, h, h^2 \bmod p, \ldots, h^{q-2} \bmod p)$, $1 \le h \le \lfloor p/2 \rfloor$, proposed by Korobov (1960). This leaves us with the problem of how to choose *h*. The method we used minimises

$$P_{q,p} = \frac{1}{p}\sum_{j=1}^{p}\prod_{i=1}^{q-1}\tilde{F}\left(\left\{\frac{j}{p}z_i\right\}\right),$$

with $\widetilde{F}(x) = F(x)/4 + ¾$ and $F(x) = 1 + 2\pi^2(x^2 - x + 1/6)$, $x \in [0, 1]$ (see Sloan and Joe, 1994, p. 173). A SAS/IML-program, which implements the lattice rule algorithm presented above together with an error estimation is available from the website with URL http://www.bioinf.uni-hannover.de/mcp_home/software.html. Fortran source codes for all three algorithms and an adaptive integration algorithm and an algorithm based on transforming (1) into a spherical coordinate system similar to the system used by Deak (1990) for multivariate normal integrals can be found the website with URL http://www.sci.wsu.edu/math/faculty/genz/homepage.

4. COMPUTATION OF SIMULTANEOUS CONFIDENCE INTERVALS

In order to illustrate the use of the transformation method presented in Section 2 for evaluating multivariate t-probabilities, we consider one important application, the computation of simultaneous confidence intervals (see, for example, Edwards and Berry, 1987, or Hsu and Nelson, 1998). An example of the use of t-probability confidence intervals will be described in detail in the next section, where one-sided intervals are required. With these problems, we are given a desired confidence level α and we need to determine an upper limit vector $\boldsymbol{t} = (t, ..., t) \in \mathbb{R}^q$, so that $T(-\infty, \boldsymbol{t}) = 1 - \alpha$. We will illustrate the use of our algorithms with one of the problems to be considered in the next section (the Dunnett contrast case), which has q = 3, and a correlation matrix $\boldsymbol{R}$ given by

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0.3636 & 0.3636 \\ 0.3636 & 1 & 0.3636 \\ 0.3636 & 0.3636 & 1 \end{pmatrix}.$$

If we want to apply our algorithms on this problem, we need to combine our algorithms with a iterated-nonlinear-equation-solving algorithm. We let $h(\boldsymbol{t}) = T(-\infty, \boldsymbol{t}) - 1 + \alpha$, so that we need to find $\boldsymbol{t}$ such that $h(\boldsymbol{t}) = 0$. We have successfully used various modified secant algorithms for this problem. The results given below were produced by the Pegasus method (see Ralston and Rabinowitz, 1978, p. 341). Using this algorithm we need a starting interval $[\boldsymbol{a}, \boldsymbol{b}]$ which contains our desired solution $\boldsymbol{t}$, where $\boldsymbol{a} = (a, \ldots, a)$ and $\boldsymbol{b} = (b, \ldots, b)$. At each stage an estimated $\boldsymbol{t}$ is computed, along with $h(\boldsymbol{t})$, and a new interval is produced. In order to simplify the start of the algorithm we used $a = -4$ and $b = 4$, assuming $h(\boldsymbol{a}) = \alpha - 1$ and $h(\boldsymbol{b}) = \alpha$. The following Table I shows the behaviour of our Lattice rule algorithm ($\varepsilon_L = 0.001$, $\gamma = 3$) combined with the Pegasus method with termination if $|\boldsymbol{b} - \boldsymbol{a}| < 0.01 = \varepsilon_P$ in each of the q components, for $\alpha = 0.05$.

⇒ insert **TABLE I** about here

The final t value was $t = 2.1664$, apparently correct to at least four digits, and this required a total of 22144 $f(\boldsymbol{w})$ values.

A commonly used algorithm for the confidence interval problem is a type of rejection algorithm (see Edwards and Berry, 1987), which we will call %-rejection. The basic idea is to generate a large number, say m, of random vectors $\boldsymbol{x}_l \sim T(.,.)$ for given correlation matrix $\boldsymbol{R}$ and degrees of freedom ν. For each $\boldsymbol{x}_l$ let $t_l = \max_{1 \le i \le q}(x_{li})$. Then sort the t_l's and let $t = t_{(r)}$,

where $r = (m+1)(1-\alpha)$. The result is that after sorting we reject $(1-\alpha)\%$ of the (smallest) t_l's and pick the smallest one left as an estimator for t. Several refinements to this basic algorithm have been suggested (see Hsu and Nelson, 1990, 1998, for discussion and further references), but we have found that rejection type algorithms are generally not very efficient for computing multivariate t-probabilities, and it is beyond the scope of this paper to provide a detailed comparison between our algorithms and %-rejection algorithms for this problem. However, we will describe how our algorithms could be used for %-rejection, too. The $\boldsymbol{x}_l$ vectors needed for %-rejection are typically generated using a combination of multivariate normal and chi variables, but our transformations can also be used to generate $\boldsymbol{x}_l$ vectors with correct distribution. The acceptance-rejection algorithm given in the last section can be easily modified to provide a simple %-rejection algorithm

1. **INPUT** m, α, ν, $\boldsymbol{R}$.
2. Compute lower triangular Cholesky factor $\boldsymbol{C}$ for $\boldsymbol{R}$.
3. **FOR** l = 1, 2, ..., m
   a) Generate uniform random w_1, ..., $w_q \in [0, 1]$.
   b) **FOR** i = 1, 2, ..., q

      Set $y_i = t^{-1}_{\nu+i-1}(w_i)\sqrt{\frac{\nu+\sum_{j=1}^{i-1} y_j^2}{\nu+i-1}}$,

      $x_{li} = \sum_{j=1}^{i} c_{i,j} y_j$.

      END FOR
   c) Set $t_l = \max_{1 \le i \le q}(x_{li})$.
4. Sort the t_l's.
5. **OUTPUT** $t = t_{(r)}$, $r = (m+1)(1-\alpha)$.

Using this algorithm with m = 22139, so that $r = (m+1)(1-\alpha) = 21033$, we found $t = 2.1725$. A standard 99% confidence interval obtained using a normal approximation to the binomial distribution is given by $\left[t_{(20995)}, t_{(21072)}\right] = [2.1523, 2.1931]$.

5. APPLICATION TO MULTIPLE CONTRAST TESTS

In the general many-to-one context of comparing several treatments with a control the following one-way layout model is widely used, consisting of one control group, for example a placebo or a negative control, and k treatments, labelled 0, 1, ..., k respectively. Let $\left\{X_{i,j},\, i = 0, \ldots, k,\; j = 1, \ldots, n_i\right\}$ be the sample values, identically and independently N(μ_i, σ^2) distributed. Denote by $\overline{X}_i$ the sample mean $\sum_j X_{i,j} \big/ n_i$ and by $s^2 = \frac{\sum_{i=0}^{k}\sum_{j=1}^{n_i}\left(X_{i,j}-\overline{X}_i\right)^2}{\nu}$ the pooled variance estimator with $\nu = \sum_i n_i - (k+1)$ degrees of freedom. For testing $H_0: \mu_0 = \mu_1 = \ldots = \mu_k$ the following two alternatives are frequently considered. On the one hand the (one-sided) simple tree ordering is given by

$$H_A{}^1: \exists i: \mu_0 < \mu_i,\ i = 1, ..., k.$$

But in practice it often occurs that several dose groups have to be compared with a negative control group, whereas a monotonous dose-response trend can be assumed a-priori. In this case, the total order restricted alternative

$$H_A{}^2\colon \mu_0 \le \mu_1 \le \ldots \le \mu_k,\ \mu_0 < \mu_k,$$

may be more appropriate. However, caution can be advisable, if one is not sure, whether this kind of underlying dose-response relationship really holds. Bauer (1997) showed that already small departures from this assumption may lead to inappropriate results if trend tests designed for such situations are used. These tests, such as the ones presented in the sequel, then are not useful in the sense that they do not control the probability of incorrectly declaring a dose to be effective when it is not.

We now consider the statistic

$$T^{MC} = \max\{T^1, \ldots, T^q\} \tag{5}$$

of multiple contrast tests according to Mukerjee et al. (1986, 1987), Westfall (1997) and Hothorn et al. (1997). This is just the maximum over q single contrasts tests

$$T = \frac{\sum_{i=0}^{k} c_i \overline{X}_i}{s\sqrt{\sum_{i=0}^{k} c_i^2 / n_i}} \propto t_\nu,$$

where t_ν denotes the central univariate t-distribution and c_i the contrast coefficients under the condition $\sum_i c_i = 0$. Besides this, the choice of the c_i's is free and numerous proposals concerning their optimal choice have been published (see, for example, Tamhane et al., 1996).

To test H_0 versus $H_A{}^1$ one can use, for example, the one-sided many-to-one test of Dunnett (1955), which can be formulated as a multiple contrast test with the weights

$$\begin{pmatrix} -1 & 0 & \ldots & 0 & 1 \\ -1 & 0 & \ldots & 1 & 0 \\ & \vdots & & & \\ -1 & 1 & \ldots & 0 & 0 \end{pmatrix}.$$

For the total restricted case $H_A{}^2$ Bretz and Hothorn (1998) developed a multiple contrast approach based on Williams' test (1971), using in the balanced set-up the k contrasts

$$\begin{pmatrix} -1 & 0 & \ldots & 0 & 1 \\ -1 & 0 & \ldots & \frac{1}{2} & \frac{1}{2} \\ & \vdots & & & \\ -1 & \frac{1}{k} & \ldots & \frac{1}{k} & \frac{1}{k} \end{pmatrix}.$$

As each T^i in formula (5) is univariate t-distributed, their joint distribution will follow per definition (Cornish, 1954; Dunn and Massey, 1965) a central q-variate t-distribution with ν degrees of freedom and correlation matrix $\boldsymbol{R} = \{\rho_{l,m}\}_{l,m}$ with entries of the form

$$\rho_{l,m} = \frac{\sum_{i=0}^{k} c_{l,i} c_{m,i} / n_i}{\sqrt{\left(\sum_{i=0}^{k} c_{l,i}^2 / n_i\right)\left(\sum_{i=0}^{k} c_{m,i}^2 / n_i\right)}}, \qquad l, m = 1, \ldots, q.$$

The power for the multiple contrast test can now be derived as follows:

$$
\begin{aligned}
\text{Power} &= \Pr\left(T^{MC} \geq t \,\middle|\, H_A\right) \\
&= \Pr\left(\max_{1\leq l\leq q} \left(\frac{\sum_{i=0}^{k} c_{l,i}\,\overline{X}_i}{s\sqrt{\sum_{i=0}^{k} c_{l,i}^2 / n_i}} \right) \geq t \,\middle|\, H_A \right) \\
&= 1 - \Pr\left(\max_{1\leq l\leq q} \left(\frac{\sum_{i=0}^{k} c_{l,i}\,\overline{X}_i}{s\sqrt{\sum_{i=0}^{k} c_{l,i}^2 / n_i}} \right) < t \,\middle|\, H_A \right) \\
&= 1 - \Pr\left(\frac{\sum_{i=0}^{k} c_{1,i}\,\overline{X}_i}{s\sqrt{\sum_{i=0}^{k} c_{1,i}^2 / n_i}} < t \;\wedge\; \ldots \;\wedge\; \frac{\sum_{i=0}^{k} c_{q,i}\,\overline{X}_i}{s\sqrt{\sum_{i=0}^{k} c_{q,i}^2 / n_i}} < t \,\middle|\, H_A \right) \\
&= 1 - \Pr\left(\frac{\frac{\sum_{i=0}^{k} c_{1,i}\left(\overline{X}_i - \mu_i\right)}{\sigma\sqrt{\sum_{i=0}^{k} c_{1,i}^2 / n_i}} + \frac{\sum_{i=0}^{k} c_{1,i}\mu_i}{\sigma\sqrt{\sum_{i=0}^{k} c_{1,i}^2 / n_i}}}{s/\sigma} < t \;\wedge\; \ldots \;\wedge\; \frac{\frac{\sum_{i=0}^{k} c_{q,i}\left(\overline{X}_i - \mu_i\right)}{\sigma\sqrt{\sum_{i=0}^{k} c_{q,i}^2 / n_i}} + \frac{\sum_{i=0}^{k} c_{q,i}\mu_i}{\sigma\sqrt{\sum_{i=0}^{k} c_{q,i}^2 / n_i}}}{s/\sigma} < t \right).
\end{aligned}
\tag{6}
$$

The probability term of equation (6) is *q*-variate noncentral t-distributed with the noncentrality parameter $\delta = \left(\frac{\sum_{i=0}^{k} c_{l,i}\mu_i}{s\sqrt{\sum_{i=0}^{k} c_{l,i}^2 / n_i}} \right)_{1\leq l\leq q}$ and ν degrees of freedom. We denote the corresponding multivariate distribution function by $T_{\delta,\boldsymbol{R},\nu}(-\infty, \boldsymbol{t})$, where a_l = -∞ $\forall l$. In equation (6) $\boldsymbol{t} = (t, \ldots, t)$ stands for the (1 - α)-equipercentage point of the central *q*-variate t-distribution, so that $T(-\infty, \boldsymbol{t}) = 1 - \alpha$, where α is the pre-specified type-I-error. We now make use of the relationship

$$
\Pr\left(\boldsymbol{t}'_\nu \leq \boldsymbol{b}\right) = \Pr\left(\boldsymbol{U} \leq \boldsymbol{b}\chi_\nu / \sqrt{\nu} - \delta\right).
$$

Here $\boldsymbol{U}$ and χ_ν are independent random variables distributed as standardised *q*-variate normal and chi with ν degrees of freedom, $\boldsymbol{t}'_\nu = \boldsymbol{t}'_\nu(\delta) = \left(\boldsymbol{U} + \delta\right) \Big/ \left(\chi_\nu / \sqrt{\nu}\right)$. This leads to the explicit representation of the noncentral *q*-variate t-distribution function $T_{\delta,\boldsymbol{R},\nu}(-\infty, \boldsymbol{b})$ by

$$
\Pr\left(\boldsymbol{t}'_\nu \leq \boldsymbol{b}\right) = \frac{1}{2^{\frac{\nu}{2}-1}\,\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty x^{\nu-1} e^{-\frac{x^2}{2}}\, \Phi_q\left(\boldsymbol{b}x/\sqrt{\nu} - \delta; \boldsymbol{0}, \boldsymbol{R}\right) dx,
\tag{7}
$$

where $\boldsymbol{\Phi}_q(.)$ is the *q*-variate normal integral with expectation $\boldsymbol{0}$, correlation matrix $\boldsymbol{R}$ and upper integration bound $\boldsymbol{b}x/\sqrt{\nu} - \delta$. Note that above formulae may be applied to both alternatives $H_A{}^1$ or $H_A{}^2$, whereas the first case for Dunnett's many-to-one test was already considered by several authors, e.g. Bristol (1989) or Hayter and Liu (1992).

A correct evaluation of the probability term in equation (6) therefore consists of the following numerical problems. First we need the ability of computing the noncentral *q*-variate t-distribution function $T_{\delta,\boldsymbol{R},\nu}(-\infty, \boldsymbol{b})$. This can be done by using the transformation indicated in equation (7). For the evaluation of the arising *q*-variate standard normal distribution $\boldsymbol{\Phi}_q(.)$ several methods from the literature can be applied. Some of them are quoted in the introductory Section 1. The remaining single integral can be computed by using standard univariate integration techniques. Finally, we need to compute the critical value $\boldsymbol{t}$ in equation

(6). The vector is the implicit solution of $h(\boldsymbol{t}) = 0$, where $h(\boldsymbol{t}) = T(-\infty, \boldsymbol{t}) - 1 + \alpha$. Therefore, the computation reduces to the problem discussed in Section 4 and the Pegasus method (or the %-rejection algorithm) stated there can be applied. The evaluation of $h(\boldsymbol{t})$, on the other hand, requires the computation of $T(\boldsymbol{a}, \boldsymbol{b})$, what can be done by using the algorithms provided in Section 3.

To illustrate the ideas discussed above, suppose k = 3 groups are to be compared with a negative control, for example increasing dose levels of a certain drug, where the control corresponds to the zero dose. Frequently in this situation it is possible to assume monotonicity a-priori from the experimenter point of view, i.e., if any effect exists, then it will increase with respect to the treatment levels. Therefore, the goal is to assess, whether a global trend exists. The interesting alternative is then chosen to be H_A^2. For this alternative, however, we do have the problem of defining appropriate contrast matrices and corresponding test statistics. While no general rule seems to be available yet, we compare some of the widely used single and multiple contrast tests throughout the literature for different dose-response relationships.

For experimental reasons we let the practical relevance shift be δ = 1 and the size be set to α = 0.05. The variances of all 4 groups are assumed to be homogeneous and estimated from preceding pilot studies as σ^2 = 1. For simplicity we allocate the total sample size of 38 experimental units roughly according to the widely used $\sqrt{k}$ -rule, resulting in the sample size vector (14, 8, 8, 8). The power, or, equivalently, the type-II-error, was calculated using a SAS/IML program based on the algorithms presented here. The results with an estimated error of 10^{-4} are summarised in Table II.

$\Rightarrow$ insert **TABLE II** about here

From the computations we can see that there is a strong dependency of the power of tests on the underlying dose-response shape. With increasing number of contrasts covering more possible dose-response shapes, the problem is reduced but not alleviated. Hence, it would be desirable to have some measure of goodness for a pre-chosen (multiple) contrast test. But even if Hayter and Liu (1992) succeeded in determining the least favourable configuration for Dunnett's two-sided test in the partial balanced set-up, yielding the unweighted minimum power over all shapes, in the general case no solution exists. Several other aspects, such as the number of contrasts, sample size allocation or the ratio variance/relevance shift, influence the power function in complicated fashion as well. So, even if we managed to calculate the power of a multiple contrast in a specific situation, little can be assessed about their optimal choice.

References

Bauer, P. (1997) A note on multiple testing procedures in dose finding. *Biometrics*, **53**, 1125-1128.

Bretz, F. and Hothorn, L. (1998) A powerful alternative to Williams test with application to toxicological dose-response relationships of normally distributed data. *Environmental and Ecological Statistics* (in press).

Bristol, D.R. (1989) Designing clinical trials for two-sided multiple comparisons with a control. *Controlled Clinical Trials*, **10**, 142-152.

Cornish, E.A. (1954) The multivariate t-distribution associated with a set of normal sample deviates. *Australian Journal of Physics*, **7**, 531-542.

Deak, I. (1990) *Random number generators and simulation.* Akademiai Kiado, Budapest.

Dunn, O.J. and Massey, F.J. (1965) Estimation of multiple contrasts using t-distributions. *Journal of the American Statistical Association,* **60,** 573-583.

Dunnett C. W. (1955) A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association,* **50,** 1096-1121.

Edwards, D., and Berry, J. (1987) The efficiency of simulation-based multiple comparisons. *Biometrics*, **43**, 913-928.

Genz, A. (1992) Numerical computation of the multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, **1**, 141-150.

Genz, A. (1993) Numerical computation of the multivariate t probabilities. Talk held on SIAM Simulation and MC Methods, San Francisco.

Hajivassiliou, V., McFadden, D. and Ruud, P. (1996) Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results. *Journal of Econometrics,* **72**, 85-134.

Hayter, A.J. and Liu, W. (1992) A method of power assessment for tests comparing several treatments with a control. *Communications in Statistics - Theory and Methods*, **21**, 1871-1889.

Hochberg, Y. and Tamhane, A. (1987) *Multiple comparison procedures.* Wiley, New York.

Hothorn, L.A., Neuhäuser, M. and Koch, H.F. (1997) Analysis of randomised dose-finding studies: closure test modifications based on multiple contrast tests. *Biometrical Journal,* **39,** 467-479.

Hsu, J.C. (1992) The factor analytic approach to simultaneous inference in the general linear model. *Journal of Computational and Graphical Statistics*, **1**, 151-168.

Hsu, J.C. and Nelson, B.L. (1990) Control variates for quantile estimation. *Management Science*, **36**, 835-951.

Hsu, J.C. and Nelson, B.L. (1998) Multiple comparisons in the general linear model. *Journal of Computational and Graphical Statistics*, **7**, 23-41.

Joe, H. (1995) Approximations to multivariate normal rectangle probabilities based on conditional expectations. *Journal of the American Statistical Association*, **90**, 957-64.

Joe, S. (1990) Randomization of lattice rules for numerical multiple integration. *Journal of Computational and Applied Mathematics,* **31**, 299-304.

Korobov, N.M. (1960) Properties and calculation of optimal coefficients. *Soviet Math. Doklady,* **1**, 696-700.

Mukerjee, H., Robertson, T. and Wright F. T. (1986) Multiple contrast tests for testing against a simple tree ordering. In *Advances in order restricted statistical inference,* Dykstra, R (ed), Springer, Berlin, pp. 203-230.

Mukerjee, H., Roberston, T. and Wright, F. T. (1987) Comparison of several treatments with a control using multiple contrasts. *Journal of the American Statistical Association,* **82,** 902-910.

Niederreiter, H. (1992) *Random number generation and quasi-Monte Carlo methods.* Society for Industrial and Applied Mathematics, Philadelphia.

Ralston, A. and Rabinowitz, P. (1978) *A first course in numerical analysis*. McGraw-Hill, New York.

Schervish, M.J. (1984) Algorithm AS 195: multivariate normal probabilities with error bound. *Applied Statistics,* **33**, 81-94. Correction: **34** (1985), 103-104.

Sloan, I.H. and Joe, S. (1994) *Lattice methods for multiple integration.* Clarendon Press, Oxford.

Solow, A.R. (1990) A method for approximating multivariate normal orthant probabilities. *Journal of Statistical Computation and Simulation*, **37**, 225-229.

Somerville, P. (1997) Multiple testing and simultaneous confidence intervals: calculation of constants. Computational Statistics & Data Analysis, **25**, 217-233.

Somerville, P. (1998) Numerical computation of multivariate normal and multivariate t-probabilities. *Journal of Computational and Graphical Statistics* (submitted).

Tamhane, A.C., Dunnett, C. W. and Hochberg, Y. (1996) Multiple test procedures for dose finding. *Biometrics,* **52,** 21-37.

Wang, O. and Kennedy, W.J. (1997) Application of numerical interval analysis to contain self-validating results for multivariate probabilities in a massively parallel environment. *Statistics and Computing*, **7**, 163-171.

Westfall, P.H. (1997) Multiple testing of general contrasts using logical constraints and correlations. *Journal of the American Statistical Association,* **92,** 299-306.

Williams, D. A. (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics,* **27,** 103-117.

TABLE I Iteration steps for the Pegasus method using the lattice rule algorithm $(\varepsilon_P = 0.01)$.

Step i	a_i	t_i	b_i	$h(t_i)$	Number of $f(\boldsymbol{w})$ values
1	–4.0000	3.6000	4.0000	0.0485	2768
2	–4.0000	2.9045	3.6000	0.0409	2768
3	–4.0000	1.9709	2.9045	–0.0244	2768
4	1.9709	2.3196	2.9045	0.0140	2768
5	1.9709	2.1924	2.3196	0.0027	2768
6	1.9709	2.1669	2.1924	0.0001	2768
7	1.9709	2.1664	2.1669	0.0000	2768
8	1.9709	2.1664	2.1664	0.0000	2768
final	2.1664	2.1664	2.1664		22144

TABLE II Power of several contrast tests for different dose-response profiles.

Contrast test	Dimension	Contrast definition	Convex (0, 0, 0, 1)	Linear $(0, \frac{1}{3}, \frac{2}{3}, 1)$	Semi-concave (0, 0, 1, 1)	Concave (0, 1, 1, 1)
Helmert contrast	1	$(-\frac{1}{3} \quad -\frac{1}{3} \quad -\frac{1}{3} \quad 1)$	0.7880	0.4940	0.4940	0.2033
reverse Helmert contrast	1	$(-1 \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3})$	0.2504	0.6171	0.6171	0.8977
Linear contrast	1	$(-1 \quad -\frac{1}{3} \quad \frac{1}{3} \quad 1)$	0.6645	0.7437	0.8674	0.6645
Bivariate contrast (Helm. & rev. Helm.)	2	$(-\frac{1}{3} \quad -\frac{1}{3} \quad -\frac{1}{3} \quad 1)$ $(-1 \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3})$	0.7131	0.6358	0.6358	0.8379
Trivariate contrast (Helm., rev. Helm. and linear)	3	$(-\frac{1}{3} \quad -\frac{1}{3} \quad -\frac{1}{3} \quad 1)$ $(-1 \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3})$ $(-1 \quad -\frac{1}{3} \quad \frac{1}{3} \quad 1)$	0.7129	0.6893	0.7909	0.8300
Dunnett contrast	3	$(-1 \quad 0 \quad 0 \quad 1)$ $(-1 \quad 0 \quad 1 \quad 0)$ $(-1 \quad 1 \quad 0 \quad 0)$	0.5453	0.6205	0.7241	0.8103
Williams contrast	3	$(-1 \quad 0 \quad 0 \quad 1)$ $(-1 \quad 0 \quad \frac{1}{2} \quad \frac{1}{2})$ $(-1 \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3})$	0.6187	0.7154	0.7971	0.8648