Journal of Statistical Planning and Inference 27 (1991) 271-290
North-Holland

Sample size determination for step-down multiple test procedures: Orthogonal contrasts and comparisons with a control

Anthony J. Hayter

Department of Mathematics, University of Bath, Bath BA2 7AY, U.K.

Ajit C. Tamhane

Department of Statistics, Northwestern University, Evanston, IL 60208, U.S.A.

Received 1 June 1989; revised manuscript received 18 January 1990
Recommended by S. Panchapakesan

Abstract: We address the problem of sample size determination for step-down multiple comparison procedures (MCP's) for two nonhierarchical families – orthogonal contrasts and comparisons with a control, in order to guarantee a specified requirement on their power. The results for the corresponding single-step MCP's are obtained as special cases.

Numerical calculations of the sample sizes to guarantee a specified power requirement are carried out for the one-sided comparisons with a control problem in selected cases. These calculations show that for the cases considered, about 10% to 20% savings can be achieved in the total sample size by using the step-down MCP of Miller (1966, pp. 85-86) instead of the single-step MCP of Dunnett (1955). The percentage savings increase, as expected, with the number of treatments being compared with the control.

In the process of determining the smallest total sample size for each MCP to guarantee the specified power requirement, we also determine the optimum allocation of this sample size among the treatments and the control. We find that the square root allocation rule recommended by Dunnett (1955) provides a reasonable approximation to the optimum allocation for both the MCP's.

AMS Subject Classification: Primary 62J15; secondary 62K99.

Key words and phrases: Multiple comparisons; single-step procedures; step-down procedures; power; sample size determination; orthogonal contrasts; comparisons with a control; multivariate t-distribution; normal distribution.

1. Introduction

It is well-known that step-down multiple comparison procedures (MCP's) are more powerful than their single-step counterparts for simultaneous testing problems. (For the terminology used in this article we refer the reader to Hochberg and Tamhane (1987).) Although step-down MCP's have been used for this reason for a long time, the important design problem of sample size determination to guarantee a specified power requirement has not been yet addressed for these procedures. The sample size determination problem for single-step MCP's is considered in Chapter 6 of Hochberg and Tamhane (1987) and more recently in Hsu (1988).

In the present article we solve the sample size determination problem for step-down and single-step MCP's for the families of orthogonal contrasts and comparisons with a control under the normality setup. This enables us to make a quantitative assessment of the savings in the total sample size achieved by step-down MCP's relative to their single-step counterparts. We have carried out these computations for the family of one-sided comparisons with a control, where Dunnett's (1955) single-step MCP is compared with the step-down MCP originally proposed by Miller (1966, pp. 85–86). The same step-down MCP was independently proposed by Naik (1975) and was shown to satisfy the type I familywise error rate requirement by Marcus, Peritz and Gabriel (1976).

The following is the outline of the present article: Section 2 gives the preliminaries. These include the distributional setup assumed, the hypotheses and the associated probability requirements (type I familywise error rate and power), and the descriptions of the single-step and step-down MCP's under consideration. Two families of multiple comparisons of particular interest to us, namely, the families of orthogonal contrasts and comparisons with a control are also described. Section 3 states the principal theoretical results concerning the least favorable configurations of the MCP's; the proofs of these results are postponed to the Appendix. Section 4 applies these results to the problem of one-sided comparisons with a control, and gives tables of the 'optimal' sample sizes that must be taken on each one of the treatments and the control to guarantee the specified power requirement. These sample sizes are tabulated both for the single-step MCP of Dunnett (1955) and the step-down MCP of Miller (1966), and they are used to make assessments of the relative savings achieved by the latter MCP. Finally some concluding remarks are given in Section 5.

2. Preliminaries

2.1. Distributional setup

Consider the standard linear model setting and let $\hat{\boldsymbol{\theta}}=(\hat{\theta}_1,\dots,\hat{\theta}_k)$ be the best linear unbiased estimator of the unknown parameter vector of interest $\boldsymbol{\theta}=(\theta_1,\dots,\theta_k)$. We assume that the $\hat{\theta}_i$ are jointly normally distributed with means θ_i, a common variance $\sigma^2 \upsilon$ and a common correlation $\varrho=\text{corr}(\hat{\theta}_i,\hat{\theta}_j)$. $(1\le i\ne j\le k)$ where $\upsilon>0$ and $0\le\varrho<1$ are known constants and $\sigma^2>0$ is an unknown scalar. We further assume that an unbiased estimator S^2 of σ^2 is available based on ν degrees of freedom (d.f.) which is distributed independently of $\hat{\boldsymbol{\theta}}$ as a $\sigma^2\chi_v^2/v$ random variable (r.v.). Two examples of this distributional setup which are of particular interest to us are given in Section 2.4.

2.2. Hypotheses and probability requirements

We consider the following two families of hypotheses testing problems:

$$\text{(I)} \qquad \textit{One-sided}: \quad H_{0i}^{(1)}: \theta_i \le 0 \text{ vs. } H_{1i}^{(1)}: \theta_i > 0 \quad (1 \le i \le k). \tag{2.1}$$

$$\text{(II)} \qquad \textit{Two-sided}: \quad H_{0i}^{(2)}: \theta_i = 0 \text{ vs. } H_{1i}^{(2)}: \theta_i \ne 0 \quad (1 \le i \le k). \tag{2.2}$$

In the case of (2.2), it is often desired to make a *directional decision* (i.e., decide that $\theta_i > 0$ or < 0) for any $H_{0i}^{(2)}$ that is rejected $(1 \le i \le k)$.

For each testing problem we want the MCP that we use to satisfy the following *type I familywise error rate (FWE) requirement*:

$$P_{\boldsymbol{\theta}}\{\text{any true } H_{0i}^{(j)} \text{ is rejected } (1 \le i \le k)\} \le \alpha \quad \text{for all } \boldsymbol{\theta} \ (j=1,2), \tag{2.3}$$

where $0 < \alpha < 1$ is specified. In addition, for design purposes we postulate the following *power requirements* for (2.1) and (2.2); in these requirements $\delta > 0$ and $0 < \beta < 1\text{-}\alpha$ are specified constants.

(I) *One-sided tests*:

$$P_{\boldsymbol{\theta}}\{\text{all false } H_{0i}^{(1)} \text{ with } \theta_i \ge \delta\sigma \text{ are rejected}\} \ge 1-\beta \quad \text{for all } \boldsymbol{\theta}. \tag{2.4}$$

(IIa) *Two-sided tests (without directional decisions)*:

$$P_{\boldsymbol{\theta}}\{\text{all false } H_{0i}^{(2)} \text{ with } |\theta_i| \ge \delta\sigma \text{ are rejected}\} \ge 1-\beta \quad \text{for all } \boldsymbol{\theta}. \tag{2.5}$$

(IIb) *Two-sided tests (with directional decisions)*:

$$P_{\boldsymbol{\theta}}\{\text{all false } H_{0i}^{(2)} \text{ with } |\theta_i| \ge \delta\sigma \text{ are rejected with correct directional decisions}\} \ge 1-\beta \quad \text{for all } \boldsymbol{\theta}. \tag{2.6}$$

2.3. Procedures

The common step-down and single-step MCP's (which can be shown to have certain optimality properties) for the testing problems (2.1) and (2.2) are based on the test statistics

$$T_i = \frac{\hat{\theta}_i}{S\sqrt{v}} \quad (1 \le i \le k). \tag{2.7}$$

Under the configuration $\theta_1 = \cdots = \theta_m = 0$ $(1 \le m \le k)$, the r.v.'s $T_1, T_2, \ldots, T_m$ have a joint m-variate central t-distribution with v d.f. and common associated correlation ϱ (see Hochberg and Tamhane (1987, Appendix 3) for a definition of this distribution); for $m=1$, this of course reduces to the univariate Student's t-distribution. We denote the upper α point of $\max_{1 \le i \le m} T_i$ by $t_{m,v,\varrho}^{(\alpha)}$ (for $m=1$ simply by $t_v^{(\alpha)}$) and that of $\max_{1\le i\le m}|T_i|$ by $|t|^{(\alpha)}_{m,\nu,\varrho}$ (which equals $t^{(\alpha/2)}_{\nu}$ for $m=1$). The critical points $t^{(\alpha)}_{m,\nu,\varrho}$ and $|t|^{(\alpha)}_{m,\nu,\varrho}$ have been tabulated by Bechhofer and Dunnett (1988) for selected values of α, m, ν and ϱ.

The single-step MCP for the one-sided testing problem (2.1) rejects any $H^{(1)}_{0i}$ in favor of $H^{(1)}_{1i}$ if

$$T_i > t^{(\alpha)}_{k,\nu,\varrho} \quad (1\le i\le k). \tag{2.8}$$

We shall refer to this MCP as SS1. Similarly the single-step MCP for the two-sided testing problem (2.2) rejects any $H^{(2)}_{0i}$ in favor of $H^{(2)}_{1i}$ if

$$|T_i| > |t|^{(\alpha)}_{k,\nu,\varrho} \quad (1\le i\le k). \tag{2.9}$$

Rejection of any $H^{(2)}_{0i}$ can be accompanied by a directional decision that $\theta_i>0$ if $T_i>0$ and $\theta_i<0$ if $T_i<0$ $(1\le i\le k)$. We shall refer to this MCP as SS2.

It can be shown that both SS1 and SS2 satisfy the FWE requirement (2.3) (see Hochberg and Tamhane (1987, Ch. 3, Sec. 2)).

The step-down MCP for the one-sided testing problem (2.1) begins by ordering the statistics (2.7) as $T_{(1)}\le T_{(2)}\le\cdots\le T_{(k)}$. Let $H^{(1)}_{0(1)}, H^{(1)}_{0(2)},\ldots,H^{(1)}_{0(k)}$ be the corresponding hypotheses. At step 1, $H^{(1)}_{0(k)}$ is tested by comparing $T_{(k)}$ with $t^{(\alpha)}_{k,\nu,\varrho}$. If $T_{(k)}\le t^{(\alpha)}_{k,\nu,\varrho}$ then all $H^{(1)}_{0i}$ are retained (i.e., not rejected) without further tests. If $T_{(k)}> t^{(\alpha)}_{k,\nu,\varrho}$ then $H^{(1)}_{0(k)}$ is rejected and $H^{(1)}_{0(k-1)}$ is tested next by comparing $T_{(k-1)}$ with $t^{(\alpha)}_{k-1,\nu,\varrho}$. In general, $H^{(1)}_{0(m)}$ is tested and rejected iff $T_{(l)}>t^{(\alpha)}_{l,\nu,\varrho}$ for $l=k,k-1,\ldots,m$; if $H^{(1)}_{0(m)}$ is tested and not rejected because $T_{(m)}\le t^{(\alpha)}_{m,\nu,\varrho}$ then the hypotheses $H^{(1)}_{0(1)},\ldots,H^{(1)}_{0(m-1)}$ are retained without actually testing them. We shall refer to this MCP as SD1.

The step-down MCP for the two-sided testing problem (2.2) operates exactly in the same manner as above except now the ordered *absolute* values of the statistics (2.7), viz., $|T|_{(1)}\le|T|_{(2)}\le\cdots\le|T|_{(k)}$, are used and the critical constants $t^{(\alpha)}_{m,\nu,\varrho}$ are replaced by $|t|^{(\alpha)}_{m,\nu,\varrho}$ $(1\le m\le k)$. Rejection of any $H^{(2)}_{0i}$ can be accompanied by a directional decision in the usual manner. We shall refer to this MCP as SD2.

It can be shown that both SD1 and SD2 satisfy the FWE requirement (2.3) (see Hochberg and Tamhane (1987), Ch. 3, Sec. 4.2)). Also note that since $t^{(\alpha)}_{k,\nu,\varrho}>t^{(\alpha)}_{m,\nu,\varrho}$ and similarly $|t|^{(\alpha)}_{k,\nu,\varrho}>|t|^{(\alpha)}_{m,\nu,\varrho}$ for $m=1,\ldots,k-1$ and for any α, ν and ϱ, the step-down MCP's SD1 and SD2 are more powerful than the respective single-step MCP's SS1 and SS2.

Remark 1. If type III errors associated with directional decisions are included in the definition of the FWE then it is not known whether or not SD2 satisfies (2.3) in all cases (see, e.g., Shaffer (1980)) although SS2 clearly does so (see Hochberg and Tamhane (1987, Ch. 2, Sec. 2.3.2)). However, type III errors are considered as part of the power requirement (2.6) and not as part of the FWE requirement in the present formulation. This seems to us a more natural way of treating type III errors.

2.4. Examples

Example 1 (*Orthogonal contrasts*). Consider a completely randomized experiment with $p \geq 2$ factorial treatment combinations having cell means μ_j $(1 \leq j \leq p)$, a common error variance σ^2 and $n \geq 2$ replications per cell. Let

$$\theta_i = \sum_{j=1}^{p} c_{ij} \mu_j \quad (1 \leq i \leq k)$$

where the coefficients c_{ij} satisfy

$$\sum_{j=1}^{p} c_{ij} = 0, \quad \sum_{j=1}^{p} c_{ij}^2 = 1 \quad \text{and} \quad \sum_{j=1}^{p} c_{ij} c_{i'j} = 0 \quad (1 \leq i \neq i' \leq k).$$

Thus the θ_i are normalized orthogonal contrasts. Let

$$\hat{\theta}_i = \sum_{j=1}^{p} c_{ij} \bar{X}_j \quad (1 \leq i \leq k)$$

where the $\bar{X}_j$ are the cell sample means, and let S^2 be the usual analysis of variance (ANOVA) mean square error. Then the $\hat{\theta}_i$ are independent normal (i.e., the common correlation $\varrho = 0$) with means θ_i and a common variance σ^2/n, and S^2 is distributed as $\sigma^2 \chi_v^2 / v$ independent of the $\hat{\theta}_i$'s with $v = p(n-1)$ d.f.

The critical points $t_{m,v,0}^{(\alpha)}$ and $|t|_{m,v,0}^{(\alpha)}$ needed to apply the MCP's of the previous section in this case are referred to as Studentized maximum and Studentized maximum modulus critical points, respectively. These critical points are also tabulated in Bechhofer and Dunnett (1988).

Example 2 (*Comparisons with a control*). Suppose that there are $k \geq 2$ treatment groups labelled $1, 2, \ldots, k$ which are to be compared with a control group labelled 0. The observations from the i-th group are assumed to be independent and identically distributed (i.i.d.) $N(\mu_i, \sigma^2)$ r.v.'s where the μ_i are unknown group means and σ^2 is a common unknown error variance. A random sample of size n is drawn from each one of the treatment groups and another random sample of size n_0 from the control group.

The parameters of interest are $\theta_i = \mu_i - \mu_0$ whose best linear unbiased estimators are $\hat{\theta}_i = \bar{X}_i - \bar{X}_0$ $(1 \leq i \leq k)$ where $\bar{X}_i$ is the sample mean for the i-th group $(0 \leq i \leq k)$. Here the $\hat{\theta}_i$ are correlated normals with means $\theta_i = \mu_i - \mu_0$, a common variance $\sigma^2(1/n + 1/n_0)$ and a common correlation $\varrho = n/(n + n_0)$. The ANOVA mean square error S^2 is distributed as $\sigma^2 \chi_v^2 / v$ with $v = N - (k+1)$ d.f. where $N = n_0 + kn$ is the total sample size.

3. Least favorable configuration results

Our next task is to determine for each MCP described in Section 2.3 the LFC of the θ_i's at which the power (as defined by the l.h.s. of (2.4) for the one-sided testing problem and the l.h.s. of (2.5) or (2.6) for the two-sided testing problem) of that MCP is minimized. For the given design one can then find the smallest total sample size N, which makes this minimum greater than or equal to the specified lower bound $1-\beta$, thus guaranteeing the appropriate power requirement.

Note that for given k and specified α and δ, the minimum power at the LFC is a function of υ, ϱ and ν, which in turn are functions of N depending on the design employed. By varying some of the design parameters it is sometimes possible to maximize this minimum for given N. This evaluation of the *max-min power* enables us to obtain the smallest N to guarantee a specified power requirement. For example, in the comparisons with a control problem described in Example 2, if we let $r=n_0/n$ and $N=n_0+kn$ then $\upsilon=\{(r+k)(r+1)/rN\}^{1/2}$, $\varrho=1/(1+r)$ and $\nu=N-(k+1)$. Thus ν and ϱ depend on the design parameter r, which can be chosen to maximize the minimum power for given N. We shall do this in the computation of the tables of N for this problem in Section 4.

We now state the result for the one-sided testing problem (2.1). Proofs of all the results are given in the Appendix.

Theorem 1. *For the step-down* MCP SD1, *the minimum over* $\boldsymbol{\theta}$ *of*

$$P_{\boldsymbol{\theta}}\{\text{reject } \mathrm{H}_{01}^{(1)},\ldots,\mathrm{H}_{0m}^{(1)} \mid \theta_i \geq \delta\sigma\ (1\leq i\leq m),\ \theta_i<\delta\sigma\ (m+1\leq i\leq k)\} \quad (3.1)$$

for fixed m $(1\leq m\leq k)$ *is attained when* $\theta_i=\delta\sigma$ *for* $1\leq i\leq m$ *and* $\theta_i=-\infty$ *for* $m+1\leq i\leq k$. *Denoting this minimum by* $P_m(\text{SD1})$, *the overall minimum of the l.h.s. of* (2.4) *for* SD1 *is given by* $\min_{1\leq m\leq k} P_m(\text{SD1})$.

Corollary 1. *Theorem* 1 *also holds for the single-step* MCP SS1. *Moreover, if* $P_m(\text{SS1})$ *denotes the associated minimum value of* (3.1) $(1\leq m\leq k)$ *then* $\min_{1\leq m\leq k} P_m(\text{SS1})=P_k(\text{SS1})$, *i.e., for* SS1 *the overall minimum of the l.h.s. of* (2.4) *is attained when* $\theta_i=\delta\sigma$ *for* $i=1,\ldots,k$.

We next state the result for the two-sided testing problem (2.2).

Theorem 2. *If* $\varrho=0$ *then for the step-down* MCP SD2, *the minimum over* $\boldsymbol{\theta}$ *of*

$$P_{\boldsymbol{\theta}}\{\text{reject } \mathrm{H}_{01}^{(2)},\ldots,\mathrm{H}_{0m}^{(2)} \mid |\theta_i| \geq \delta\sigma\ (1\leq i\leq m),\ |\theta_i|<\delta\sigma\ (m+1\leq i\leq k)\} \quad (3.2)$$

for fixed m $(1\leq m\leq k)$ *is attained when* $|\theta_i|=\delta\sigma$ *for* $1\leq i\leq m$ *and* $|\theta_i|=0$ *for* $m+1\leq i\leq k$. *Denoting this minimum by* $P_m(\text{SD2})$, *the overall minimum of the l.h.s. of* (2.5) *is given by* $\min_{1\leq m\leq k} P_m(\text{SD2})$.

Corollary 2. *Theorem* 2 *also holds for the single-step* MCP SS2. *Moreover, if* $P_m(\text{SS2})$ *denotes the associated minimum value of* (3.2) *then* $\min_{1\leq m\leq k} P_m(\text{SS2})=$

P_k(SS2), *i.e., for* SS2 *the overall minimum of the l.h.s. of* (2.5) *is attained when* $|\theta_i| = \delta\sigma$ *for* $i = 1, \ldots, k$.

Corollary 3. *Theorem* 2 *and Corollary* 2 *also hold if* (3.2) *is replaced with*

$$P_\theta\{\text{reject } \mathrm{H}_{01}^{(2)}, \ldots, \mathrm{H}_{0m}^{(2)} \text{ with correct directional decisions} \mid |\theta_i| \geq \delta\sigma\ (1 \leq i \leq m),\ |\theta_i| < \delta\sigma\ (m+1 \leq i \leq k)\}. \tag{3.3}$$

Denoting by P_m(SD2) *and* P_m(SS2) *the associated minimum values of* (3.3) *we obtain that for* SD2 *the overall minimum of the l.h.s. of* (2.6) *is given by* $\min_{1 \leq m \leq k} P_m$(SD2) *while that for* SS2 *is given by* $\min_{1 \leq m \leq k} P_m(\text{SS2}) = P_k(\text{SS2})$, *i.e., for* SS2 *the overall minimum of the l.h.s. of* (2.6) *is attained when* $|\theta_i| = \delta\sigma$ *for* $i = 1, \ldots, k$.

Remark 2. Theorem 2 does not hold when $\varrho > 0$. In this case the LFC is not simple and must be numerically determined. It is worth noting, however, that there are examples where two-sided multiple tests on independent estimates are of interest. One such example is provided by multiple regression model building where the model is parametrized so that the effects are orthogonal, e.g., in orthogonal polynomial regression. Here often no directional decisions are required.

4. Tables of sample sizes for one-sided comparisons with a control

In this section we determine the smallest $N = n_0 + kn$ and the associated optimal allocation of the sample sizes, (n_0, n), required by SD1 and SS1 for guaranteeing the power requirement (2.4) for selected values of k, α, δ and $1-\beta$. Toward this end we first derive expressions for P_m(SD1) $(1 \leq m \leq k)$ and P_k(SS1).

4.1. Expressions for minimum powers P_m

For convenience of notation henceforth we shall use c_i to denote $t_{i,\nu,\varrho}^{(\alpha)}$ $(1 \leq i \leq k)$; here $\nu = N - (k+1)$ and $\varrho = n/(n+n_0)$. To evaluate P_m(SD1) we can take $\mu_0 = 0$, $\mu_i = \delta\sigma$ $(1 \leq i \leq m)$ and $\mu_i = -\infty$ $(m+1 \leq i \leq k)$. In that case we can represent the statistics T_i as

$$T_i = \frac{Z_i\sqrt{1-\varrho} + \delta\sqrt{n(1-\varrho)} - Z_0\sqrt{\varrho}}{U} \quad (1 \leq i \leq m) \tag{4.1}$$

and $T_{m+1} = \cdots = T_k = -\infty$ with probability 1. Here the Z_i are i.i.d. N(0,1) r.v.'s independent of U, which is distributed as a $(\chi_\nu^2/\nu)^{1/2}$ r.v. Using the notation (introduced in the Appendix) $(x_1, \ldots, x_p) >^* (y_1, \ldots, y_p)$ to mean $x_{(i)} > y_{(i)}$ for $i = 1, \ldots, p$, where the $x_{(i)}$ and $y_{(i)}$ are the ordered x_i's and y_i's, respectively, we can write

$$P_m(\text{SD1}) = P\{(T_1, \ldots, T_m) >^* (c_{k-m+1}, \ldots, c_k)\}$$

$$= P\{(Z_1, \ldots, Z_m) \overset{*}{>} (D_{k-m+1}, \ldots, D_k)\} \tag{4.2}$$

where

$$D_i = c_i U + Z_0\sqrt{\varrho} - \delta\sqrt{n(1-\varrho)} \quad (k-m+1 \le i \le k).$$

By conditioning on $U=u$ and $Z_0=z_0$ and thus on $D_i = d_i = c_i u + z_0\sqrt{\varrho} - \delta\sqrt{n(1-\varrho)}$, we can write (4.2) as

$$\int_0^\infty \int_{-\infty}^\infty P\{(Z_1, \ldots, Z_m) \overset{*}{>} (d_{k-m+1}, \ldots, d_k)\}\,\phi(z_0) f_\nu(u)\,\mathrm{d}z_0\,\mathrm{d}u \tag{4.3}$$

where $\phi(\cdot)$ is the standard normal density function and $f_\nu(\cdot)$ is the density function of a $(\chi_\nu^2/\nu)^{1/2}$ r.v.

We still need a computable expression for the probability term in the integrand of (4.3). This is obtained in a recursive form in the following lemma.

Lemma. *For $m=1$, $P\{Z_1 > d_k\} = 1 - \Phi(d_k)$ and for $m>1$,*

$$\begin{aligned} &P\{(Z_1, \ldots, Z_m) \overset{*}{>} (d_{k-m+1}, \ldots, d_k)\} \\ &\quad = \{\Phi(d_{k-m+2}) - \Phi(d_{k-m+1})\} P\{(Z_1, \ldots, Z_{m-1}) \overset{*}{>} (d_{k-m+2}, \ldots, d_k)\} \\ &\qquad + \{\Phi(d_{k-m+3}) - \Phi(d_{k-m+2})\} P\{(Z_1, \ldots, Z_{m-1}) \\ &\qquad\qquad \overset{*}{>} (d_{k-m+1}, d_{k-m+3}, \ldots, d_k)\} \\ &\qquad + \cdots + \{1 - \Phi(d_k)\} P\{(Z_1, \ldots, Z_{m-1}) \overset{*}{>} (d_{k-m+1}, \ldots, d_{k-1})\} \end{aligned} \tag{4.4}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Proof. Condition on Z_m to lie in successive intervals: $(d_{k-m+1}, d_{k-m+2}], \ldots, (d_k, \infty)$ which leads to (4.4). Note that we do not condition on Z_m to lie in the interval $(-\infty, d_{k-m+1}]$ because in that case $\mathrm{H}_{0m}^{(1)}$ cannot be rejected. □

By combining (4.3) with (4.4) a feasible algorithm can be developed to compute $P_m(\mathrm{SD1})$ $(1 \le m \le k)$.

Next consider

$$\begin{aligned} P_k(\mathrm{SS1}) &= P\{T_1 > c_k, \ldots, T_k > c_k\} \\ &= P\{Z_1 > D_k, \ldots, Z_k > D_k\} \\ &= \int_0^\infty \int_{-\infty}^\infty [1 - \Phi(d_k)]^k \phi(z_0) f_\nu(u)\,\mathrm{d}z_0\,\mathrm{d}u \end{aligned} \tag{4.5}$$

where $d_k = c_k u + z_0\sqrt{\varrho} - \delta\sqrt{n(1-\varrho)}$.

Remark 3. For the problem of orthogonal contrasts, we put $\varrho = 0$ in (4.3) and (4.5) and see that they reduce to one-dimensional integrals since the $d_i = c_i u - \delta\sqrt{n}$ $(1 \le i \le k)$ become free of z_0. The expressions for $P_m(\mathrm{SD2})$ and $P_k(\mathrm{SS2})$ when $\varrho = 0$

can also be readily obtained both when type III errors are ignored (see (3.2)) and when they are not ignored (see (3.3)).

4.2. Tables

Tables 1, 2 and 3 give the smallest N and the associated optimal allocation (n_0, n) required to satisfy the power requirement (2.4) using SD1 and SS1 for $k=2$, 3 and 4, respectively. The calculations are done for $\alpha=0.05$, $1-\beta=0.70$, 0.80, 0.90, 0.95, 0.99 and $\delta=0.5$, 1.0. These tables also give the critical constants $c_i = t^{(\alpha)}_{i,\nu,\varrho}$ required to implement SD1 and the critical constant $c_k = t^{(\alpha)}_{k,\nu,\varrho}$ required to implement SS1. The last column in each table gives the percentage relative saving (RS) in the total sample size N achieved by SD1 over SS1 which is given by

$$\mathrm{RS} = \frac{N(\mathrm{SS1}) - N(\mathrm{SD1})}{N(\mathrm{SS1})} \times 100.$$

For given k, α, δ and (n_0, n), the power of SD1 under the LFC was computed by evaluating $P_m(\mathrm{SD1})$ $(1 \le m \le k)$ using (4.3) and (4.4) and then taking their minimum, that for SS1 was computed using (4.5). For each given total sample size N, the maximum of this minimum power over all allocations (n_0, n) subject to $N = n_0 + kn$ was found for SD1 and SS1. The smallest N for which this max–min power $\ge 1-\beta$ and the associated optimal allocation (n_0, n) were then determined for these MCP's. The critical constants $c_i = t^{(\alpha)}_{i,\nu,\varrho}$ $(1 \le i \le k)$ required in this computation were obtained as follows: For $i=1$, the corresponding critical point is the univariate Student's t critical point, which was obtained using a NAG subroutine. For $i>1$, the desired critical points were obtained by interpolating in the tables of Bechhofer and Dunnett (1988). As suggested by these authors, interpolation with

Table 1
Optimal sample sizes and critical constants required by SD1 and SS1 to guarantee the power requirement (2.4) ($k=2$, $\alpha=0.05$)

δ	$1-\beta$	Step-down procedure SD1					Single-step procedure SS1				RS (%)
		N	n	n_0	c_2	c_1	N	n	n_0	c_2	
1.0	0.70	40	12	16	1.985	1.687	49	15	19	1.972	18.4
	0.80	49	15	19	1.972	1.679	59	18	23	1.963	16.9
	0.90	64	19	26	1.962	1.670	74	22	30	1.957	13.5
	0.95	77	23	31	1.955	1.666	89	27	35	1.950	13.5
	0.99	108	32	44	1.947	1.659	120	36	48	1.944	10.0
0.5	0.70	155	47	61	1.939	1.655	187	58	71	1.935	17.1
	0.80	191	58	75	1.936	1.653	228	70	88	1.935	16.2
	0.90	248	74	100	1.937	1.651	291	88	115	1.936	14.8
	0.95	302	89	124	1.938	1.650	350	105	140	1.937	13.7
	0.99	426	126	174	1.938	1.648	474	141	192	1.937	10.1

Table 2
Optimal sample sizes and critical constants required by SD1 and SS1 to guarantee the power requirement (2.4) ($k=3$, $\alpha=0.05$)

δ	$1-\beta$	Step-down procedure SD1						Single-step procedure SS1				RS (%)
		N	n	n_0	c_3	c_2	c_1	N	n	n_0	c_3	
1.0	0.70	63	14	21	2.126	1.973	1.671	78	18	24	2.112	19.2
	0.80	76	17	25	2.117	1.959	1.666	92	20	32	2.123	17.4
	0.90	96	21	33	2.120	1.954	1.662	113	25	38	2.106	15.0
	0.95	114	25	39	2.107	1.950	1.659	132	29	45	2.104	13.6
	0.99	153	33	54	2.103	1.946	1.655	173	37	62	2.102	11.6
0.5	0.70	244	54	82	2.095	1.940	1.651	306	69	99	2.093	20.3
	0.80	296	65	101	2.096	1.941	1.650	362	81	119	2.094	18.2
	0.90	376	81	133	2.099	1.942	1.649	447	98	153	2.096	15.9
	0.95	451	97	160	2.099	1.942	1.648	525	114	183	2.098	14.1
	0.99	610	130	220	2.100	1.943	1.647	691	149	244	2.098	11.7

Table 3
Optimal sample sizes and critical constants required by SD1 and SS1 to guarantee the power requirement (2.4) ($k=4$, $\alpha=0.05$)

δ	$1-\beta$	Step-down procedure SD1							Single-step procedure SS1				RS (%)
		N	n	n_0	c_4	c_3	c_2	c_1	N	n	n_0	c_4	
1.0	0.70	86	15	26	2.229	2.120	1.960	1.664	109	19	33	2.221	21.1
	0.80	102	18	30	2.221	2.113	1.954	1.661	126	22	38	2.216	19.0
	0.90	126	22	38	2.216	2.109	1.951	1.658	153	26	49	2.216	17.6
	0.95	149	25	49	2.218	2.110	1.950	1.656	176	30	56	2.213	15.3
	0.99	198	33	66	2.214	2.106	1.947	1.653	227	39	71	2.209	12.8
0.5	0.70	340	60	100	2.205	2.099	1.942	1.649	431	77	123	2.203	21.1
	0.80	403	70	123	2.208	2.101	1.944	1.649	500	88	148	2.206	19.4
	0.90	501	86	157	2.209	2.102	1.945	1.648	606	105	186	2.208	17.3
	0.95	591	100	191	2.211	2.104	1.946	1.647	703	121	219	2.209	15.9
	0.99	790	133	258	2.212	2.105	1.946	1.647	906	154	290	2.211	12.8

respect to ν was linear in $1/\nu$, while that with respect to ϱ was linear in $1/(1-\varrho)$. Both these interpolations result in approximations to the exact critical points that are accurate for practical purposes. All computations were done in Fortran on a Sun workstation.

Remark 4. It is of interest to know how the computing effort grows with k. To find the minimum power of SD1, one must evaluate k $P_m(\mathrm{SD1})$ terms (for $m=1,\ldots,k$), each of which is a bivariate integral (4.3) (for $\varrho>0$) regardless of the value of k

(assuming $\Phi(\cdot)$ is computed without resorting integration, many such algorithms being available). For each m, the integrand of the bivariate integral is computed recursively using (4.4) which involves N_m calculations of terms of the form $\{\Phi(d_i)-\Phi(d_j)\}$ where $N_m=m(1+N_{m-1})$ and $N_1=1$; thus $N_2=4$, $N_3=15$, $N_4=64$, etc. Although N_m increases rapidly with m, these calculations are straightforward and not time-consuming, once all the $\Phi(d_i)$ are computed and stored. (It is worth noting that the programming of the expression is simplified by the fact that the N_m terms constituting expression (4.4) need not be explicitly written out; rather they are recursively generated by the algorithm.) Therefore the total computing effort for evaluating the minimum power grows only somewhat faster than linearly with k.

Computations for determining the smallest total sample size to guarantee given power are, however, further complicated by the necessity to vary N and the ratio n_0/n because this changes the k critical constants c_i, which must be obtained by interpolation in the Bechhofer–Dunnett tables as explained above. If the tabled values are stored in the memory then interpolation poses no problem. In conclusion, computations are easily feasible for k not too large, say, about 10.

Remark 5. We should like to point out that which one of the k P_m(SD1) terms yields the minimum cannot be determined analytically - only numerically, although in our computations we did observe that, as a function of m, P_m(SD1) is unimodal, first decreasing and then increasing with m. Moreover, the index m^* corresponding to the smallest P_m(SD1) is a function of N among other things. As an example, for $k=4$, $\delta=1.0$, $\alpha=0.05$ and $n_0=2n$ we found that $m^*=3$ for $30\le N\le 120$, while $m^*=2$ for $180\le N\le 200$. (We did not attempt to determine the exact N-value at which m^* changes from 3 to 2 which will be some number between 120 and 180.) This pattern of m^* jumping to the next lower integer as N is increased was observed in all of our calculations.

4.3. Discussion

First note that for the cases considered, the relative saving due to SD1 ranges between 10% to 20% with high values achieved at low powers $1-\beta$ and vice-versa. The relative saving is roughly constant with respect to the choice of δ but, as expected, increases with k. Thus very substantial savings in N are possible even for modest values of k if we use a step-down MCP instead of a single-step MCP.

Next consider the behavior of the optimal allocation (n_0, n) both for SD1 and SS1. Table 4 gives the values of the ratio $r=n_0/n$ for these MCP's computed from the entries in Tables 1-3. We see that both for $\delta=1.0$ and 0.5, the r-values are quite close to and in all cases except one (when it is equal) less than $\sqrt{k}$. Thus the square root allocation rule of Dunnett (1955) appears to work reasonably well both for SD1 and SS1. For $\delta=0.5$, the r-values seem to approach $\sqrt{k}$ from below as $1-\beta$ increases; the lack of monotone behavior of r for $\delta=1.0$ may be explained by the discreteness effects associated with small sample sizes. Note also that the r-values

Table 4
Values of $r = n_0/n$

δ	$1-\beta$	Step-down procedure SD1			Single-step procedure SS1		
		$k=2$	$k=3$	$k=4$	$k=2$	$k=3$	$k=4$
1.0	0.70	1.333	1.500	1.733	1.267	1.333	1.737
	0.80	1.267	1.471	1.667	1.278	1.600	1.727
	0.90	1.368	1.571	1.727	1.364	1.520	1.885
	0.95	1.348	1.560	1.960	1.296	1.552	1.867
	0.99	1.375	1.636	2.000	1.333	1.676	1.821
0.5	0.70	1.298	1.519	1.667	1.224	1.435	1.597
	0.80	1.293	1.554	1.757	1.257	1.469	1.682
	0.90	1.351	1.642	1.826	1.307	1.561	1.771
	0.95	1.393	1.649	1.910	1.333	1.605	1.810
	0.99	1.381	1.692	1.940	1.362	1.638	1.883

for SD1 are almost always higher than the corresponding r-values for SS1, and are thus closer to $\sqrt{k}$.

One other matter deserving comment is that in a number of cases, $c_i = t^{(\alpha)}_{i,\nu,\varrho}$ for $i > 1$ increases as $1-\beta$ (and hence N and $\nu = N-(k+1)$) increases. This apparently anomalous behavior is explained by the fact that $\varrho = 1/(1+r)$ decreases (since r increases) as $1-\beta$ increases, and it is well-known that $t^{(\alpha)}_{i,\nu,\varrho}$ increases as ϱ decreases (Slepian (1962)). Thus the effect of decreasing ϱ dominates that of increasing ν especially when ν is large.

Finally we remark that the tables of sample sizes given here are not intended to be exhaustive. For practical application it would perhaps be more convenient to plot graphs of the max–min power versus δ for different values of N for given k and α. These graphs can be supplemented by some simple rules for choosing nearly optimum $r = n_0/n$.

5. Concluding remarks

This article gives for the first time a method for determining the sample sizes to guarantee a specified power requirement for step-down MCP's and a quantitative assessment of their power advantage over the corresponding single-step MCP's. The numerical results demonstrate that quite substantial savings in the sample sizes can be achieved by using step-down MCP's and hence they are recommended when the problem is one of simultaneous testing.

Recently Bofinger (1987) and Stefánsson, Kim and Hsu (1988) have shown how SD1 can be inverted to obtain $100(1-\alpha)\%$ lower one-sided simultaneous confidence limits on the θ_i. These limits have the form $\theta_i > 0$ for the rejected $H^{(1)}_{0i}$ and

$\theta_i > \hat{\theta}_i - t_{p,\nu,\varrho}^{(\alpha)} s\sqrt{v}$ for the accepted $H_{0i}^{(1)}$ where $p \le k$ is the number of accepted $H_{0i}^{(1)}$. Comparing these confidence limits with those obtained by inverting SS1, viz., $\theta_i > \hat{\theta}_i - t_{k,\nu,\varrho}^{(\alpha)} s\sqrt{v}$ for $1 \le i \le k$, we see that the SD1 limits are at least as sharp (strictly sharper when $p < k$) for the accepted $H_{0i}^{(1)}$. However, in practice one would typically want sharp and informative confidence limits on the θ_i for the *rejected* $H_{0i}^{(1)}$. Unfortunately, the SD1 limits fail this test; on the other hand, the SS1 limits on the θ_i are strictly positive in this case, and hence are sharper and informative. Analogous problems arise when SD2 is inverted to yield two-sided confidence limits; see Bofinger (1987). Therefore use of step-down MCP's for simultaneous confidence estimation is not recommended at this time.

In this article we have restricted to the so-called balanced experiments satisfying the conditions that the $\text{var}(\hat{\theta}_i)$ are equal and the $\text{corr}(\hat{\theta}_i, \hat{\theta}_j)$ are also equal. There can arise situations in practice where the experiment is unbalanced and one wishes to assess the guaranteed minimum power. The unbalance may result either because the experiment is not a designed one, or even if it is designed to be balanced, unforeseen losses of data can occur in practice making it unbalanced. The LFC results can be extended to the case of unequal $\text{var}(\hat{\theta}_i)$ in a straightforward manner but unequal $\text{corr}(\hat{\theta}_i, \hat{\theta}_j)$ introduce difficulties, which cannot be handled with the current methods of proof. Furthermore, evaluation of the exact critical points poses computational difficulties if the correlations are arbitrary. For a step-down MCP such critical points need to be calculated repeatedly which makes its application difficult. A solution to this problem is provided by the sequentially rejective Bonferroni procedure of Holm (1979). This is a step-down MCP, which uses conservative critical constants $c_i = t_\nu^{(\alpha/i)}$ and $c_i = t_\nu^{(\alpha/2i)}$ $(1 \le i \le k)$ for the one-sided and two-sided testing problems, respectively. Use of these conservative critical points diminishes its power advantage over the competing single-step MCP. However, if the $\text{corr}(\hat{\theta}_i, \hat{\theta}_j)$ are all of the form $\gamma_i \gamma_j$ for some $\gamma_i \ge 0$ (which holds true in the comparisons with the control problem with unequal number of observations on the treatments), then a multivariate t integral can be reduced to a bivariate integral simplifying the computational problem tremendously; see Hochberg and Tamhane (1987, Ch. 5, Sec. 2.1.1). Dunnett and Tamhane (1991) have provided an exact method for carrying out SD1 and SD2 in this case.

Although numerical results have been obtained only for the problem of one-sided comparisons with a control, we expect qualitatively similar conclusions for one-sided and two-sided tests on orthogonal contrasts. Computations for the orthogonal case are simpler because (1) the integrals (4.3) and (4.5) become univariate when $\varrho = 0$, and (2) the critical constants $t_{i,\nu,0}^{(\alpha)}$ and $|t|_{i,\nu,0}^{(\alpha)}$ are found directly from existing tables, there being no need for interpolation with respect to ϱ. Also there is no search required to find the optimal allocation of the samples among the treatments. However, computations for the two-sided case are complicated by the fact that $P_m(\text{SD2})$ is evaluated under the least favorable configuration $|\theta_i| = \delta\sigma$ for $1 \le i \le m$ and $|\theta_i| = 0$ for $m+1 \le i \le k$; see Theorem 2. Therefore the statistics $T_{m+1}, \ldots, T_k$ do not drop out from the probability calculation as they do in the one-sided case.

Therefore the integral (4.3) is replaced by an expression, which is much harder to write down and compute. This problem is under study at present.

In the power requirements (2.4)–(2.6) the quantity $\delta\sigma$ is in the units of unknown σ. If this quantity were to be specified in absolute units, as some practitioners might prefer, then the given power requirements cannot be guaranteed using single-stage MCP's. Two-stage or more generally sequential sampling schemes must then be employed; see Hochberg and Tamhane (1987, Ch. 6, Sec. 2).

Finally we note that step-down MCP's such as those of Duncan (1955) and of Newman (1939) and Keuls (1952) are widely used in practice for testing pairwise homogeneity hypotheses and more generally for testing subset homogeneity hypotheses. (As is well-known, these step-down MCP's in their classical form do not satisfy the type I FWE requirement and need to be modified so that they do; see Hochberg and Tamhane (1987, Ch. 2, Sec. 4.3.3.3).) The problems of power evaluation and sample size determination are also important for these MCP's. Here the competing single-step MCP's are those of Tukey (1953) and Scheffé (1953). The proofs of the LFC's in the present article do not readily extend to these MCP's because the subset homogeneity hypotheses bear implication relations with each other. Such a family of hypotheses is referred to as a *hierarchical* family, while the two families considered in the present article are *nonhierarchical*. Sample size determination for step-down MCP's in the case of hierarchical families remains a problem for future research.

Appendix

Before stating the proofs of the theorems, we shall make some simplifications. First we may assume, without loss of generality, that σ^2 is known. This will eliminate the need for first conditioning on S^2 and then unconditioning on it. (In the final expressions for power used for computational purposes this operation involving one additional integration must, of course, be done as in (4.3) and (4.5).) Thus the T_i defined in (2.7) can be regarded as equicorrelated $N(\lambda_i, 1)$ r.v.'s with common correlation ϱ where $\lambda_i = \theta_i/\sigma\sqrt{v}$ $(1 \le i \le k)$. We denote the critical constants used in SD1 and SD2 generically by $c_k \ge c_{k-1} \ge \cdots \ge c_1$. The critical constants used in SS1 and SS2 are then generically denoted by c_k. We also use the following notation:

$$(x_1, \ldots, x_p) \stackrel{*}{>} (y_1, \ldots, y_p)$$

means that $x_{(i)} > y_{(i)}$ for $i = 1, \ldots, p$ where $x_{(p)} \ge \cdots \ge x_{(1)}$ and $y_{(p)} \ge \cdots \ge y_{(1)}$ are the ordered x_i's and y_i's, respectively.

Proof of Theorem 1. We first prove the result for $\varrho = 0$ in which case the T_i are independent $N(\lambda_i, 1)$ r.v.'s $(1 \le i \le k)$. Consider a parameter configuration of the θ_i's or equivalently that of the λ_i's, which for fixed m $(1 \le m \le k)$ satisfies $\lambda_i \ge \Delta = \delta/\sqrt{v}$ $(1 \le i \le m)$, $\lambda_i \le 0$ $(m+1 \le i \le k)$. We can write

$$P_{\boldsymbol{\lambda}}\{\text{reject } \mathrm{H}_{01}^{(1)}, \dots, \mathrm{H}_{0m}^{(1)}\}$$
$$= \int \cdots \int A(\lambda_1, \dots, \lambda_m; t_{m+1}, \dots, t_k) \prod_{i=m+1}^{k} \phi_i(t_i)\, \mathrm{d}t_i. \tag{A.1}$$

Here $\phi_i(t_i)$ is the density function of a $\mathrm{N}(\lambda_i, 1)$ r.v. $(1 \le i \le k)$ and

$$A(\lambda_1, \dots, \lambda_m; t_{m+1}, \dots, t_k) = P_{\boldsymbol{\lambda}}\{(T_1, \dots, T_m) \succ^{*} (d_1, \dots, d_m)\} \tag{A.2}$$

where $(d_1, \dots, d_m)$ is some subset of the critical constants $(c_1, \dots, c_k)$ depending on the conditioned values $t_{m+1}, \dots, t_k$. For instance, suppose $k=2$ and $m=1$. If $t_2 < c_2$ then in order to reject $\mathrm{H}_{01}^{(1)}$ using SD1 we must have $T_1 > c_2$, i.e., $d_1 = c_2$. On the other hand, if $t_2 > c_2$ then we only need $T_1 > c_1$, i.e., $d_1 = c_1$.

Now $P_{\boldsymbol{\lambda}}\{(T_1, \dots, T_m) \succ^{*} (d_1, \dots, d_m)\}$ is decreasing in each d_i $(1 \le i \le m)$. Hence (A.1) is minimized when $(d_1, \dots, d_m) = (c_{k-m+1}, \dots, c_k)$ which leads to the inequality

$$P_{\boldsymbol{\lambda}}\{\text{reject } \mathrm{H}_{01}^{(1)}, \dots, \mathrm{H}_{0m}^{(1)}\} \ge P_{\boldsymbol{\lambda}}\{(T_1, \dots, T_m) \succ^{*} (c_{k-m+1}, \dots, c_k)\} \tag{A.3}$$

with equality holding iff $\lambda_{m+1} = \cdots = \lambda_k = -\infty$.

Next we wish to find the minimum of the r.h.s. of (A.3) subject to $\lambda_i \ge \Delta$ $(1 \le i \le m)$. If $m=1$ then the r.h.s. of (A.3) equals $P_{\lambda_1}\{T_1 > c_k\}$ which is minimized subject to $\lambda_1 \ge \Delta$ when $\lambda_1 = \Delta$ because of the stochastically increasing property of T_1 in λ_1. (This is the only property of the normal distribution needed in this proof.) If $m > 1$ then fix j $(1 \le j \le m)$, say $j = 1$. The r.h.s. of (A.3) can be written as

$$\int \cdots \int B(\lambda_1; t_2, \dots, t_m) \prod_{i=2}^{m} \phi_i(t_i)\, \mathrm{d}t_i$$

where $B(\lambda_1; t_2, \dots, t_m) = 0$ or $P_{\lambda_1}\{T_1 > d_1\}$ for some $d_1 \in \{c_{k-m+1}, \dots, c_k\}$, the actual value of d_1 being a function of the conditioned values $t_2, \dots, t_m$. For instance, for $m=3$ we have $B=0$ if $\max(t_2, t_3) \le c_{k-1}$ or if $\min(t_2, t_3) \le c_{k-2}$. Now suppose that $\max(t_2, t_3) > c_{k-1}$ and $\min(t_2, t_3) > c_{k-2}$. In that case if $\max(t_2, t_3) \le c_k$ then $B = P_{\lambda_1}\{T_1 > d_1\}$ with $d_1 = c_k$; on the other hand, if $\max(t_2, t_3) > c_k$ and $\min(t_2, t_3) \le c_{k-1}$ then $B = P_{\lambda_1}\{T_1 > d_1\}$ with $d_1 = c_{k-1}$; finally if $\max(t_2, t_3) > c_k$ and $\min(t_2, t_3) > c_{k-1}$ then $B = P_{\lambda_1}\{T_1 > d_1\}$ with $d_1 = c_{k-2}$. Now by the same argument as before it follows that B and hence the r.h.s. of (A.3) is minimized with respect to λ_1 when $\lambda_1 = \Delta$. By repeating this argument for the other λ_i's $(2 \le i \le m)$ it follows that for fixed m $(1 \le m \le k)$, the LFC of the λ_i's occurs when $\lambda_i = \Delta$ for $1 \le i \le m$ and $\lambda_i = -\infty$ for $m+1 \le i \le k$, i.e., when $\theta_i = \delta\sigma$ for $1 \le i \le m$ and $\theta_i = -\infty$ for $m+1 \le i \le k$.

Next consider the case $0 < \varrho < 1$. As in (4.1) we can express

$$T_i = Z_i\sqrt{1-\varrho} - Z_0\sqrt{\varrho} \quad (1 \le i \le k) \tag{A.4}$$

where the Z_i are independent $N(\xi_i, 1)$ r.v.'s with $\xi_0 = 0$ and $\xi_i = \lambda_i/\sqrt{1-\varrho}$ $(1 \le i \le k)$. Note that finding the LFC of the θ_i's or of the λ_i's is equivalent to finding the LFC of the ξ_i's since the θ_i's, λ_i's and ξ_i's are common known multiples of each other. Now (A.1) changes to

$$P_\xi\{\text{reject } H_{01}^{(1)}, \dots, H_{0m}^{(1)}\} = \int \cdots \int A(\xi_1, \dots, \xi_m; z_{m+1}, \dots, z_k) \prod_{i=m+1}^{k} \phi_i(z_i)\, dz_i\, \phi_0(z_0)\, dz_0 \qquad (A.5)$$

and (A.2) changes to

$$A(\xi_1, \dots, \xi_m; z_{m+1}, \dots, z_k) = P_\xi\{(Z_1, \dots, Z_m) \overset{*}{>} (d_1 + z_0\sqrt{\varrho}, \dots, d_m + z_0\sqrt{\varrho})\} \qquad (A.6)$$

where $\phi_i(z_i)$ is the density function of Z_i $(0 \le i \le k)$ and $(d_1, \dots, d_m)$ has the same meaning as in the case of $\varrho = 0$.

To find the LFC of the ξ_i's over $\xi_i \ge \Delta/\sqrt{1-\varrho}$ $(1 \le i \le m)$ and $\xi_i \le 0$ $(m+1 \le i \le k)$ for fixed m $(1 \le m \le k)$, we can consider z_0 fixed. This takes us back to the setup for $\varrho = 0$, and hence the LFC result for $\varrho > 0$ follows. □

Proof of Corollary 1. For SS1 the above proof of the LFC of the θ_i's for fixed m $(1 \le m \le k)$ works since SS1 is a special case of SD1 with $c_1 = c_2 = \cdots = c_k$. Furthermore, it is also clear that under any LFC $\boldsymbol{\theta}$ such that $\theta_i = \delta\sigma$ for $1 \le i \le m$ and $\theta_i = -\infty$ for $m+1 \le i \le k$,

$$P_m(\text{SS1}) = P_{\boldsymbol{\theta}}\{T_1 > c_k, \dots, T_m > c_k\}$$

is minimum when $m = k$. □

Proof of Theorem 2. We are assuming $\varrho = 0$ in which case the T_i are independent $N(\lambda_i, 1)$ r.v.'s $(1 \le i \le k)$. Consider a parameter configuration of the θ_i's or equivalently that of the λ_i's, which for fixed m $(1 \le m \le k)$ satisfies

$$|\lambda_i| \ge \Delta = \frac{\delta}{\sqrt{v}} \quad (1 \le i \le m), \qquad |\lambda_i| < \Delta \quad (m+1 \le i \le k).$$

We can write

$$P_{\boldsymbol{\lambda}}\{\text{reject } H_{01}^{(2)}, \dots, H_{0m}^{(2)}\} = \int \cdots \int C(\lambda_1, \dots, \lambda_m; |t_{m+1}|, \dots, |t_k|) \prod_{i=m+1}^{k} \phi_i(t_i)\, dt_i, \qquad (A.7)$$

where

$$C(\lambda_1, \dots, \lambda_m; |t_{m+1}|, \dots, |t_k|) = P_{\boldsymbol{\lambda}}\{(|T_1|, \dots, |T_m|) \overset{*}{>} (d_1, \dots, d_m)\}, \qquad (A.8)$$

and as before, $(d_1, \ldots, d_m)$ is some subset of the critical constants $(c_1, \ldots, c_k)$ depending on the conditioned values $|t_{m+1}|, \ldots, |t_k|$. Now the r.h.s. of (A.8) can be written as

$$\int \cdots \int D(\lambda_1; |t_2|, \ldots, |t_m|) \prod_{i=2}^{m} \phi_i(t_i)\, \mathrm{d}t_i \tag{A.9}$$

where, as in the definition of B, $D(\lambda_1; |t_2|, \ldots, |t_m|) = 0$ or $P_{\lambda_1}\{|T_1| > e_1\}$ for some $e_1 \in \{d_1, \ldots, d_m\}$, the actual value of e_1 being a function of the conditioned values $|t_2|, \ldots, |t_m|$. Now $P_{\lambda_1}\{|T_1| > e_1\}$ is increasing in $|\lambda_1|$ because the distribution of T_1 is unimodal and symmetric about λ_1. (This is the only property of the normal distribution needed in the proof.) Therefore it follows that (A.9) is minimized over $|\lambda_1| \geq \Delta$ when $|\lambda_1| = \Delta$. By repeating this argument for the other λ_i's $(2 \leq i \leq m)$ it follows that at the LFC we have $|\lambda_1| = \cdots = |\lambda_m| = \Delta$.

We next address the problem of minimization of (A.7) with respect to $\lambda_{m+1}, \ldots, \lambda_k$. Toward this end, fix $|\lambda_1|, \ldots, |\lambda_m|$ (say, equal to their least favorable value Δ) and $|t_{m+1}|, \ldots, |t_{k-1}|$, and regard C defined by (A.8) as a function of $|t_k|$ only, denoting it by $F(|t_k|)$. To study the behavior of $F(|t_k|)$, first suppose $|t_k| = 0$. Since $|t_{m+1}|, \ldots, |t_{k-1}|$ are fixed, this fixes a particular choice of the critical constants $d_1, d_2, \ldots, d_m$ in (A.8), say, $d_1^*, d_2^*, \ldots, d_m^*$ (recall that $(d_1^*, \ldots, d_m^*) \subseteq \{c_1, \ldots, c_k\}$) where, without loss of generality, we can take $d_1^* \leq d_2^* \leq \cdots \leq d_m^*$. Thus

$$\begin{aligned} F(0) &= P_{\lambda}\{(|T_1|, |T_2|, \ldots, |T_m|) \succ^* (d_1^*, d_2^*, \ldots, d_m^*)\} \\ &= G(d_1^*, d_2^*, \ldots, d_m^*) \quad \text{(say).} \end{aligned} \tag{A.10}$$

In fact, for $0 \leq |t_k| \leq d_1^*$, we see that

$$F(|t_k|) = G(d_1^*, d_2^*, \ldots, d_m^*).$$

As soon as $|t_k|$ exceeds d_1^* with $|t_{m+1}|, \ldots, |t_{k-1}|$ remaining fixed, the choice of $(d_1, d_2, \ldots, d_m)$ changes to $(d_0^*, d_2^*, \ldots, d_m^*)$ where $d_0^* \leq d_1^*$. This choice remains unchanged until $|t_k|$ exceeds d_2^* when it changes to $(d_0^*, d_1^*, d_3^*, \ldots, d_m^*)$. Continuing in this manner, we obtain

$$F(|t_k|) = \begin{cases} F_0 = G(d_1^*, d_2^*, \ldots, d_m^*) & \text{for} \quad 0 \leq |t_k| \leq d_1^*, \\ F_1 = G(d_0^*, d_2^*, \ldots, d_m^*) & \text{for} \quad d_1^* < |t_k| \leq d_2^*, \\ F_2 = G(d_0^*, d_1^*, d_3^*, \ldots, d_m^*) & \text{for} \quad d_2^* < |t_k| \leq d_3^*, \\ \quad \vdots & \qquad \vdots \\ F_{m-1} = G(d_0^*, d_1^*, \ldots, d_{m-2}^*, d_m^*) & \text{for} \quad d_{m-1}^* < |t_k| \leq d_m^*, \\ F_m = G(d_0^*, d_1^*, \ldots, d_{m-1}^*) & \text{for} \quad |t_k| > d_m^*. \end{cases} \tag{A.11}$$

For instance, suppose $k=4$ and $m=3$. If $|t_4| \leq c_2$ then we have $(d_1, d_2, d_3) = (d_1^*, d_2^*, d_3^*) = (c_2, c_3, c_4)$ and $F(|t_4|) = G(c_2, c_3, c_4)$. If $c_2 < |t_4| \leq c_3$ then $(d_1, d_2, d_3) = (d_0^*, d_2^*, d_3^*) = (c_1, c_3, c_4)$ and $F(|t_4|) = G(c_1, c_3, c_4)$, and so on.

It is easy to see that

$$F_0 \leq F_1 \leq \cdots \leq F_m \tag{A.12}$$

since F_{j+1} is obtained from F_j by replacing d_{j+1}^* by d_j^* in the argument of $G(\cdot,\cdot,\ldots,\cdot)$ $(0\le j\le m-1)$, and $d_j^*\le d_{j+1}^*$ while $G(\cdot,\cdot,\ldots,\cdot)$ is decreasing in each one of its arguments. Now (A.7) can be written as

$$\int\cdots\int\left\{\int F(|t_k|)\phi_k(t_k)\,\mathrm{d}t_k\right\}\prod_{i=m+1}^{k-1}\phi_i(t_i)\,\mathrm{d}t_i$$
$$=\int\cdots\int\left\{\sum_{j=0}^{m}F_j\,p_j(\lambda_k)\right\}\prod_{i=m+1}^{k-1}\phi_i(t_i)\,\mathrm{d}t_i \tag{A.13}$$

where $p_0(\lambda_k)=P_{\lambda_k}\{|T_k|\le d_1^*\}$,

$$p_j(\lambda_k)=P_{\lambda_k}\{d_j^*<|T_k|\le d_{j+1}^*\}\quad(1\le j\le m-1)$$

and $p_m(\lambda_k)=P_{\lambda_k}\{|T_k|>d_m^*\}$. Note that each $p_j(\lambda_k)=p_j(-\lambda_k)$ and $\sum_{j=0}^m p_j(\lambda_k)=1$. We can minimize (A.13) with respect to λ_k (where we may take $\lambda_k\ge 0$ without loss of generality) by minimizing $\sum_{j=0}^m F_j\,p_j(\lambda_k)$. Now for $0\le\lambda_k'<\lambda_k$ we have

$$\sum_{j=l}^{m}p_j(\lambda_k')<\sum_{j=l}^{m}p_j(\lambda_k)\quad\text{for } 1\le l\le m. \tag{A.14}$$

Therefore

$$\sum_{j=0}^{m}F_j\,p_j(\lambda_k')=F_0+\sum_{l=1}^{m}(F_l-F_{l-1})\sum_{j=l}^{m}p_j(\lambda_k')$$
$$\le F_0+\sum_{l=1}^{m}(F_l-F_{l-1})\sum_{j=l}^{m}p_j(\lambda_k)\quad\text{(using (A.12) and (A.14))}$$
$$=\sum_{j=0}^{m}F_j\,p_j(\lambda_k). \tag{A.15}$$

Therefore $\sum_{j=0}^m F_j\,p_j(\lambda_k)$ is minimized with respect $\lambda_k\ge 0$ when $\lambda_k=0$. By repeating this argument for $i=m+1,\ldots,k-1$ we obtain that the LFC of the λ_i's occurs when $|\lambda_i|=\Delta$ for $1\le i\le m$ and $\lambda_i=0$ for $m+1\le i\le k$, i.e., when $|\theta_i|=\delta\sigma$ for $1\le i\le m$ and $\theta_i=0$ for $m+1\le i\le k$. □

Proof of Corollary 2. Analogous to the proof of the corollary to Theorem 1. □

Proof of Corollary 3. In this case we must make a correct directional decision for every θ_i with $|\theta_i|\ge\delta\sigma$. Because of the symmetry of the distribution of the T_i in the λ_i and hence in the θ_i, without loss of generality, we can take all such θ_i's to be positive. Hence (A.8) changes to

$$C(\lambda_1,\ldots,\lambda_m;|t_{m+1}|,\ldots,|t_k|)=P_\lambda\{(T_1,\ldots,T_m)\overset{*}{>}(d_1,\ldots,d_m)\}.$$

The proof then carries through essentially as before with only the obvious modifications necessary to take into account one-sided decisions for the θ_i's with $\theta_i\ge\delta\sigma$. □

For $0<\varrho<1$ it can be shown for fixed m $(1\leq m\leq k)$ that the LFC will have $|\theta_i|=\delta\sigma$ $(1\leq i\leq m)$ but it will not in general be true that $|\theta_i|=0$ $(m+1\leq i\leq k)$. The values of the θ_i $(m+1\leq i\leq k)$ at the LFC depend in a complicated way on the critical constants c_i and ϱ. To see this, suppose we use the transformation (A.4) to reduce the problem to the independence case. Then $|t_k|$ in (A.11) is replaced by $|z_k\sqrt{1-\varrho}-z_0\sqrt{\varrho}|$. For fixed $z_0\neq 0$, the intervals of z_k-values over which $F_0, F_1, \ldots, F_m$ are defined in (A.11) are not symmetric around zero. Therefore the inequalities (A.14) and (A.15) do not hold. Another way to see this is the following: Suppose we do not reduce the problem to the independence case. Then in (A.8) the distribution of $(|T_1|, \ldots, |T_m|)$ depends on the conditioned values $|t_{m+1}|, \ldots, |t_k|$. As a result, the F_j defined in (A.11) are functions of $|t_k|$, and hence the ordering given in (A.12) does not in general hold.

Acknowledgement

This work was carried out when the second author was visiting the University of Bath in the summer of 1988 with support from a Visiting Research Fellowship awarded by the Science and Engineering Research Council of U.K. The authors thank the referee for some useful suggestions.

References

Bechhofer, R.E. and C.W. Dunnett (1988). Tables of percentage points of multivariate Student t distributions. *Selected Tables in Mathematical Statistics, Vol.* **11**. Amer. Math. Soc., Providence, RI.

Bofinger, E. (1987). Step-down procedures for comparison with a control. *Austral. J. Statist.* **29**, 348–364.

Duncan, D.B. (1955). Multiple range and multiple F tests. *Biometrics* **11**, 1–42.

Dunnett, C.W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50**, 1096–1121.

Dunnett, C.W. and A.C. Tamhane (1991). Step-down multiple tests for comparing treatments with a control in unbalanced one-way layouts. *Statist. in Medicine*, to appear.

Hochberg, Y. and A.C. Tamhane (1987). *Multiple Comparison Procedures*. John Wiley, New York.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6**, 65–70.

Hsu, J.C. (1988). Sample size computation for designing multiple comparison experiments. *Comput. Statist. Data Anal.* **7**, 79–91.

Keuls, M. (1952). The use of the 'Studentized range' in connection with an analysis of variance. *Euphytica* **1**, 112–122.

Marcus, R., E. Peritz and K.R. Gabriel (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* **63**, 655–660.

Miller, R.G., Jr. (1966). *Simultaneous Statistical Inference*. McGraw Hill, New York.

Naik, U.D. (1975). Some selection rules for comparing p processes with a standard. *Commun. Statist. Ser. A* **4**, 519–535.

Newman, D. (1939). The distribution of the range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika* **31**, 20–30.

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

Shaffer, J.P. (1980). Control of directional errors with stagewise multiple test procedures. *Ann. Statist.* **8**, 1342–1348.

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Syst. Tech. J.* **41**, 463–501.

Stefánsson, G., W. Kim and J.C. Hsu (1988). Confidence sets in multiple comparisons. In: S.S. Gupta and J.O. Berger, Eds., *Statistical Decision Theory and Related Topics, IV, Vol.* **2**, 89–104.

Tukey, J.W. (1953). *The Problem of Multiple Comparisons*. Mimeographed monograph, Princeton University.