Biom. J. 31 (1989) 5, 545–561 Akademie-Verlag Berlin

Two-Stage Procedures for Comparing Treatments with a Control: Elimination at the First Stage and Estimation at the Second Stage

ROBERT E. BECHHOFER

Cornell University

CHARLES W. DUNNETT

McMaster University

AJIT C. TAMHANE

Northwestern University

Summary

We consider the problem of comparing a set of p_1 test treatments with a control treatment. This is to be accomplished in two stages as follows: In the first stage, N_1 observations are allocated among the p_1 treatments and the control, and the subset selection procedure of GUPTA and SOBEL (1958) is employed to eliminate "inferior" treatments. In the second stage, N_2 observations are allocated among the (randomly) selected subset of $p_2(\leqq p_1)$ treatments and the control, and joint confidence interval estimates of the treatment versus control differences are calculated using DUNNETT's (1955) procedure. Here both N_1 and N_2 are assumed to be fixed in advance, and the so-called square root rule is used to allocate observations among the treatments and the control in each stage.

Dunnett's procedure is applied using two different types of estimates of the treatment versus control mean differences: The unpooled estimates are based on only the data obtained in the second stage, while the pooled estimates are based on the data obtained in both stages. The procedure based on unpooled estimates uses the critical point from a p_2-variate Student t-distribution, while that based on pooled estimates uses the critical point from a p_1-variate Student t-distribution. The two procedures and a composite of the two are compared via Monte Carlo simulation. It is shown that the expected value of p_2 determines which procedure yields shorter confidence intervals on the average. Extensions of the procedures to the case of unequal sample sizes are given. Applicability of the proposed two-stage procedures to a drug screening problem is discussed.

Key words: Subset selection; Joint confidence interval estimation; Multiple comparisons with a control; Drug screening; Gupta-Sobel procedure; Dunnett procedure; Pooled estimates; Unpooled estimates.

1. Introduction

Two types of inferential goals have been proposed in the literature for use in problems involving test treatments versus control comparisons. One of these pertains to the *elimination* of test treatments that are "inferior" to the control treatment. The test treatments that are selected as being "superior" (or "equal") to the control treatment can then be studied more intensively in later experimentation. The other goal pertains to the *joint estimation* of the test treatment versus control differences with stated precision. The reasons for employing *joint* rather than *separate* estimation are explained in BECHHOFER and TAMHANE (1988); also see HOCHBERG and TAMHANE (1987, Chapter 1).

These two inferential goals have been treated separately in the literature. For the first goal, GUPTA and SOBEL (1958) proposed a subset selection procedure (referred to herein as the *GS-procedure*), while for the second goal, DUNNETT (1955) proposed a joint confidence interval estimation procedure (referred to herein as the *D-procedure*). In this paper we study a two-stage approach: The first stage uses the *GS*-procedure to eliminate the apparently inferior test treatments, while the second stage uses the *D*-procedure to estimate by joint confidence intervals (one-sided or two-sided) the performances of the retained test treatments relative to the control or placebo. In fact, we study two different procedures for second stage estimation. The first uses only the data obtained in the second stage for constructing the joint confidence intervals, and is referred to as the *Not Pool D-* (*ND-*) *procedure*. The second pools the data obtained in both stages, and is referred to as the *Pool D-* (*PD-*) *procedure*. Relative performances of the two procedures are studied via simulation.

The outline of the paper is as follows: Section 2 introduces the notation and states the basic assumptions. Section 3 provides descriptions of the two two-stage procedures. Section 4 discusses the so-called square root rule for allocating the total number of observations in each stage among the test treatments and the control treatment. Section 5 gives a numerical example to illustrate the procedures. Section 6 compares of the *ND-* and *PD*-procedures, and two of their variants, using numerical and simulation results. Section 7 describes extensions of the procedures to the case of arbitrary, unequal sample sizes on the test treatments. Section 8 discusses the application of the proposed two-stage procedures in a problem of drug screening.

2. Notation and Assumptions

We assume that at the first (elimination) stage of experimentation there are available $p_1 \geqq 2$ test treatments labelled 1, 2, ..., p_1 and a control treatment labelled 0. Let $\{Y_{ij1}(1 \leqq j \leqq n_{i1})\}$ denote a random sample of size n_{i1} on the ith treatment $(0 \leqq i \leqq p_1)$ with $N_1 = \sum_{i=0}^{p_1} n_{i1}$ being the given total sample size used at the first stage. As in the usual fixed-effects one-way layout model, the random sample on the ith treatment is assumed to be drawn from an $N(\mu_i, \sigma^2)$ distribution $(0 \leqq i \leqq p_1)$. Here the μ_i and σ^2 are *unknown* parameters. Let $\bar{Y}_{i1} = \sum_{j=1}^{n_{i1}} Y_{ij1}/n_{i1}$ denote the first stage sample mean for the ith treatment $(0 \leqq i \leqq p_1)$ and let

$$S_{\nu_1}^2 = \frac{\sum_{i=0}^{p_1} \sum_{j=1}^{n_{i1}} (Y_{ij1} - \bar{Y}_{i1})^2}{N_1 - (p_1 + 1)}$$

denote the first stage pooled sample variance based on $\nu_1 = N_1 - (p_1 + 1)$ degrees of freedom (d.f.).

The corresponding quantities in the second (estimation) stage are denoted by substituting subscript 2 in place of 1 in the above. Thus p_2 denotes the (random) number of test treatments retained for experimentation in the second stage. (If $p_2 = 0$ then there is no second stage experiment.) Without loss of generality we assume that the test treatments are labelled so that the first p_2 test treatments are retained. The total sample size N_2 for the second stage is assumed to be *fixed* in advance; this is allocated among the p_2 test treatments and the control treatment so that n_{i2} observations are taken on the ith treatment $(0 \leqq i \leqq p_2)$ with $N_2 = \sum_{i=0}^{p_2} n_{i2}$. Let $\bar{Y}_{i2} = \sum_{j=1}^{n_{i2}} Y_{ij2}/n_{i2}$ denote the second stage sample mean for the ith treatment $(0 \leqq i \leqq p_2)$ and let

$$S_{\nu_2}^2 = \frac{\sum_{i=0}^{p_2} \sum_{j=1}^{n_{i2}} (Y_{ij2} - \bar{Y}_{i2})^2}{N_2 - (p_2 + 1)}$$

denote the second stage pooled sample variance based on $\nu_2 = N_2 - (p_2 + 1)$ d.f. Note that both the n_{i2} and ν_2 are random variables, although $N_2 = \sum_{i=0}^{p_2} n_{i2}$ is fixed.

We will assume in Sections 3 through 6 that, based on symmetry considerations, $n_{i1} = n_1$ (say) for $i = 1, \ldots, p_1$ and $n_{i2} = n_2$ (say) for $i = 1, \ldots, p_2$. Thus $N_1 = n_{01} + p_1 n_1$ and $N_2 = n_{02} + p_2 n_2$. In Section 7 we will show how our procedures can be implemented when the sample sizes on the test treatments are not equal.

3. Two-Stage Procedures

In this section we describe our two two-stage procedures. Both procedures have the same goals for Stage 1 and for Stage 2. The goal for Stage 1 (*Goal* 1) is to select a subset of the p_1 test treatments which contains all of the treatments having means $\mu_i \geqq \mu_0$. (These test treatments are referred to as "superior.") If this goal is achieved then a *correct selection* (*CS*) is said to have been made. The goal for Stage 2 (*Goal* 2) is to estimate by means of joint confidence intervals the p_2 differences $\mu_i - \mu_0$ $(1 \leqq i \leqq p_2)$. For one-sided intervals this latter goal is referred to as Goal 2-*I* and for two-sided intervals it is referred to as Goal 2-II.

The *probability requirement* for Goal 1 is:

$$P(CS) \geqq 1 - \alpha_1 \quad \text{for all} \quad (\mu_0, \mu_1, \ldots, \mu_{p_1}; \sigma^2)\,, \tag{3.1}$$

and that for Goal 2 is:

$$\text{Joint Confidence Coefficient} \geqq 1 - \alpha_2 \text{ for all } (\mu_0, \mu_1, \ldots, \mu_{p_2}; \sigma^2)\,. \tag{3.2}$$

Here $1-\alpha_1$ and $1-\alpha_2$ are prespecified numbers between 0 and 1. Notice that $\{\mu_1, \ldots, \mu_{p_2}\}$ is a *random* subset of $\{\mu_1, \ldots, \mu_{p_1}\}$, and the requirement (3.2) must be guaranteed *unconditionally*. This can be achieved by guaranteeing (3.2) *conditionally* for every possible subset.

Both two-stage procedures use the *GS*-procedure in Stage 1 to guarantee (3.1). The *GS*-procedure retains the test treatment i in the selected subset for second stage experimentation iff

$$\bar{Y}_{i1} \geqq \bar{Y}_{01} - g_{\nu_1} S_{\nu_1} \left(\frac{1}{n_1} + \frac{1}{n_{01}}\right)^{1/2} \quad (1 \leqq i \leqq p_1) . \tag{3.3}$$

Here $g_{\nu_1} = g_{\nu_1, p_1, \varrho_1, \alpha_1}$ is the $100\alpha_1$ equicoordinate percentage point of the p_1-variate Student t-distribution with ν_1 d.f. and associated common correlation coefficient $\varrho_1 = n_1/(n_1+n_{01})$. We refer to the quantity $g_{\nu_1} S_{\nu_1} (1/n_1 + 1/n_{01})^{1/2}$ as the *allowance* associated with the *GS*-procedure.

The *ND*-procedure employs the following joint one-sided and two-sided confidence intervals, respectively, in the second stage:

$$\left\{\mu_i - \mu_0 \geqq \bar{Y}_{i2} - \bar{Y}_{02} - g_{\nu_2} S_{\nu_2} \left(\frac{1}{n_2} + \frac{1}{n_{02}}\right)^{1/2} \quad (1 \leqq i \leqq p_2)\right\} \quad \text{(for Goal 2-I)} \tag{3.4a}$$

and

$$\left\{\mu_i - \mu_0 \in \left[\bar{Y}_{i2} - \bar{Y}_{02} \pm h_{\nu_2} S_{\nu_2} \left(\frac{1}{n_2} + \frac{1}{n_{02}}\right)^{1/2}\right] \quad (1 \leqq i \leqq p_2)\right\} \tag{3.4b}$$

(for Goal 2-II) .

Here $g_{\nu_2} = g_{\nu_2, p_2, \varrho_2, \alpha_2}$ and $h_{\nu_2} = h_{\nu_2, p_2, \varrho_2, \alpha_2}$ are the upper $100\alpha_2$ equicoordinate percentage points of the p_2-variate Student t- and $|t|$-distributions, respectively, with ν_2 d.f. and associated common correlation coefficient $\varrho_2 = n_2/(n_2+n_{02})$. The quantities $g_{\nu_2} S_{\nu_2} (1/n_2 + 1/n_{02})^{1/2}$ and $h_{\nu_2} S_{\nu_2} (1/n_2 + 1/n_{02})^{1/2}$ are referred to as the *allowances* associated with the joint confidence intervals (3.4a) and (3.4b), respectively. To date the most complete and accurate tables of the critical points g_ν and h_ν have been given by BECHHOFER and DUNNETT (1988).

We now describe the *PD*-procedure. The pooled estimates on which the *PD*-procedure is based are calculated as follows: Let

$$D_i = \frac{D_{i1}\tau_2^2 + D_{i2}\tau_1^2}{\tau_1^2 + \tau_2^2} \tag{3.5}$$

denote the pooled estimate of $\mu_i - \mu_0$ $(1 \leqq i \leqq p_2)$ where

$$D_{ik} = \bar{Y}_{ik} - \bar{Y}_{0k} \quad (1 \leqq i \leqq p_2,\ k = 1, 2) \tag{3.6}$$

and

$$\tau_k^2 = \frac{1}{n_k} + \frac{1}{n_{0k}} \quad (k = 1, 2) . \tag{3.7}$$

Also let

$$S_v^2 = \frac{\nu_1 S_{\nu_1}^2 + \nu_2 S_{\nu_2}^2}{\nu_1 + \nu_2} \tag{3.8}$$

be a pooled estimator of σ^2 based on $\nu = \nu_1 + \nu_2$ d.f.

Finally let

$$\tau^2 = (1/\tau_1^2 + 1/\tau_2^2)^{-1} \tag{3.9}$$

and

$$\varrho = \frac{\varrho_1 \tau_2^2 + \varrho_2 \tau_1^2}{\tau_1^2 + \tau_2^2} . \tag{3.10}$$

The *PD*-procedure employs the following joint confidence intervals at the second stage:

$$\{\mu_i - \mu_0 \geqq D_i - g_{\nu,p_1,\varrho,\alpha_2} S_\nu \tau \ (1 \leqq i \leqq p_2)\} \quad \text{(for Goal 2-I)} \tag{3.11a}$$

and

$$\{\mu_i - \mu_0 \in [D_i \pm h_{\nu,p_2,\varrho,\alpha_2} S_\nu \tau] \ (1 \leqq i \leqq p_2)\} \quad \text{(for Goal 2-II)} . \tag{3.11b}$$

The quantities $g_{\nu,p_1,\varrho,\alpha_2} S_\nu \tau$ and $h_{\nu,p_1,\varrho,\alpha_2} S_\nu \tau$ are referred to as the *allowances* associated with the joint confidence intervals (3.11a) and (3.11b), respectively.

The intuitive reasoning behind these "pooled" intervals is as follows: If the random nature of p_2 (and hence that of n_{02}, n_2 and ν_2) is ignored, then D_i given by (3.5) is the "best" (minimum variance unbiased) pooled estimator of $\mu_i - \mu_0$ among all linear combinations of D_{i1} and D_{i2}; this minimum variance is equal to $\sigma^2\tau^2$ where τ^2 is given by (3.9). Also note that in this case the D_i are equicorrelated with common correlation coefficient ϱ given by (3.10).

We now turn to the question of whether or not the joint confidence intervals (3.4) and (3.11) associated with the *ND*- and *PD*-procedures, respectively, guarantee the probability requirement (3.2) for Goal 2. We can restrict the discussion to the one-sided intervals in each case since the same arguments apply to the two-sided intervals. For the joint one-sided confidence intervals (3.4a) associated with the *ND*-procedure, it is easy to see that they have an unconditional joint confidence coefficient $\geqq 1-\alpha_2$. This is so because conditional on the subset selected (assuming it is nonempty), p_2 and hence n_2, n_{02} are fixed. Therefore conditionally, the random variables

$$\frac{\bar{Y}_{i2} - \bar{Y}_{02} - (\mu_i - \mu_0)}{S_{\nu_2}\left(\frac{1}{n_2} + \frac{1}{n_{02}}\right)^{1/2}} \quad (1 \leqq i \leqq p_2)$$

cave a joint p_2-variate Student t-distribution with ν_2 d.f. and common correlation hoefficient $= \varrho_2 = n_2/(n_2 + n_{02})$. Hence the conditional joint confidence coefficient for the intervals (3.4a) is $1-\alpha_2$ if $p_2 \geqq 1$, and it may be taken to be unity if an

empty subset is selected, i.e., if $p_2=0$. Therefore the unconditional joint confidence coefficient is $\geqq 1-\alpha_2$.

No such rigorous argument can be given for the joint one-sided confidence intervals (3.11a) associated with the *PD*-procedure. The reason for this is that conditional on the subset selected (assuming it is nonempty), the random variables

$$\frac{D_i-(\mu_i-\mu_0)}{S_\nu\tau} \quad (1\leqq i\leqq p_2)$$

do not have a p_2-variate Student t-distribution even though p_2 and hence n_2, n_{02} are fixed. This is so because conditioning on a subset selected using (3.3) restricts the D_{i1} to be greater than or equal to $-g_{\nu_i}S_{\nu_i}(1/n_1+1/n_{01})^{1/2}$ for $1\leqq i\leqq p_2$, and hence the conditional distribution of the D_i is not p_2-variate normal. In fact, the individual D_i are not even conditionally univariate normal. It should also be noted that in (3.11a) we use the percentage point from the p_1-variate Student t-distribution even though the joint confidence statement is made concerning only $p_2(\leqq p_1)$ differences $\mu_i-\mu_0$ $(1\leqq i\leqq p_2)$. This is needed to compensate for the fact that the pooled estimates D_i are based in part on the first stage data, which have already been used to select the treatments for the first stage. Note that this compensation tends to make the procedure conservative. In the simulation experiment described in Section 6, we will examine the effect of using the percentage point from the p_2-variate Student t-distribution instead of one from the p_1-variate.

4. Allocation of Observations

In this section we discuss the choice of (n_{0k}, n_k) to be used in each stage $k=1, 2$. The particular choice that we recommend is based on the well-known square root allocation rule (DUNNETT (1955)) which yields

$$(4.1)\qquad n_{0k}=n_{0k}^*=\frac{N_k}{1+\sqrt{p_k}}, \quad n_k=n_k^*=\frac{N_k}{\sqrt{p_k}\,(1+\sqrt{p_k})} \quad (k=1, 2);$$

As discussed below, for stage k this choice approximately minimizes the expected allowance associated with (3.3) (for $k=1$) and (3.4) (for $k=2$) and exactly minimizes (ignoring the integer restrictions on n_{0k} and n_k) the common value of var $(\bar{Y}_{ik}-\bar{Y}_{0k})=\tau_k^2\sigma^2$ subject to given N_k and p_k and specified $1-\alpha_k$ $(k=1, 2)$.

The expected allowance associated with the *GS*-procedure is given by

$$(4.2)\qquad g_{\nu_1,p_1,\varrho_1,\alpha_1}\left(\frac{1}{n_{01}}+\frac{1}{n_1}\right)^{1/2}E(S_{\nu_1})$$

$$=g_{\nu_1,p_1,\varrho_1,\alpha_1}\{(r_1+p_1)(r_1+1)/r_1\}^{1/2}E(S_{\nu_1})/\sqrt{N_1}$$

where we have let $r_1=n_{01}/n_1$ and $\varrho_1=1/(1+r_1)$. Note that the minimizing value of r_1 is independent of σ. If the *ND*-procedure for one-sided comparisons is used in the second stage then the criterion to be minimized is the same as (4.2) but with

subscript 1 changed to 2 everywhere. If the ND-procedure for two-sided comparisons is used in the second stage then, in addition, the critical constant g_{ν_2} must be replaced by h_{ν_2}. For the PD-procedure the corresponding criteria can be stated in an analogous manner, but they are functions of the first stage quantities as well. We shall indicate later in this section how the allocation (4.1) can be justified for the PD-procedure.

Minimization of the expected allowance criterion has a clear interpretation for the joint confidence interval estimation problem. For the subset selection problem we use the same criterion because decreasing the expected allowance has the effect of decreasing the expected number of "inferior" test treatments (i.e., those having means $\mu_i < \mu_0$) included in the selected subset.

In Bechhofer, Dunnett and Tamhane (1987) we demonstrated by extensive numerical calculations that the choice $r_k = r_k^* = \sqrt{p_k}$, which gives the square root allocation (4.1), approximately minimizes the criterion (4.2) (as well as the criterion obtained by replacing g_{ν_k} with h_{ν_k}). For this allocation rule, ϱ_k equals $1/(1+\sqrt{p_k})$. The corresponding critical constants g_ν and h_ν for $\varrho = 1/(1+\sqrt{p})$ needed to implement the GS- and ND-procedures were tabulated for selected values of p, ν and α in the aforementioned article. (A subset of these tables may also be found in Bechhofer and Dunnett (1988).) The asymptotic (as $N_k \to \infty$) optimality of (4.1) for joint confidence interval estimation was shown by Bechhofer (1969) for one-sided comparisons and by Bechhofer and Nocturne (1972) for two-sided comparisons.

As we pointed out, the square root allocation rule (4.1) exactly minimizes $\mathrm{var}(\bar{Y}_{ik} - \bar{Y}_{0k}) = \tau_k^2 \sigma^2$ subject to given N_k and p_k ($k = 1, 2$). (This results in only *approximate* minimization of the expected allowance criterion because the critical constants g_{ν_k} and h_{ν_k}, which also are functions of $r_k = n_{0k}/n_k$ through $\varrho_k = 1/(1+r_k)$, do not vary much with ϱ_k for small α_k ($k = 1, 2$).) Therefore for stage $k = 2$, (4.1) exactly minimizes $\tau^2 \sigma^2$ for any given τ_1^2, and N_2 and p_2.

The expected allowance (conditioned on p_2) associated with the PD-procedure (3.11) is proportional to τ; moreover, the critical constants $g_{\nu,p_1,\varrho,\alpha_2}$ and $h_{\nu,p_1,\varrho,\alpha_2}$ are relatively insensitive to the choice of r_2. Therefore it follows that (4.1) also approximately minimizes the expected allowance associated with the PD-procedure.

In practice, the (n_{0k}^*, n_k^*) values given by (4.1) must be rounded to one of the nearest integer values, which are

$$(n_{0k}, n_k) = \begin{cases} (N_k - p_k[n_k^*], [n_k^*]) & \text{(4.3a)} \\ (N_k - p_k[n_k^* + 1], [n_k^* + 1]) & \text{(4.3b)} \end{cases} \quad (k = 1, 2),$$

where $[x]$ denotes the integer part of x. The choice between (4.3a) and (4.3b) should be based on the minimum expected allowance criterion. To make this comparison, the critical constants associated with the two allocations (4.3a) and

(4.3b) with their respective ϱ-values are needed. Linear interpolation with respect to $1/(1-\varrho)$ in the tables of BECHHOFER and DUNNETT (1988) is recommended for this purpose if the exact values are not readily available.

5. Numerical Example

Suppose that a pharmaceutical laboratory has identified 20 chemical compounds which it wishes to test against an existing standard drug (or placebo). The testing is to be done in two stages. The purpose of the first stage is to eliminate those compounds which are indicated as being "inferior" to the control compound so that more observations can be allocated in the second stage to the compounds retained which are presumably the "superior" ones. Suppose that 70 animals are available for each stage of testing.

For the first stage it is desired to design an efficient experiment so that with probability at least 0.90 the selected subset will contain all test compounds at least as good as the control compound; thus $1-\alpha_1=0.90$. Here $p_1=20$, $N_1=70$; hence there will be $\nu_1=N_1-(p_1+1)=70-21=49$ d.f. available for estimating σ^2. From (4.1) we see that the asymptotically optimal allocation is

$$n_{01}^*=\frac{N_1}{1+\sqrt{p_1}}=\frac{70}{1+\sqrt{20}}=12.79 \quad \text{and} \quad n_1^*=\frac{N_1}{\sqrt{p_1}\,(1+\sqrt{p_1})}=\frac{70}{\sqrt{20}\,(1+\sqrt{20})}=2.86\,.$$

From (4.3) the corresponding rounded pairs of integers are given by $(n_{01}, n_1)=$ $=(30,2)$ and $(10,3)$; the associated ϱ_1-values are 0.0625 and 0.2308, respectively. The corresponding critical constants are $g_{49,20,.0625,.10}=2.633$ and $g_{49,20,.2308,.10}=2.572$. The expected allowance is proportional to $2.633\,(1/30+1/2)^{1/2}=1.923$ if $(n_{01}, n_1)=$ $=(30,2)$, and $2.572\,(1/10+1/3)^{1/2}=1.693$ if $(n_{01}, n_1)=(10,3)$. Thus the latter choice is preferred. (In fact, the only other choice for (n_{01}, n_1) is (50,1) which leads to $\varrho_1=0.0196$ and $g_{49,20,.0196,.10}=2.644$. This choice is clearly inferior to either of the other two. Hence the square root rule indeed yields the overall optimum allocation in this case.)

Thus in the first stage, 10 observations will be taken on the control compound and 3 observations on each of the 20 test compounds. Then those test compounds whose sample means $\bar{Y}_{i1}$ are no less than $\bar{Y}_{01}-1.693S_{\nu_1}$ will be retained for further experimentation.

Now suppose that 15 compounds are eliminated in the first stage, and 5 are retained in the selected subset. In the second stage it is desired to obtain, say, 95 % joint one-sided confidence interval estimates of the five differences $\mu_i-\mu_0\,(1\leqq i\leqq 5)$; thus $1-\alpha_2=0.95$. Here $p_2=5$, $N_2=70$; hence there will be $\nu_2=N_2-(p_2+1)=$ $=70-6=64$ d.f. available for estimating σ^2 from the second stage. From (4.1)

we obtain

$$n_{02}^{*}=\frac{N_2}{1+\sqrt{p_2}}=\frac{70}{1+\sqrt{5}}=21.63 \quad \text{and} \quad n_2^{*}=\frac{N_2}{\sqrt{p_2}\,(1+\sqrt{p_2})}=\frac{70}{\sqrt{5}\,(1+\sqrt{5})}=9.67\,.$$

Using (4.3) we find that the corresponding rounded pairs of integers are $(n_{02}, n_2)=$ $=(25,9)$ and $(20,10)$; the associated ϱ_2-values are 0.2647 and 1/3, respectively. Following the same steps as taken earlier, we find that $(n_{02}, n_2)=(20, 10)$ is the preferred choice for the one-sided ND-procedure, the necessary critical constant in this case being $g_{64,5,1/3,.05}=2.329$. This same choice $(n_{02}, n_2)=(20, 10)$ is also preferred for the two-sided ND-procedure, the necessary critical constant in this case being $h_{64,5,1/3,.05}=2.616$.

Now suppose that it is desired to employ the PD-procedure at the second stage. Using the square root allocation rule we are again led to choosing between $(n_{02}, n_2)=(25,9)$ or $(20,10)$. The corresponding τ_2^2-values are 0.1511 and 0.1500, respectively, and the ϱ_2-values are 0.2647 and 1/3, respectively. Assuming that $(n_{01}, n_1)=$ $=(10,3)$ is used in the first stage, we have $\tau_1^2=(1/10+1/3)=0.4333$ and $\varrho_1=0.2308$. Applying (3.9) and (3.10) we obtain $(\tau^2, \varrho)=(0.1120, 0.2558)$ for $(n_{02}, n_2)=(25,9)$ and $(\tau^2, \varrho)=(0.1114, 0.3070)$ for $(n_{02}, n_2)=(20,10)$. Since the latter choice yields a smaller τ^2 and larger ϱ, it is clear that it will yield the smaller expected allowance. The critical constants needed to implement the PD-procedure for this choice of (n_{02}, n_2) are $g_{113,20,.3070,.05}=2.784$ and $h_{113,20,.3070,.05}=3.039$ for one-sided and two-sided joint intervals, respectively; here the pooled d.f. are $\nu=49+64=113$.

6. A Comparison of Procedures

In this section we compare the performances of the ND- and PD-procedures via Monte Carlo simulation. We also study two variants of these two procedures for making joint confidence statements at the second stage. The rationale behind these variants will become clear after we make a preliminary comparison between the ND- and PD-procedures. For convenience, in this section we will refer to these two procedures as $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.

It is clear from the description of $\mathcal{S}_2$ that it will tend to be conservative if the true number, say q $(\leqq p_1)$, of "superior" test treatments is small relative to p_1. This is so because $\mathcal{S}_2$ uses the critical point from the p_1-variate Student t-distribution when, in fact, only p_2 $(\leqq p_1)$ joint confidence statements about *apparently* "superior" test treatments are made. A natural question to ask is whether a procedure that uses g_{ν,p_2,ν,α_2} in place of $g_{\nu,p_1,\varrho,\alpha_2}$ in (3.11a) will still guarantee the probability requirement (3.2) for Goal 2−I. We refer to such a procedure as $\mathcal{S}_3$. We will show by simulation that $\mathcal{S}_3$ does not guarantee (3.2) in all cases, i.e., it can be liberal.

Table I

One-Sided Allowances for the $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ Procedures
($p_1=20$, $N_1=70$, $n_{01}=10$, $n_1=3$, $N_2=70$, $1-\alpha_2=0.95$)

p_2	(n_{02}, n_2)	ϱ_2	τ_2	ϱ	τ	$g_{r_2,p_2,\varrho_2,\alpha_2}$	$g_{r,p_1,\varrho,\alpha_2}$	$g_{r,p_2,\varrho,\alpha_2}$	$\mathcal{S}_1$-Allowance $g_{r_2,p_2,\varrho_2,\alpha_2}\tau_2$	$\mathcal{S}_2$-Allowance $g_{r,p_1,\varrho,\alpha_2}\tau$	$\mathcal{S}_3$-Allowance $g_{r,p_2,\varrho,\alpha_2}\tau$
1	(35,35)	0.5000	0.2390	0.4686	0.2247	1.6676	2.7008	1.6651	0.3986	0.6082	0.3741
2	(28,21)	0.4286	0.2887	0.3967	0.2644	1.9581	2.7423	1.9501	0.5653	0.7251	0.5156
3	(28,14)	0.3333	0.3273	0.3130	0.2931	2.1326	2.7817	2.1224	0.6980	0.8153	0.6221
4	(22,12)	0.3529	0.3589	0.3250	0.3151	2.2403	2.7747	2.2272	0.8040	0.8743	0.7018
5	(20,10)	0.3333	0.3873	0.3070	0.3338	2.3285	2.7835	2.3145	0.9018	0.9291	0.7726
6	(22,8)	0.2667	0.4129	0.2565	0.3498	2.4116	2.8003	2.3894	0.9957	0.9795	0.8358
7	(21,7)	0.2500	0.4364	0.2441	0.3638	2.4730	2.8055	2.4487	1.0792	1.0206	0.8908
8	(22,6)	0.2143	0.4606	0.2197	0.3774	2.5308	2.8123	2.5005	1.1657	1.0614	0.9437
9	(16,6)	0.2727	0.4787	0.2582	0.3872	2.5606	2.8012	2.5344	1.2258	1.0846	0.9813
10	(20,5)	0.2000	0.5000	0.2113	0.3982	2.6175	2.8155	2.5815	1.3088	1.1211	1.0280
12	(22,4)	0.1539	0.5436	0.1851	0.4191	2.6960	2.8234	2.6511	1.4655	1.1833	1.1111
14	(14,4)	0.2222	0.5669	0.2259	0.4296	2.7371	2.8130	2.6949	1.5517	1.2085	1.1577
16	(22,3)	0.1200	0.6155	0.1717	0.4496	2.8129	2.8287	2.7540	1.7313	1.2718	1.2382
18	(16,3)	0.1590	0.6292	0.1933	0.4548	2.8499	2.8244	2.7895	1.7932	1.2845	1.2687
20	(10,3)	0.2308	0.6582	0.2308	0.4655	2.8689	2.8145	2.8145	1.8883	1.3101	1.3101

Based on the above discussion we can surmise that $\mathcal{S}_2$ will yield a wider allowance than $\mathcal{S}_1$ if p_2/p_1 is small and vice versa. We now compare the one-sided allowances for the two procedures for $p_1=20$, $p_2=1(1)\,10(2)\,20$, $N_1=70$, $N_2=70$, $(n_{01}, n_1)=(10,3)$ and thus $(\tau_1^2, \varrho_1)=(0.4333,\ 0.2308)$, $\nu_1=49$, $\nu_2=70-(p_2+1)$ and $1-\alpha_2=0.95$. In this comparison, for each given p_2, (n_{02}, n_2) is chosen using the square root allocation rule given in Section 4. Furthermore, the sampling variations in S_{ν_2} and S_ν are ignored because ν_2 and $\nu=\nu_1+\nu_2$ are large; both S_{ν_2} and S_ν are taken to be equal to unity, which is the assumed value of σ. The results are presented in Table I. The $\mathcal{S}_3$-allowances also are included in this table for additional information.

From Table I we see that among the three procedures, $\mathcal{S}_3$ yields the smallest allowance for all values of p_2; however, this is at the expense of not guaranteeing (3.2) in all cases, as noted above. Between $\mathcal{S}_1$ and $\mathcal{S}_2$, the former yields the smaller allowance for $p_2\leqq 5$ and the latter yields the smaller allowance for $p_2>5$. This observation suggests an adaptive composite procedure $\mathcal{S}_4$ which uses $\mathcal{S}_1$ for making the joint confidence statements at the second stage if $p_2\leqq 5$ and which uses $\mathcal{S}_2$ if $p_2>5$. (More generally, the precise value of p_2 at which the $\mathcal{S}_2$-allowance becomes smaller than the $\mathcal{S}_1$-allowance will depend on the values of p_1, N_1, N_2 and whether the square root or some other allocation rule is employed at each stage.)

In summary, the following four procedures, all of which use the *GS*-procedure for subset selection in the first stage, were compared in our simulation study:

$\mathcal{S}_1$: *ND*-procedure.
$\mathcal{S}_2$: *PD*-procedure.
$\mathcal{S}_3$: *PD*-procedure which uses $g_{\nu,p_2,\varrho,\alpha_2}$ instead of $g_{\nu,p_1,\varrho,\alpha_2}$ in (3.11a).
$\mathcal{S}_4$: Uses $\mathcal{S}_1$ if $p_2\leqq 5$ and $\mathcal{S}_2$ if $p_2>5$.

The procedures were simulated under seven different μ_i-configurations for $p_1=20$. The μ_i-configurations were chosen to cover three different values of q, the true number of "superior" test treatments ($q=5$, 10 and 20). Without loss of generality, throughout we assumed $\mu_0=0$ and $\sigma^2=1$. The μ_i-values for "superior" test treatments were taken to be equal to μ_0, but the μ_i-values for "inferior" test treatments were varied over the range -1 to -4 in different combinations. The seven configurations are listed below.

Config. 1: $\mu_1=\ldots=\mu_{20}=0$ $(q=20)$
Config. 2: $\mu_1=\ldots=\mu_{10}=-1$, $\mu_{11}=\ldots=\mu_{20}=0$ $(q=10)$
Config. 3: $\mu_1=\ldots=\mu_{10}=-2$, $\mu_{11}=\ldots=\mu_{20}=0$ $(q=10)$
Config. 4: $\mu_1=\ldots=\mu_{15}=-2$, $\mu_{16}=\ldots=\mu_{20}=0$ $(q=5)$
Config. 5: $\mu_1=\ldots=\mu_8=-4$, $\mu_9=\ldots=\mu_{15}=-2$, $\mu_{16}=\ldots=\mu_{20}=0$ $(q=5)$
Config. 6: $\mu_1=\ldots=\mu_{15}=-3$, $\mu_{16}=\ldots=\mu_{20}=0$ $(q=5)$
Config. 7: $\mu_1=\ldots=\mu_{15}=-4$, $\mu_{16}=\ldots=\mu_{20}=0$ $(q=5)$.

For each configuration a total of 50,000 independent simulation runs were performed. Each run consisted of two stages: In the first stage the mutually independent random variables $\bar{Y}_{01}\sim N(\mu_0, \sigma^2/n_{01})$, $\bar{Y}_{i1}\sim N(\mu_i, \sigma^2/n_1)$ $(1\leqq i\leqq p_1)$ and

$S^2_{\nu_1} \sim \sigma^2\chi^2_{\nu_1}/\nu_1$ were generated and the *GS*-procedure (3.3) was applied to select a subset of p_2 test treatments. Here we used $N_1 = 70$ with the associated square root allocation $(n_{01}, n_1) = (10,3)$ and $1-\alpha_1 = 0.90$. As noted in Section 5, in this case we have $\nu_1 = 49$, $\varrho_1 = 0.2308$ and $g_{49,20,.2308,.10} = 2.633$. In the second stage the mutually independent random variables (which are independent of the first stage random variables) $\bar{Y}_{02} \sim N(\mu_0, \sigma^2/n_{02})$, $\bar{Y}_{i2} \sim N(\mu_i, \sigma^2/n_2)$ $(1 \leqq i \leqq p_2)$ and $S^2_{\nu_2} \sim \sigma^2\chi^2_{\nu_2}/\nu_2$ were generated and each of the four procedures was applied to the resulting data to construct one-sided joint confidence intervals for $\mu_i - \mu_0$ for the selected test treatments. For a given procedure the proportion of runs that resulted in the correct coverage of *all* of the $\mu_i - \mu_0$ for the selected test treatments was used as an estimate of the joint confidence coefficient of that procedure. In the second stage we used $1-\alpha_2 = 0.95$ and $N_2 = 70$. For each given p_2 $(1 \leqq p_2 \leqq 20)$ we used the square root allocation (n_{02}, n_2) given in Table I such that $n_{02} + p_2 n_2 = N_2 = 70$. These allocations and the associated critical constants needed to implement the four procedures were determined in advance and stored in memory, so that they did not have to be recomputed each time.

All simulation experiments were performed on McMaster University's VAX-8600 computer using a Fortran program. IMSL subroutines GGNPM and GGCHS were used to generate the normal and chi-square random variables, respectively. A single simulation experiment consisting of 50,000 runs of the four procedures took approximately 5 minutes of CPU time at a rate of $ 20 per hour. The simulation results are reported in Table II.

Table II

Simulation Estimates of the Joint Confidence Coefficient and Expected Subset Size

Config. No.	$E(p_2)$	Procedures $\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$
1	19.87	.9505	.9500	.9499	.9500
2	18.37	.9510	.9499	.9476	.9499
3	13.14	.9497	.9489	.9311	.9489
4	9.60	.9507	.9508	.9183	.9470
5	7.22	.9524	.9697	.9321	.9609
6	5.37	.9494	.9737	.9278	.9437
7	4.97	.9498	.9843	.9472	.9490

The primary quantities of interest in Table II are the estimated joint confidence coefficients of the procedures $\mathcal{S}_i$. These are to be compared with the nominal level $1-\alpha_2 = 0.95$. In making this comparison it must be kept in mind that the standard error of each estimate is approximately $(.05 \times .95/50{,}000)^{1/2} \cong 0.0010$. Thus the estimated values would be expected to lie in the interval $0.95 \pm 2 \times 0.0010$ if the corresponding joint confidence coefficients are controlled at the nominal level of 0.95. Using this criterion we find that $\mathcal{S}_1$ controls the joint confidence

coefficient quite accurately at the nominal level; this is, of course, to be expected in view of the proof of this fact given in Section 3.

We next note that for large values of the expected subset size, $(E(p_2, \mathcal{S}_2))$ controls the joint confidence coefficient accurately at the nominal level but the conservatism of $\mathcal{S}_2$ increases with decreasing $E(p_2)$-values. $\mathcal{S}_2$ is extremely conservative for configurations 5, 6 and 7, which involve small values of q and large negative μ_i-values for the "inferior" test treatments. These latter configurations result in small expected subset sizes. This behavior of $\mathcal{S}_2$ is to be anticipated in view of our previous discussion.

Next note that $\mathcal{S}_3$ is liberal under all configurations except under configuration no. 1 and possibly under configuration no. 2. Thus $\mathcal{S}_3$ is not an acceptable procedure.

Finally $\mathcal{S}_4$ appears to control the joint confidence coefficient in most cases, but under one configuration (config. no. 6) it is liberal. Thus there is some question about its validity under all configurations.

As a matter of additional interest, in Table III we give the simulation estimates for each $\mathfrak{P}_i$ of the probability of the *joint* event that a correct selection is made in the first stage (i.e., all "superior" test treatments are included in the subset) *and*

Table III

Simulation Estimates of the Overall Probability of No Error

Config. No.	Procedures $\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$
1	.8574	.8536	.8536	.8536
2	.8981	.8954	.8933	.8954
3	.8963	.8936	.8767	.8936
4	.9215	.9207	.8890	.9175
5	.9243	.9406	.9037	.9329
6	.9196	.9426	.8976	.9142
7	.9201	.9529	.9166	.9196

all the $\mu_i - \mu_0$ for the selected test treatments are covered by their respective confidence intervals. We refer to this probability as the *overall probability of no error*.

The estimates in Table III may be compared with the nominal value $(1-\alpha_1)\times$ $\times(1-\alpha_2) = 0.90\times 0.95 = 0.855$, which is the overall probability of making no error under configuration no. 1 (the least favorable configuration for the GS-procedure) if the inferences in the two stages were *statistically independent*. However, this independence holds only for $\mathcal{S}_1$. We note that the probabilities for all the $\mathcal{S}_i$ are within two standard errors $(=2\times(0.855\times 0.145/50{,}000)^{1/2}\cong 0.0032)$ of the nominal value under configuration no. 1. For other configurations, the achieved probabilities for all the $\mathcal{S}_i$ are strictly higher than the nominal value because the first stage GS-procedure achieves $P(CS) > 1-\alpha_1 = .90$ under these more favorable configurations.

In conclusion, if the *unknown* proportion of "superior" test treatments (viz. q/p_1) is expected to be large (say, at least one quarter of the total number of test treatments are expected to be "superior") then $\mathcal{S}_2$ is the preferred procedure; otherwise $\mathcal{S}_1$ is the preferred procedure. $\mathcal{S}_4$ provides a compromise between the two procedures, and is a good practical alternative. $\mathcal{S}_3$ is not an acceptable procedure.

In practice it would seem wasteful not to pool the data from the two stages. Therefore, in future research it would be desirable to develop a less conservative version of the *PD*-procedure for small p_2-values, which can be recommended in all situations.

7. Extension to Unequal Sample Sizes

In the previous sections we assumed, based on symmetry considerations, that $n_{i1}=n_1$ for $i=1, \ldots, p_1$ and $n_{i2}=n_2$ for $i=1, \ldots, p_2$. In practice, however, even if the experiment is designed to be balanced, there are often losses of observations due to reasons which are not fully under the experimenter's control. Therefore the actual sample sizes on the treatments are not always equal. In this section we show how to implement our two-stage procedures when the sample sizes are arbitrary. We use the notation defined in Section 2.

The following modifications are necessary in the application of the *GS*-procedure used in the 1st stage and the *ND*-procedure if it is used in the 2nd stage: In (3.3), n_1 should be n_{i1} and $g_{\nu_1}=g_{\nu_1,p_1,\underset{\sim}{R}_1,\alpha_1}$. Similarly in (3.4a), n_2 should be n_{i2} and $g_{\nu_2}=$ $=g_{\nu_2,p_2,\underset{\sim}{R}_2,\alpha_2}$, and in (3.4b), n_2 should be n_{i2} and $h_{\nu_2}=h_{\nu_2,p_2,\underset{\sim}{R}_2,\alpha_2}$. Here for $g_{\nu_k,p_k,\underset{\sim}{R}_k,\alpha_k}$ ($h_{\nu_k,p_k,\underset{\sim}{R}_k,\alpha_k}$) for $k=1, 2$ is the upper $100\alpha_k$ equicoordinate percentage point of the p_k variate Student t- ($|t|$-) distribution with ν_k d.f. and associated correlation matrix $\underset{\sim}{R}_k=\{\varrho_{ijk}\}$ where

$$\varrho_{ijk}=\begin{cases}1 & (1\leqq i=j\leqq p_k)\\ \left[\dfrac{n_{ik}n_{jk}}{(n_{ik}+n_{0k})\,(n_{jk}+n_{0k})}\right]^{1/2} & (1\leqq i\neq j\leqq p_k)\,.\end{cases} \tag{7.1}$$

Remark 7.1: As noted in HOCHBERG and TAMHANE (1987, p. 141), the so-called product structure possessed by this correlation matrix makes the task of evaluating these percentage points on a computer relatively easy because the associated p_k-variate ($k=1, 2$) integral can be reduced to a bivariate integral. A computer program for this purpose has been written by DUNNETT (1984). Alternatively, a good approximation to the exact percentage point is provided by $g_{\nu_k,p_k,\bar{\varrho}_k,\alpha_k}$ where $\bar{\varrho}_k$ is the arithmetic average of the ϱ_{ijk} for $1\leqq i\neq j\leqq p_k$, $k=1, 2$; see HOCHBERG and TAMHANE (1987, p. 145). Extensive numerical studies by DUNNETT (1985) show that this approximation is slightly on the conservative side.

If the *PD*-procedure is used in the 2nd stage than the following modifications

are necessary in its application: The pooled estimate of $\mu_i - \mu_0$ is now given by

$$D_i = \frac{D_{i1}\tau_{i2}^2 + D_{i2}\tau_{i1}^2}{\tau_{i1}^1 + \tau_{i2}^2} \quad (1 \leqq i \leqq p_2) \tag{7.2}$$

where

$$\tau_{ik}^2 = \frac{1}{n_{ik}} + \frac{1}{n_{0k}} \quad (1 \leqq i \leqq p_k,\ k = 1, 2)\,. \tag{7.3}$$

Further in (3.11 a), τ is replaced by τ_i and $g_{\nu,p_1,\varrho,\alpha_2}$ is replaced by $g_{\nu,p_1,\underset{\sim}{R},\alpha_2}$; similarly in (3.11 b), τ is replaced by τ_i and $h_{\nu,p_1,\varrho,\alpha_2}$ is replaced by $h_{\nu,p_1,\underset{\sim}{R},\alpha_2}$. Here, as in (3.9), we have

$$\tau_i^2 = (1/\tau_{i1}^2 + 1/\tau_{i2}^2)^{-1} \qquad (1 \leqq i \leqq p_2)\,, \tag{7.4}$$

and $\underset{\sim}{R} = \{\varrho_{ij}\}$ where, as in (3.10), we have

$$\varrho_{ij} = \frac{\tau_{i2}\tau_{j2}\varrho_{ij1} + \tau_{i1}\tau_{j1}\varrho_{ij2}}{\{(\tau_{i2}^2 + \tau_{i2}^2)\,(\tau_{j1}^2 + \tau_{j2}^2)\}^{1/2}} \quad (1 \leqq i \neq j \leqq p_2)\,. \tag{7.5}$$

Note that the formulae (7.1)–(7.5) reduce to the corresponding formulae given in Section 3 in the balanced case.

8. Application to Drug Screening

Drug screening is generally an on-going program consisting of a series of experiments in each of which perhaps 20 to 30 chemical compounds are tested for some specific type of activity. The total number of compounds tested is very large, but the number of compounds included in each experiment depends on the available laboratory facilities and resources. The purpose of the screening is to eliminate the compounds which have little or no activity. A false positive is a compound which, although not having the desired level of activity, nevertheless by chance gives a result on the screening test that falls in the "acceptable" range. Even though the accept/reject rule used in the screening test may be designed to have a very small probability of such an occurrence, the actual number of false positives that accrue over a period of time may be quite large, perhaps even exceeding the number of true positives. Hence, before proceeding further with more definitive testing of the compounds that have been identified by the initial screening procedure, it may be desirable to carry out a special experiment to eliminate the false positives and obtain precise estimates of the biological activity of the compounds indicated as being true positives.

We suggest that the two-stage approach described in this article may be appropriate for such an experiment. The purpose of the first stage (*GS*-procedure) would be to eliminate most of the false positives accrued in previous screening tests, while the purpose of the second stage would be to estimate the activity levels of the retained treatments in the first stage relative to a reference standard. For

the latter, a known active compound would be used if one were available; otherwise an inactive control could be used.

An important design problem now arises, namely, how to allocate a fixed total amount of resources (e.g., a fixed total number of animals available to carry out the entire experiment) between the two stages. In other words, what are the "optimal" values of N_1 and N_2 for fixed given $N=N_1+N_2$. This problem is not easy to formulate mathematically. A reasonable formulation would involve the expected number of "inferior" test treatments retained at the first stage, and errors of estimation at the second stage as measured by the expected values of the allowances of joint confidence intervals (for specified values of $1-\alpha_1$ and $1-\alpha_2$). The solution would depend on unknown parameters, e.g., the actual proportion of "inferior" test treatments. For instance, if this proportion were thought to be small, it might be desirable to omit the first stage entirely and allocate all the available experimental resources to the second stage.

The goal associated with the *GS*-procedure (Goal 1) states that *all* test compounds with means $\mu_i \geqq \mu_0$ be included in the selected subset. (If an active compound is used as a reference standard then μ_0 would be its unknown mean. However, if an inactive control is used as a reference standard then μ_0 should be its mean plus a specified constant $\delta > 0$; here δ is the minimum threshold that the mean of the test compound must exceed that of the inactive control in order for it to be considered "superior.") In practice, the number of such compounds and also the number of compounds with means $\mu_i < \mu_0$ is unknown. The constants necessary to implement the *GS*-procedure are derived under the so-called "least favorable" configuration in which *all* test compounds are assumed to have means $\mu_i = \mu_0$ $(1 \leqq i \leqq p_1)$. However, this assumption may be much too conservative if the experimenter has reason to believe that a number of the compounds actually have mean values $\mu_i < \mu_0$. In this case, it may be modified as follows: Prior to the start of experimentation the experimenter may be prepared to state an upper bound $m_1 \leqq p_1$ on the number of true positive test compounds. Then the asymptotically optimal allocation is still given by (4.1) but the critical constant to be used in the associated *GS*-procedure (3.3) is reduced from $g_{\nu_1,p_1,\varrho_1,\alpha_1}$ to $g_{\nu_1,m_1,\varrho_1,\alpha_1}$ where ϱ_1 still equals $1/(1+\sqrt{p_1})$. In particular, if $m_1=1$ then this latter critical constant equals t_{ν_1,α_1} — the upper α_1-point of the *univariate* Student's t-distribution with ν_1 d.f. Tables of $g_{\nu_1,m_1,\varrho_1,\alpha_1}$ for $m_1=2, \ldots, p_1-1$ are not available. The case $m_1=p_1$ dealt with in the present paper would correspond to the situation in which a series of structurally related compounds are submitted together for testing. However, even in this situation the experimenter may not require that the selected subset contain *all* "superior" compounds; i.e., he may be satisfied with selecting only a specified fraction.

Another difficulty with the use of the *GS*-procedure is that since the number of treatments in the selected subset is random, problems may arise in the second stage (the *D*-procedure) if the total amount of experimentation (N_2) that can be carried out in that stage is fixed in advance, as is assumed in the present article.

Therefore one may wish to use a procedure with a prespecified upper bound on the number of test treatments in the selected subset as proposed by SANTNER (1975) for a different problem.

Acknowledgement

R. E. Bechhofer's research was partially supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University. C. W. Dunnett's research was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

BECHHOFER, R. E., 1969: Optimal allocation of observations when comparing several treatment with a control. In: *Multivariate Analysis-II* (ed., P. R. Krishnaiah), New York: Academic Press, 463–473.

BECHHOFER, R. E. and DUNNETT, C. W., 1988: Percentage points of multivariate Student t distributions. *Selected Tables in Mathematical Statistics*, Vol. 11, Providence, R. I.: American Mathematical Society.

BECHHOFER, R. E., DUNNETT, C. W. and TAMHANE, A. C., 1987: Optimal allocation of observations in subset selection and multiple comparisons with a control, and associated tables (with application to drug screening). Tech. Rep. No. 745, School of OR & IE, College of Engineering, Cornell University, Ithaca, NY 14853.

BECHHOFER, R. E. and NOCTURNE, D. J. M., 1972: Optimal allocation of observations when comparing several treatments with a control, II: Two-sided comparisons. Technometrics **14**, 423–436.

BECHHOFER, R. E. and TAMHANE, A. C., 1988: Discussion of 'Optimal designs for comparing treatments with controls,' by A. S. Hedayat, M. Jacroux and D. Majumdar. Statistical Science **3**, 477–480.

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., 1984: Probability integral of multivariate t. Unpublished manuscript.

DUNNETT, C. W., 1985: Multiple comparisons between several treatments and a specified treatment. Invited talk at the Spring ENAR Meeting, Raleigh, NC.

GUPTA, S. S. and SOBEL, M., 1958: On selecting a subset which contains all populations better than a standard. Ann. Math. Statist. **29**, 235–244.

HOCHBERG, Y. and TAMHANE, A. C., 1987: *Multiple Comparison Procedures*. New York: John Wiley.

SANTNER, T. J., 1975: A restricted subset selection approach to ranking and selection problems. Ann. Statist. **3**, 334–349.

Received May 1988
Revised Aug. 1988

Professor ROBERT E. BECHHOFER
School of OR & IE
Upson Hall
Cornell University
Ithaca, NY 14853, U.S.A.

Professor CHARLES W. DUNNETT
Mathematics and Statistics Dept.
Burke Science Building
McMaster University
Hamilton, Ontario
CANADA L8S 4K1

Professor AJIT C. TAMHANE
Department of Statistics
Northwestern University
2006 Sheridan Road
Evanston, IL 60208, U.S.A.