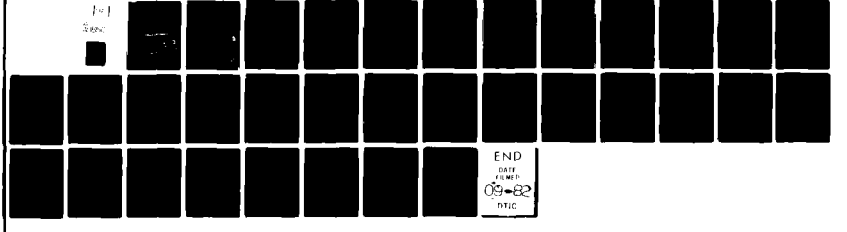
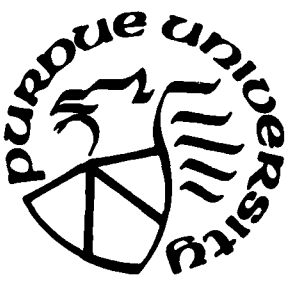
AD-A118 501 PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS F/G 12/1
A TWO-STAGE ELIMINATION TYPE PROCEDURE FOR SELECTING THE LARGES--ETC(U)
AUG 82 S S GUPTA, W KIM N00014-75-C-0455
UNCLASSIFIED TR-82-27 NL



END
DATE FILMED
09-82
DTIC

PURDUE UNIVERSITY



DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

(12)

A TWO-STAGE ELIMINATION TYPE PROCEDURE FOR SELECTING THE LARGEST OF SEVERAL NORMAL MEANS WITH A COMMON UNKNOWN VARIANCE

Shanti S. Gupta[1] and Woo-Chul Kim[2]
Purdue University Seoul National University

Technical Report #82-27

August 1982

[1]Research supported by ONR contract N00014-75-C-0455 at Purdue University, West Lafayette, IN, U.S.A. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2]Research supported in part by the Ministry of Educations, Korean Government, through the Research Institute of Basic Sciences, Seoul National University, Seoul, Korea.

A TWO-STAGE ELIMINATION TYPE PROCEDURE FOR SELECTING THE LARGEST OF SEVERAL NORMAL MEANS WITH A COMMON UNKNOWN VARIANCE

Shanti S. Gupta
Purdue University, W. Lafayette, Indiana, U.S.A.

Woo-Chul Kim
Seoul National University, Seoul, Korea

ABSTRACT

The problem of selecting the normal population with the largest mean from several populations with common unknown variance is considered. A new simple two-stage elimination type procedure is proposed which guarantees the same probability requirement using the indifference-zone approach as does the two-stage non-elimination type procedure of Bechhofer, Dunnett and Sobel (1954). The observations in the first stage are used not only to estimate the common unknown vairance but also to screen out non-contending populations. The proposed procedure has some similarities with the procedure of Tamhane (1975), but uses a new design criterion and a sharper lower bound on the probability of a correct selection. A table of constants necessary to implement the procedure is provided. The results of a Monte Carlo study of comparison of the elimination type procedure with the non-elimination type procedure are also given. The results strongly indicate that the proposed elimination type procedure performs much better than the non-elimination type procedure of Bechhofer, Dunnett and Sobel (1954) in terms of the expected total sample size.

1. INTRODUCTION

Following the work of Bechhofer (1954), considerable amount of research has been carried out for the problem of selecting the normal population with the largest mean from $k(k \geq 2)$ normal populations using the so-called indifference-zone approach.

For the case of common known variance, Alam (1970) proposed a two-stage elimination type procedure, in which he uses Gupta's (1956, 1965) subset selection procedure in the first stage to screen out non-contending populations and Bechhofer's (1954) indifference-zone approach to all populations retained in the second stage. However, most of his work was limited to the special case of $k=2$ populations. Recently, Tamhane and Bechhofer (1977, 1979) studies in detail a two-stage elimination type procedure using a minimax design criterion, and they found it more efficient than the single-stage procedure of Bechhofer (1954). For relevant results regarding various two-stage procedures, see Gupta and Miescke (1982a).

For the case of common unknown variance, it is known that there does not exist a single-stage procedure which guarantees a "minimum probability requirement", if one adopts the indifference-zone approach. Bechhofer, Dunnett and Sobel (1954) (henceforth referred to as BDS) proposed a two-stage non-elimination type procedure, in which the observations in the first stage are only used to obtain an estimate of the common unknown variance. Tamhane (1976) proposed a three-stage elimination type procedure, and carried out a Monte Carlo

study to compare his procedure with the non-elimination type procedure of BDS (1954). However, as pointed out by Hochberg and Marcus (1981), Tamhane's argument is not quite correct. Tamhane (1975) also considered a two-stage elimination type procedure for the case of common unknown variance. However, in his Monte Carlo sampling results, the performance of his two-stage elimination type procedure was found to be inferior to that of the non-elimination type procedure of BDS (1954).

In this paper, we propose and study a two-stage elimination type procedure for the case of common unknown variance. In Section 2, a precise formulation of the problem and a two-stage elimination type procedure are given. Section 3 discusses some possible design criteria, and the tables to implement the procedure are given. Section 4 gives the results of the Monte Carlo study in which we compare the two-stage elimination type procedure with the non-elimination type procedure of BDS (1954). The results show that the two-stage elimination type procedures perform much better than the non-elimination type procedure of BDS (1954).

It is worth noting that Paulson (1964) studies a sequential elimination type procedure for the case of common unknown variance, and Hochberg and Marcus (1981) considered three-stage elimination type procedures for the case of unequal and unknown variances.

Recently, Gupta and Miescke (1982b) and Miescke (1980) have studied two-stage selection procedures when we have a control population. For recent reviews of the literature, see Miescke (1982).

The reader is referred to the relevant book by Gupta and Panchapakesan (1979) for a complete overview.

2. A TWO-STAGE PROCEDURE

Let $\pi_1,\ldots,\pi_k (k\geq 2)$ be k normal populations with unknown means $\mu_1,\ldots,\mu_k$ and a common unknown variance σ^2. The ordered values of $\mu_1,\ldots,\mu_k$ are denoted by $\mu_{[1]}\leq\ldots\leq\mu_{[k]}$. No prior knowledge is assumed concerning the correct pairing between π_i and $\mu_{[j]}$ $(1\leq i,j\leq k)$.

The goal of the experimenter is to select the "best" population, i.e., the one associated with the largest mean $\mu_{[k]}$. Following the indifference-zone approach by Bechhofer (1954), the experimenter, prior to the experiment, specifies two constants $\delta^*>0$ and $P^*(1/k<P^*<1)$ which are incorporated into a probability requirement

$$P_{\underline{\theta}}\{CS\} \geq P^* \text{ for all } \underline{\theta}\in\Omega(\delta^*) \tag{2.1}$$

where CS denotes the event of selecting the best population and

$$\Omega(\delta^*) = \{\underline{\theta} = (\mu_1,\ldots,\mu_k,\sigma) \mid \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}.$$

The two-stage non-elimination type procedure of BDS (1954) and Paulson's (1964) sequential elimination type procedure satisfy the probability requirement (2.1). However, Tamhane's (1976) three-

stage procedure needs to be re-examined for the probability requirement (2.1) because of the error as pointed out previously.

The following two-stage elimination type procedure P can be viewed as a generalization of the non-elimination type procedure of BDS (1954).

Stage 1: Take n_1 ($n_1 \geq 2$) independent observations X_{ij} ($j=1,\ldots,n_1$) from each π_i ($i=1,\ldots,k$) and compute the sample means $\bar{X}_i^{(1)} = \sum_{j=1}^{n_1} X_{ij}/n_1$ ($i=1,\ldots,k$) and the pooled sample variance $S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_1} (X_{ij} - \bar{X}_i^{(1)})^2/k(n_1 - 1)$. Determine a subset I of $\{1,\ldots,k\}$ where

$$I = \{i \mid \bar{X}_i^{(1)} \geq \max_{1 \leq j \leq k} \bar{X}_i^{(1)} - (dS/\sqrt{n_1} - \delta^*)^+\},$$

the symbol a^+ denoting the positive part of a, i.e., $a^+ = \max(a,0)$; here d is a positive constant defiend in (2.2) below.

(a) If I has only one element, then stop sampling and assert that the population associated with $\max_{1 \leq j \leq k} \bar{X}_j^{(1)}$ is the best.

(b) If I has more than one element, then proceed to the second stage.

Stage 2: Take $N-n_1$ additional observations X_{ij} from each π_i ($i \in I$) where

$$N = \max\{n_1, [(hS/\delta^*)^2]\},$$

the symbol [y] denoting the smallest integer equal to or greater than y; here h is a positive constant defined in (2.2) below. Then compute the overall sample means

$$\bar{X}_i = \sum_{j=1}^{N} X_{ij}/N \quad (i \in I)$$

and assert that the population associated with $\max_{i \in I} \bar{X}_i$ is the best.

Remark 2.1: This procedure is essentially of the same type as the one in Tamhane (1975) even though we have a different type of screening procedure. However, it is important to note that we are using an improved lower bound on $P_{\underline{\theta}}\{CS\}$ to determine the constands d and h so as to satisfy the probability requirement (2.1).

Remark 2.2: Note that in the limiting case $d=\infty$, this procedure reduces to that of BDS (1954).

In the sequel, $\Phi(\cdot)$ denotes the c.d.f. of the standard normal distribution and $G_\nu(\cdot)$ denotes the c.d.f. of $(\chi^2(\nu)/\nu)^{1/2}$ with $\chi^2(\nu)$ denoting the chi-square random variable with $\nu=k(n_1-1)$ degrees of freedom.

Theorem 2.1. For the procedure P, we have the following inequalities:

$$\inf_{\Omega(\delta^*)} P_{\underline{\theta}}\{CS\} \geq \int_0^\infty \{\int_{-\infty}^\infty \Phi^{k-1}(x+dw)d\Phi(x) \int_{-\infty}^\infty \Phi^{k-1}(y+hw)d\Phi(y)\}dG_\nu(w)$$

$$\geq \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+dw)d\Phi(x)dG_\nu(w) \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+hw)d\Phi(x)dG_\nu(w).$$

Proof. Let $\delta_{ki} = \mu_{[k]} - \mu_{[i]}$, and let $\bar{X}_{(i)}^{(1)}$ and $\bar{X}_i$ denote respectively the first stage sample mean and the overall sample mean, i=1,...,k. Then, from Tamhane and Bechhofer (1979), we have the following inequalities for all $\underline{\theta} = (\mu_1,\ldots,\mu_k,\sigma) \in \Omega(\delta^*)$;

$$P_{\underline{\theta}}\{CS\} \geq P_{\underline{\theta}}\{\bar{X}_{(k)}^{(1)} \geq \bar{X}_{(i)}^{(1)} - (dS/\sqrt{n_1} - \delta^*)^+,\ \bar{X}_{(k)} \geq \bar{X}_{(i)},\ i=1,\ldots,k-1\}$$

$$= P_{\underline{\theta}}\{U_i \leq \frac{\sqrt{n_1}}{\sigma}\delta_{ki} + (d\frac{S}{\sigma} - \frac{\sqrt{n_1}}{\sigma}\delta^*)^+,\ V_i \leq \frac{\sqrt{N}}{\sigma}\delta^*,\ i=1,\ldots,k-1\}$$

where $U_i = (\bar{X}_{(i)}^{(1)} - \bar{X}_{(k)}^{(1)} + \delta_{ki})\sqrt{n_1}/\sigma$ and $V_i = (\bar{X}_{(i)} - \bar{X}_{(k)} + \delta_{ki})\sqrt{N}/\sigma$, i=1,...,k-1. Since $\sqrt{N} \geq hs/\delta^*$ and $\frac{\sqrt{n_1}}{\sigma}\delta_{ki} + (d\frac{S}{\sigma} - \frac{\sqrt{n_1}}{\sigma}\delta^*)^+ \geq d\frac{S}{\sigma}$ for i=1,...,k-1, it follows that for all $\underline{\theta} \in \Omega(\delta^*)$,

$$P_{\underline{\theta}}\{CS\} \geq P_{\underline{\theta}}\{U_i \leq d\frac{S}{\sigma},\ V_i \leq h\frac{S}{\sigma},\ i=1,\ldots,k-1\}.$$

Now, conditioning on $S/\sigma = w$ and applying Slepian's inequality as in Tamhane and Bechhofer (1979), we obtain the first inequality in the theorem. We note that the lower bound on $P_{\underline{\theta}}\{CS\}$ by the first inequality can be written as $E\{F_d(W)F_h(W)\}$ where $F_d(w) = \int_{-\infty}^{\infty} \Phi^{k-1}(x+dw)d\Phi(x)$ and $F_h(w) = \int_{-\infty}^{\infty} \Phi^{k-1}(x+hw)d\Phi(x)$. Since F_d and F_h are non-decreasing in w, it follows from Tchebysheff's inequality (Hardy, Littlewood and Polya (1934), Theorem 43) that

$$E\{F_d(W)F_h(W)\} \geq E\{F_d(W)\}E\{F_h(W)\}$$

which gives the second inequality in the theorem.

Now, the constants d and h in the procedure ρ can be chosen by setting either of the two lower bounds in Theorem 2.1 to be P* so as to satisfy the probability requirement (2.1).

Even though the second lower bound is less sharp than the first, it is much easier to evaluate. Furthermore, the functions in the second lower bound are the c.d.f.'s at equicoordinate point of a multivariate central t-distribution, which have been studied and tabluated by Dunnett (1955), Gupta and Sobel (1957) and Krishnaiah and Armitage (1966). Thus, we used the second lower bound to compare our two-stage elimination type procedure ρ with the two-stage non-elimination type procedure of BDS (1954). We state this in the next result.

Corollary 2.1. If $d > 0$ and $n > 0$ are chosen to satisfy

$$\int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+dw)d\Phi(x)dG_\nu(w) \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+hw)d\Phi(x)dG_\nu(w) = P^*, \quad (2.2)$$

then the procedure ρ satisfies the probability requirement (2.1).

We note that the lower bound on $P_\theta\{CS\}$ we are using is sharper than the one in Tamhane (1975), as can be seen in Section 11 of Tamhane and Bechhofer (1977).

3. A DESIGN CRITERION

The constants n_1, d and h in the procedure ρ can be considered

as design constants. As in the case of known σ^2, one might try to minimize the maximum of the expected total sample size. However, it can be easily seen that the maximum of the expected total sample size w.r.t. $\mu_1,\ldots,\mu_k$ depends on unknown σ^2 and increases indefinitely as σ^2 increases. Thus, as an alternative, we might ask for the procedure P to have a high probability of including the best in the first stage.

More precisely, we ask for the procedure P to satisfy the following requirement in addition to (2.1): For all $\underline{\theta} \in \Omega(\delta^*)$,

$$P_{\underline{\theta}}\{\text{including the best in the first stage}\} \geq P_1^* \tag{3.1}$$

where $P_1^*(P^*<P_1^*<1)$ is a preassigned constant.

It can be easily shown that, for the procedure $\mathcal{P}$,

$$\inf_{\Omega(\delta^*)} P_{\underline{\theta}}\{\text{including the best in the first stage}\}$$

$$= \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+dw)\,d\Phi(x)\,dG_\nu(w).$$

Therefore, if $d > 0$ and $h > 0$ are chosen to satisfy

$$\begin{aligned} \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+dw)\,d\Phi(x)\,dG_\nu(w) &= P_1^* \\ \int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+hw)\,d\Phi(x)\,dG_\nu(w) &= P^*/P_1^*, \end{aligned} \tag{3.2}$$

then the procedure $\mathcal{P}$ satisfies (2.1) and (3.1).

Note that the additional requirement (3.1) for the procedure ρ makes the choice of d and h unique once n_1 is given, and it can be regarded as a design criterion for d and h.

From now on, we will denote by $\rho_1(\rho_2)$ the procedure ρ satisfying (3.1) with $P_1^* = 0.99$ (0.975). Table 3.1 and Table 3.2 give the constants d and h for ρ_1 and ρ_2, respectively, for selected values of n_1, k and P*. In making these tables, the integrals on the right side of (3.2) were computed by using the result of Hartley (1943-46) and Gauss-Hermite quadrature formula as done in Gupta and Panchapakesan (1982).

Another simple way to make the choice of d and h unique is to assume that d=h. In such a case, (2.2) reduced to

$$\int_0^\infty \int_{-\infty}^\infty \Phi^{k-1}(x+dw)d\Phi(x)dG_\nu(w) = \sqrt{P^*} \qquad (3.3)$$

Note that this is equivalent to taking $P_1^* = \sqrt{P^*}$ in (3.2). In fact, this was the choice made in Tamhane's (1975) Monte Carlo study even though Tamhane used a lower bound different from the one used in (3.3). We will call by ρ_3 the procedure ρ satisfying (3.3). The constants d(=h) in procedure ρ_3 are given in Table 3.3 for selected values of n_1, k and P*.

Also, we give the constnats h in the procedure P(BDS) of BDS (1954) in Table 3.4. Some of these entries can also be found in Dunnett (1955), Gupta and Sobel (1957) and Krishnaiah and Armitage (1966). The values seem to agree well in most of the common cases

up to the second decimal point except a very few cases.

Finally, we remark that we, by no means, claim that (3.2) and (3.3) are the only possible design criteria. For example, the expected number, say S, of populations to be sampled at the second stage can be used in devising a design criterion. Furthermore, following the arguments in Section 6 of Tamhane and Bechhofer (1977), it can be shown that

$$\sup_{\Omega} E_{\underline{\theta}}\{S\} = k \int_{0}^{\infty} \int_{-\infty}^{\infty} \{\Phi^{k-1}(x+dw) - \Phi^{k-1}(x-dw)\} d\Phi(x) dG_{\nu}(w) \quad (3.4)$$

Hence, if one wishes to use S in devising a design criterion, the above equation might be useful.

TABLE 3.1: Constants d and h for the procedure $\mathcal{P}_1$

(a) constants d $\qquad P_1^* = 0.99$

k \ n_1	3	5	7	9	10	15	20	25
2	5.064	4.140	3.807	3.660	3.614	3.490	3.435	3.404
3	5.026	4.277	4.030	3.917	3.881	3.782	3.737	3.712
4	4.953	4.330	4.135	4.045	4.016	3.935	3.897	3.876
5	4.896	4.368	4.207	4.131	4.106	4.038	4.006	3.988
10	4.791	4.502	4.413	4.370	4.356	4.316	4.297	4.286
15	4.795	4.596	4.533	4.502	4.492	4.463	4.450	4.442

TABLE 3.1: (Continued)

(b) constants h

P* = 0.90

k \ n_1	3	5	7	9	10	15	20	25
2	2.291	2.068	2.005	1.974	1.964	1.936	1.923	1.916
3	2.677	2.473	2.413	2.384	2.375	2.348	2.335	2.328
4	2.847	2.672	2.620	2.594	2.586	2.562	2.551	2.544
5	2.953	2.802	2.756	2.733	2.726	2.704	2.694	2.689
10	3.225	3.134	3.105	3.090	3.085	3.071	3.065	3.061
15	3.368	3.301	3.279	3.269	3.265	3.255	3.250	3.247

P* = 0.95

k \ n_1	3	5	7	9	10	15	20	25
2	3.446	2.834	2.698	2.636	2.617	2.562	2.537	2.522
3	3.616	3.161	3.046	2.993	2.976	2.927	2.905	2.892
4	3.666	3.315	3.220	3.174	3.156	3.118	3.099	3.088
5	3.702	3.416	3.334	3.295	3.282	3.246	3.229	3.219
10	3.837	3.682	3.634	3.610	3.602	3.580	3.569	3.563
15	3.934	3.824	3.789	3.771	3.766	3.749	3.741	3.737

TABLE 3.2: Constants d and h for the procedure P_2

(a) constants d $\quad P_1^* = 0.975$

k \ n_1	3	5	7	9	10	15	20	25
2	4.152	3.284	3.087	3.000	2.972	2.897	2.863	2.844
3	4.177	3.553	3.396	3.324	3.301	3.237	3.208	3.191
4	4.151	3.675	3.547	3.488	3.468	3.415	3.390	3.375
5	4.139	3.755	3.648	3.597	3.580	3.534	3.512	3.500
10	4.178	3.978	3.916	3.886	3.876	3.848	3.835	3.828
15	4.243	4.104	4.060	4.038	4.031	4.010	4.001	3.995

(b) constants h $\quad P^* = 0.90$

k \ n_1	3	5	7	9	10	15	20	25
2	2.503	2.229	2.153	2.117	2.106	2.073	2.058	2.049
3	2.862	2.619	2.550	2.516	2.505	2.474	2.460	2.452
4	3.014	2.810	2.750	2.720	2.711	2.684	2.671	2.664
5	3.108	2.934	2.881	2.855	2.847	2.823	2.812	2.805
10	3.356	3.253	3.220	3.204	3.198	3.183	3.176	3.172
15	3.491	3.415	3.391	3.379	3.375	3.363	3.358	3.355

TABLE 3.2: (Continued)

(b) constants h $P^* = 0.95$

k \ n_1	3	5	7	9	10	15	20	25
2	4.120	3.260	3.066	2.981	2.954	2.880	2.846	2.827
3	4.149	3.533	3.378	3.307	3.284	3.221	3.192	3.176
4	4.126	3.656	3.530	3.472	3.453	3.400	3.395	3.361
5	4.116	3.737	3.631	3.581	3.565	3.519	3.498	3.485
10	4.160	3.962	3.902	3.872	3.862	3.835	3.822	3.814
15	4.227	4.089	4.046	4.024	4.017	3.997	3.987	3.982

TABLE 3.3: Constants d(=h) for the procedure P_3

$P^* = 0.90$

k \ n_1	3	5	7	9	10	15	20	25
2	3.070	2.610	2.500	2.449	2.432	2.386	2.365	2.353
3	3.329	2.963	2.867	2.821	2.806	2.765	2.746	2.735
4	3.422	3.132	3.050	3.011	2.998	2.963	2.946	2.936
5	3.482	3.241	3.171	3.137	3.126	3.094	3.080	3.071
10	3.663	3.528	3.485	3.464	3.457	3.438	3.429	3.423
15	3.774	3.677	3.646	3.631	3.626	3.611	3.604	3.600

TABLE 3.3: (Continued)

P* = 0.95

k \ n_1	3	5	7	9	10	15	20	25
2	4.136	3.272	3.076	2.991	2.963	2.889	2.855	2.835
3	4.163	3.543	3.387	3.315	3.293	3.229	3.200	3.183
4	4.139	3.665	3.539	3.480	3.461	3.407	3.382	3.368
5	4.128	3.746	3.639	3.589	3.573	3.526	3.505	3.493
10	4.169	3.970	3.909	3.879	3.869	3.842	3.828	3.821
15	4.235	4.097	4.053	4.031	4.024	4.004	3.994	3.988

TABLE 3.4: Constants h for $P^{}$ (BDS)

P* = 0.90

k \ n_1	3	5	7	9	10	15	20	25
2	2.173	1.976	1.918	1.890	1.881	1.856	1.844	1.838
3	2.573	2.389	2.334	2.307	2.298	2.273	2.262	2.255
4	2.752	2.593	2.544	2.520	2.513	2.490	2.480	2.474
5	2.864	2.726	2.682	2.661	2.654	2.635	2.625	2.620
10	3.149	3.064	3.037	3.023	3.019	3.006	3.000	2.996
15	3.297	3.234	3.214	3.204	3.200	3.191	3.186	3.184

TABLE 3.4: (Continued)

$P^* = 0.95$

k \ n_1	3	5	7	9	10	15	20	25
2	3.110	2.634	2.522	2.469	2.453	2.406	2.384	2.372
3	3.360	2.985	2.886	2.840	2.825	2.783	2.763	2.752
4	3.448	3.152	3.069	3.029	3.016	2.979	2.963	2.953
5	3.506	3.260	3.189	3.154	3.143	3.111	3.096	3.088
10	3.682	3.544	3.502	3.480	3.473	3.454	3.444	3.439
15	3.792	3.694	3.662	3.646	3.641	3.627	3.620	3.616

4. RESULTS OF MONTE CARLO COMPARISON

The elimination type procedure P in Section 2 can, in fact, be regarded as a class of procedures, since there are many ways of choosing the constants d and h satisfying (2.2). In particular, the procedures P_1, P_2, P_3 in Section 3 can be regarded as various versions of the elimination type procedure P. Hence, a comparison of the procedures P_1, P_2, P_3 with the non-elimination type procedure $P(BDS)$ can shed light on the effect of elimination in the first stage.

This section gives the results of Monte Carlo comparisons of the elimination type procedures P_1, P_2, P_3 with the non-elimination type procedure $P(BDS)$ in terms of the expected total sample sizes.

We will use respectively T and S as generic notations for the total sample size and the number of populations to be sampled in the second stage in using the elimination type procedure. Then, the total sample size of the elimination type procedure is given by

$$T = kn_1 + S(N-n_1) \tag{4.1}$$

where S is an integer-valued random variable which can take values 1 to k, inclusive. Now the total sample size of the non-elimination type procedure R(BDS) is given by T=kN. Expressions for the expected total sample sizes can be found in BDS (1954) and Tamhane (1975).

For the elimination type procedures in this simulation study, the estimates of the expected total sample sizes $E_{\underline{\theta}}\{T\}$ were computed at the so-called equal means configuration (EMC), $\mu_{[1]}=\ldots=\mu_{[k]}$, and at the least favorable configuration (LFC) $\mu_{[1]}=\ldots=\mu_{[k-1]}=\mu_{[k]}-\delta^*$. Note that $E_{\underline{\theta}}\{T\}$ of the procedure R(BDS) does not depend on $\mu_1,\ldots,\mu_k$.

The Monte Carlo comparisons were carried out for k=5,10,15, P*=0.90, 0.95 and δ^*=0.5,1.0. The choice of n_1, the first sample size for each population, is optional and we considered the procedures with n_1=5,10,15,20,25 to get some idea about the relationship between $E_{\underline{\theta}}\{T\}$ and n_1. For each (k,P^*,δ^*,n_1)-combination, we considered the cases of σ=1,2,3 and 1000 simulations were carried out. In each simulation, k independent standard normal random variables were generated by the random number generator of H. Rubin and C. Hinkle

at Purdue University. In this paper, we have reported the results only for $P^*=0.90$ and $\delta^*=1.0$ and the results for other cases indicate similar conclusions and are available upon request.

The estimates of the expected total sample sizes of the procedures R_1, R_2, R_3 and $R(BDS)$ are given in Table 4.1 along with their respective estimated standard errors. Throughout all the cases considered, the elimination type procedures R_1, R_2, R_3 have substantially smaller $E\{T\}$ values than corresponding $E\{T\}$ values of the non-elimination type procedure $R(BDS)$, except when the procedures stop at the first stage (these cases are denoted by a symbol + in Table 4.1 and Table 4.2). It can also be observed from Table 4.1 that the reduction in $E\{T\}$ by using the elimination type procedure increases with k for fixed n_1 and σ.

Remark 4.1: The superiority of the two-stage elimination type procedure to the two-stage non-elimination type procedure is what one might expect intuitively; however, this was not revealed in Tamhane's (1975) Monte Carlo study. We think that it was because he used a lower bound on $P_{\underline{\theta}}\{CS\}$ less sharp than the one in this paper, as pointed out in Remark 2.1.

From Table 4.1, we also observe that, as n_1 increases, $E\{T\}$ values decrease at the beginning and then increase. Such a relationship between n_1 and $E\{T\}$ for the procedure $R(BDS)$ has been given in BDS (1954). A glance at the results in Table 4.1shows what n_1 should be for the range of σ-values considered.

TABLE 4.1.

Monte Carlo estimates of the expected total sample size

The numbers in parentheses are the estimated standard errors.

(a) $k = 5$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E\{T\}$	$E_{LFC}\{T\}$			$E_{EMC}\{T\}$		
		$P(BDS)$	P_1	P_2	P_3	P_1	P_2	P_3
5	1	54.58 (0.52)	35.40 (0.38)	33.94 (0.37)	33.25 (0.41)	39.51 (0.39)	39.69 (0.43)	40.84 (0.54)
	2	210.70 (2.09)	148.27 (1.78)	152.00 (2.08)	164.89 (2.63)	152.49 (1.71)	158.93 (1.98)	177.77 (2.55)
	3	470.74 (4.64)	340.84 (3.75)	358.02 (4.38)	406.26 (5.76)	344.06 (3.68)	364.42 (4.29)	418.22 (5.67)
10	1	54.27 (0.21	50.07 (0.02)	50.11 (0.03)	50.09 (0.03)	50.24 (0.04)	50.41 (0.05)	50.60 (0.07)
	2	196.48 (1.31)	127.16 (1.30)	121.93 (1.50)	123.44 (1.88)	139.67 (1.18)	140.22 (1.41)	147.63 (1.83)
	3	441.16 (2.98)	309.84 (2.95)	316.83 (3.51)	346.34 (4.53)	319.76 (2.70)	330.52 (3.21)	366.72 (4.24)
15	1†	75.02 (0.01)	75.00 (0.00)	75.00 (0.00)	75.00 (0.00)	75.00 (0.00)	75.00 (0.00)	75.00 (0.00)
	2	192.64 (1.03)	119.17 (1.03)	114.86 (1.12)	113.16 (1.34)	134.06 (0.95)	132.92 (1.12)	136.49 (1.43)
	3	430.63 (2.23)	289.97 (2.50)	284.76 (3.06)	295.71 (4.04)	307.90 (2.20)	312.37 (2.72)	335.92 (3.64)
20	1†	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)
	2	191.39 (0.88)	120.44 (0.68)	117.62 (0.70)	116.23 (0.81)	134.88 (0.72)	134.36 (0.82)	136.53 (1.06)
	3	426.21 (1.91)	274.68 (2.38)	263.81 (2.79)	266.61 (3.64)	300.35 (2.06)	301.03 (2.51)	316.96 (3.43)
25	1†	125.00 (0.00)	125.00 (0.00)	125.00 (0.00)	125.00 (0.00)	125.00 (0.00)	125.00 (0.00)	125.00 (0.00)
	2	192.48 (0.79)	132.87 (0.38)	131.81 (0.39)	131.51 (0.47)	141.72 (0.50)	143.15 (0.57)	145.48 (0.78)
	3	425.19 (1.69)	264.75 (2.28)	253.33 (2.67)	252.59 (3.33)	295.78 (1.93)	292.60 (2.37)	303.53 (3.19)

TABLE 4.1.(Continued)

Monte Carlo estimates of the total expected sample size

(b) $k = 10$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E\{T\}$	$E_{LFC}\{T\}$			$E_{EMC}\{T\}$		
		$P(BDS)$	P_1	P_2	P_3	P_1	P_2	P_3
5	1	128.83 (0.83)	78.98 (0.76)	74.16 (0.74)	71.88 (0.80)	90.81 (0.76)	87.53 (0.80)	86.08 (0.92)
	2	498.20 (3.52)	364.09 (3.45)	362.10 (4.00)	381.45 (4.94)	373.56 (3.24)	377.95 (3.83)	404.01 (4.80)
	3	1123.91 (8.12)	846.50 (7.21)	880.14 (8.42)	965.32 (10.67)	863.58 (7.10)	894.05 (8.19)	986.66 (10.38)
10	1	124.11 (0.52)	100.91 (0.10)	100.81 (0.09)	100.74 (0.09)	102.81 (0.17)	103.04 (0.17)	103.34 (0.20)
	2	482.72 (2.27)	320.89 (2.81)	303.13 (3.17)	299.42 (3.75)	348.96 (2.47)	340.46 (2.93)	348.10 (3.61)
	3	1073.48 (5.12)	788.84 (5.63)	791.42 (6.83)	835.46 (8.71)	807.93 (5.22)	317.81 (6.33)	871.55 (8.15)
15	1†	150.18 (0.01)	150.00 (0.00)	150.00 (0.00)	150.00 (0.00)	150.00 (0.00)	150.00 (0.00)	150.01 (0.01)
	2	475.94 (1.76)	291.34 (2.41)	272.30 (2.50)	266.25 (2.82)	329.69 (2.14)	316.70 (2.43)	316.62 (2.90)
	3	1069.97 (4.11)	756.80 (5.25)	735.80 (6.47)	749.44 (8.06)	793.17 (4.51)	786.83 (5.66)	821.71 (7.28)
20	1†	200.00 (0.00)	200.00 (0.00)	200.00 (0.00)	200.00 (0.00)	200.00 (0.00)	200.00 (0.00)	200.00 (0.00)
	2	473.53 (1.61)	281.61 (1.84)	266.50 (1.78)	260.35 (1.92)	324.48 (1.71)	311.06 (1.91)	307.57 (2.21)
	3	1059.65 (3.34)	708.56 (5.35)	675.93 (6.17)	668.80 (7.38)	762.71 (4.37)	746.07 (5.52)	761.71 (7.07)
25	1†	250.00 (0.00)	250.00 (0.00)	250.00 (0.00)	250.00 (0.00)	250.00 (0.00)	250.00 (0.00)	250.00 (0.00)
	2	472.87 (1.39)	292.70 (1.21)	283.93 (1.17)	279.78 (1.26)	329.93 (1.32)	320.96 (1.42)	316.83 (1.68)
	3	1056.94 (3.01)	673.23 (5.20)	634.37 (5.88)	622.93 (6.97)	741.29 (4.23)	719.07 (5.19)	718.23 (6.53)

TABLE 4.1. (Continued)

Monte Carlo estimates of the total expected sample size

(c) $k = 15$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E\{T\}$	$E_{LFC}\{T\}$			$E_{EMC}\{T\}$		
		$P(BDS)$	P_1	P_2	P_3	P_1	P_2	P_3
5	1	210.78 (1.13)	126.81 (1.16)	117.35 (1.13)	112.85 (1.16)	146.52 (1.14)	138.25 (1.20)	134.84 (1.31)
	2	808.71 (4.66)	597.24 (4.93)	589.65 (5.85)	607.75 (7.90)	611.84 (4.62)	612.01 (5.55)	640.31 (6.89)
	3	1821.95 (10.60)	1411.98 (9.96)	1438.80 (11.76)	1550.68 (14.88)	1422.77 (9.75)	1455.05 (11.43)	1576.57 (14.39)
10	1	203.87 (0.74)	152.77 (0.19)	152.13 (0.16)	151.92 (0.17)	158.21 (0.32)	157.59 (0.29)	157.26 (0.31)
	2	795.60 (3.10)	536.59 (4.20)	504.62 (4.81)	493.40 (5.60)	578.50 (3.60)	559.16 (4.26)	561.12 (5.21)
	3	1769.78 (6.61)	1323.44 (7.88)	1314.81 (9.75)	1367.52 (12.47)	1347.80 (7.27)	1350.95 (9.17)	1413.69 (11.90)
15	1†	226.05 (0.14)	225.00 (0.00)	225.00 (0.00)	225.00 (0.00)	225.00 (0.00)	225.00 (0.00)	225.03 (0.01)
	2	787.41 (2.33)	485.11 (3.86)	447.10 (4.04)	430.15 (4.35)	543.54 (3.18)	514.51 (3.73)	503.75 (4.36)
	3	1772.42 (5.50)	1276.49 (7.65)	1233.22 (9.59)	1234.36 (11.90)	1324.92 (6.69)	1303.19 (8.52)	1328.29 (10.97)
20	1†	300.00 (0.00)	300.00 (0.00)	300.00 (0.00)	300.00 (0.00)	300.00 (0.00)	300.00 (0.00)	300.00 (0.00)
	2	785.07 (2.12)	458.84 (3.10)	426.93 (2.97)	412.68 (3.02)	530.50 (2.72)	500.33 (2.99)	485.76 (3.29)
	3	1757.37 (4.49)	1199.67 (8.30)	1133.44 (9.71)	1118.41 (11.36)	1281.32 (6.66)	1239.73 (8.42)	1245.92 (10.59)
25	1†	375.00 (0.00)	375.00 (0.00)	375.00 (0.00)	375.00 (0.00)	375.00 (0.00)	375.00 (0.00)	375.00 (0.00)
	2	784.77 (1.85)	465.40 (2.25)	444.78 (2.07)	433.39 (2.07)	532.02 (2.16)	508.85 (2.24)	496.63 (2.47)
	3	1753.50 (4.03)	1136.27 (8.13)	1061.76 (9.27)	1027.48 (10.73)	1241.22 (6.43)	1185.11 (7.91)	1172.15 (9.83)

To summarize the comparison, we computed the estimated relative savings RSAVE expressed in percentages in using an elimination type procedure P over the non-elimination type procedure $P(BDS)$, i.e.,

$$\text{RSAVE}\{P \mid P(BDS)\} = \frac{E\{T \mid P(BDS)\} - E\{T \mid P\}}{E\{T \mid P(BDS)\}} \times 100(\%)$$

These values are based on Table 4.1, and are given in Table 4.2.

It is observed from Table 4.2 that there is a more pronounced improvement of the elimination type procedure over the procedure $P(BDS)$ at the LFC than at the EMC. In fact, it can shown that, for fixed σ^2, $E_{\underline{\theta}}\{T\}$ of the elimination type procedure is maximized at the EMC when $\underline{\theta}=(\mu_1,\ldots,\mu_k,\sigma)$ vaires over the whole parameter space. Thus, even greater savings can be obtained by using the elimination type procedure than those indicated by the values in Table 4.2. Finally, it should be pointed out that the superiority of one procedure over others among the procedures P_1, P_2, P_3 are not quite clear from Table 4.2. Thus, it is recommended that the choice of d and h satisfying (2.2) may be made to depend on the experimenter's other requirements such as those described in Section 3. In this respect, we have computed the expected numbers of populations to be sampled in the second stage, $E\{S\}$, as by-products of a simulation study, which are given in Table 4.3.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. Joong Kweon Sohn for much help in computing the tables in Section 3.

TABLE 4.2.

Estimated relative savings expressed in percentages in using procedures P_1, P_2, P_3 over the procedure $P(BDS)$.

(a) $k = 5$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	LFC			EMC		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	35.1	37.8	39.1	27.6	27.3	25.2
	2	29.6	27.9	21.7	27.6	24.6	15.6
	3	27.6	23.9	13.7	26.9	22.6	11.2
10	1	7.7	7.7	7.7	7.4	7.1	6.7
	2	35.3	37.9	37.2	28.9	28.6	24.9
	3	29.8	28.2	21.5	27.5	25.1	16.9
15	1†	0.03	0.03	0.03	0.02	0.02	0.02
	2	38.1	40.4	41.3	30.4	31.0	29.1
	3	32.7	33.9	31.3	28.5	27.5	22.0
20	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	37.1	38.5	39.3	29.5	29.8	28.7
	3	35.6	38.1	37.5	29.5	29.4	25.6
25	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	31.0	31.5	31.7	26.4	25.6	24.4
	3	37.7	40.4	40.6	30.4	31.2	28.6

TABLE 4.2.(Continued)

Estimated relative savings expressed in percentages in using procedures P_1, P_2, P_3 over the procedure $P(BDS)$.

(b) $k = 10$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	LFC			EMC		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	38.7	42.4	44.2	29.5	32.1	33.2
	2	26.9	27.3	23.4	25.0	24.1	18.9
	3	23.8	21.7	14.1	23.2	20.5	12.2
10	1	18.7	18.8	18.8	17.2	17.0	16.7
	2	33.5	37.2	38.0	27.7	29.5	27.9
	3	26.5	26.3	22.2	24.7	23.8	18.8
15	1†	0.1	0.1	0.1	0.1	0.1	0.1
	2	38.8	42.8	44.1	30.7	33.5	33.5
	3	29.3	31.2	30.0	25.9	26.5	23.2
20	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	40.5	43.7	45.0	31.5	34.3	35.0
	3	33.1	36.2	36.9	28.0	29.6	28.1
25	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	38.1	40.0	40.8	30.2	32.1	33.0
	3	36.3	40.0	41.1	29.9	32.0	32.0

TABLE 4.2. (Continued)

Estimated relative savings expressed in percentages in using procedures P_1, P_2, P_3 over the procedure P(BDS).

(c) k = 15, δ^* = 1.0, P* = 0.90

n_1	σ	LFC			EMC		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	39.8	44.3	46.5	30.5	34.4	36.0
	2	26.1	27.1	24.9	24.3	24.3	20.8
	3	22.5	21.0	14.9	21.9	20.1	13.5
10	1	25.1	25.3	25.9	22.4	22.7	22.9
	2	32.6	36.6	38.0	27.3	29.7	29.5
	3	25.2	25.7	22.7	23.8	23.7	20.1
15	1†	0.5	0.5	0.5	0.5	0.5	0.5
	2	38.4	43.2	45.3	31.0	34.7	36.0
	3	28.0	30.4	30.4	25.2	26.5	25.1
20	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	41.6	45.6	47.4	32.4	36.3	38.1
	3	31.7	35.5	36.4	27.1	29.5	29.1
25	1†	0.0	0.0	0.0	0.0	0.0	0.0
	2	41.0	43.3	44.8	32.2	35.2	36.7
	3	35.2	39.4	41.4	29.2	32.4	33.2

TABLE 4.3.

Monte Carlo estimates of the expected number of populations to be sampled in the second stage. The numbers in parentheses are the estimated standard errors.

(a) $k = 5$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E_{LFC}\{S\}$			$E_{EMC}\{T\}$		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	2.46 (0.06)	1.70 (0.05)	1.06 (0.04)	3.80 (0.04)	3.00 (0.05)	2.10 (0.05)
	2	4.50 (0.03)	4.11 (0.04)	3.56 (0.05)	4.69 (0.02)	4.37 (0.03)	3.91 (0.04)
	3	4.76 (0.02)	4.51 (0.03)	4.14 (0.04)	4.82 (0.02)	4.61 (0.03)	4.28 (0.03)
10	1	0.39 (0.03)	0.16 (0.02)	0.03 (0.01)	1.99 (0.05)	0.92 (0.04)	0.20 (0.02)
	2	3.72 (0.04)	3.03 (0.05)	2.38 (0.05)	4.38 (0.03)	3.84 (0.04)	3.19 (0.05)
	3	4.47 (0.03)	4.12 (0.04)	3.69 (0.05)	4.66 (0.02)	4.36 (0.03)	3.96 (0.04)
15	1†	0.03 (0.01)	0.00 (0.00)	0.00 (0.00)	0.53 (0.03)	0.03 (0.01)	0.00 (0.00)
	2	2.86 (0.05)	2.17 (0.05)	1.51 (0.05)	3.91 (0.04)	3.21 (0.05)	2.46 (0.05)
	3	4.16 (0.03)	3.62 (0.04)	3.02 (0.05)	4.53 (0.02)	4.12 (0.03)	3.60 (0.04)
20	1†	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
	2	1.98 (0.05)	1.34 (0.05)	0.82 (0.04)	3.49 (0.04)	2.66 (0.05)	1.85 (0.05)
	3	3.80 (0.04)	3.15 (0.05)	2.50 (0.05)	4.37 (0.03)	3.88 (0.04)	3.27 (0.04)
25	1†	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	2	1.36 (0.05)	0.79 (0.04)	0.41 (0.03)	3.20 (0.05)	2.26 (0.05)	1.36 (0.04)
	3	3.40 (0.05)	2.72 (0.05)	2.07 (0.05)	4.18 (0.03)	3.59 (0.04)	2.92 (0.05)

TABLE 4.3. (Continued)

Monte Carlo estimates of the expected number of populations to be sampled in the second stage.

(b) $k = 10$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E_{LFC}\{S\}$			$E_{EMC}\{S\}$		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	4.98 (0.09)	3.56 (0.08)	2.45 (0.07)	7.26 (0.07)	5.71 (0.08)	4.13 (0.08)
	2	8.96 (0.05)	8.12 (0.06)	7.15 (0.07)	9.28 (0.04)	8.57 (0.06)	7.67 (0.07)
	3	9.54 (0.03)	9.03 (0.05)	8.35 (0.06)	9.64 (0.03)	9.20 (0.04)	8.56 (0.05)
10	1	1.03 (0.05)	0.55 (0.04)	0.19 (0.02)	3.91 (0.08)	2.21 (0.06)	0.89 (0.04)
	2	7.59 (0.07)	6.32 (0.08)	5.06 (0.08)	8.60 (0.06)	7.51 (0.07)	6.33 (0.07)
	3	9.05 (0.05)	8.35 (0.06)	7.45 (0.07)	9.31 (0.04)	8.68 (0.05)	7.83 (0.06)
15	1†	0.08 (0.01)	0.01 (0.01)	0.00 (0.00)	1.37 (0.05)	0.23 (0.02)	0.01 (0.00)
	2	5.98 (0.09)	4.61 (0.08)	3.47 (0.08)	7.64 (0.07)	6.29 (0.07)	4.97 (0.07)
	3	8.53 (0.05)	7.53 (0.07)	6.43 (0.07)	9.08 (0.04)	8.22 (0.05)	7.23 (0.06)
20	1†	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.07 (0.01)	0.00 (0.00)	0.00 (0.00)
	2	4.40 (0.09)	3.11 (0.08)	2.13 (0.06)	6.81 (0.07)	5.24 (0.07)	3.83 (0.07)
	3	7.80 (0.07)	6.65 (0.08)	5.39 (0.08)	8.64 (0.05)	7.63 (0.06)	6.47 (0.07)
25	1†	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	2	3.15 (0.08)	2.05 (0.07)	1.27 (0.05)	6.04 (0.07)	4.38 (0.07)	2.90 (0.06)
	3	7.03 (0.08)	5.77 (0.08)	4.55 (0.08)	8.19 (0.06)	7.07 (0.07)	5.74 (0.07)

TABLE 4.3. (Continued)

Monte Carlo estimates of the expected number of populations to be sampled in the second stage.

(c) $k = 15$, $\delta^* = 1.0$, $P^* = 0.90$

n_1	σ	$E_{LFC}\{S\}$			$E_{EMC}\{S\}$		
		P_1	P_2	P_3	P_1	P_2	P_3
5	1	7.55 (0.13)	5.47 (0.12)	3.87 (0.10)	10.63 (0.10)	8.29 (0.11)	6.19 (0.10)
	2	13.38 (0.07)	12.14 (0.09)	10.63 (0.10)	13.79 (0.06)	12.70 (0.08)	11.30 (0.10)
	3	14.29 (0.04)	13.52 (0.06)	12.48 (0.08)	14.41 (0.04)	13.69 (0.06)	12.72 (0.08)
10	1	1.67 (0.08)	0.88 (0.05)	0.42 (0.03)	5.52 (0.10)	3.32 (0.08)	1.65 (0.06)
	2	11.55 (0.10)	9.71 (0.11)	7.83 (0.11)	12.83 (0.07)	11.22 (0.09)	9.40 (0.10)
	3	13.62 (0.06)	12.54 (0.08)	11.18 (0.10)	13.91 (0.05)	12.93 (0.07)	11.61 (0.09)
15	1†	0.17 (0.02)	0.04 (0.01)	0.00 (0.00)	2.12 (0.06)	0.59 (0.04)	0.03 (0.01)
	2	9.23 (0.12)	7.14 (0.12)	5.38 (0.11)	11.36 (0.09)	9.33 (0.10)	7.32 (0.10)
	3	12.90 (0.07)	11.43 (0.09)	9.70 (0.10)	13.52 (0.06)	12.25 (0.08)	10.63 (0.09)
20	1†	0.01 (0.01)	0.00 (0.00)	0.00 (0.00)	0.19 (0.02)	0.00 (0.00)	0.00 (0.00
	2	6.89 (0.12)	4.89 (0.11)	3.42 (0.09)	10.07 (0.10)	7.75 (0.10)	5.65 (0.09)
	3	11.88 (0.10)	10.13 (0.11)	8.36 (0.11)	12.97 (0.07)	11.43 (0.09)	9.66 (0.10)
25	1†	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	2	4.95 (0.11)	3.29 (0.09)	2.06 (0.07)	8.77 (0.10)	6.43 (0.09)	4.38 (0.08)
	3	10.78 (0.10)	8.90 (0.11)	7.03 (0.11)	12.29 (0.08)	10.52 (0.09)	8.61 (0.10)

BIBLIOGRAPHY

Alam, K. (1970). A two-sample procedure for selecting the population with largest mean from k normal populations. Ann. Inst. Statist. Math. 22, 127-136.

Bechhofer, R. E. (1954). A single-sample multiple decision procedures for ranking means of normal populations with known variances. Ann. Math. Statist. 25, 16-39.

Bechhofer, R. E., Dunnett, C. W., and Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. Biometrika 41, 170-176.

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

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeo. Ser. No. 150). Inst. of Statist., Univ. of North Carolina, Chapel Hill.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 7, 225-245.

Gupta, S. S., and Miescke, K. J. (1982a). On the least favorable configurations in certain two-stage selection procedures. Statistics and Probability: Essays in Honor of C. R. Rao (Eds. Kallianpur, G. et al.), Amsterdam-New York-Oxford, North-Holland, pp. 295-305.

Gupta, S. S., and Miescke, K. J. (1982b). On the problem of finding a best population with respect to a control in two-stages. Statistical Decision Theory and Related Topics III, Vol. 1 (Eds. S. S. Gupta and J. O. Berger), New York, Academic Press, pp. 473-496.

Gupta, S. S., and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. New York, John Wiley.

Gupta, S. S., and Panchapakesan, S. (1982). Edgeworth Expansions in Statistics: A Brief Review. Technical Report #82-18, Dept. of Statist., Purdue University, W. Lafayette, IN.

Gupta, S. S., and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist. 28, 957-967.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). Inequalities, Cambridge University Press, Cambridge.

Hartley, H. O. (1943-46). Studentization or the elimination of standard deviation of the parent population from the random sample distribution of Statistics. Biometrika 33, 173-180.

Hochberg, Y., and Marcus, R. (1981). Three stage elimination type procedures for selecting the best normal population when variances are unknown. Commun. Statist. A10, 597-612.

Krishnaiah, P. R., and Armitage, J. V. (1966). Tables for multivariate t-distribution. Sankhyā, Ser. B 28, 31-56.

Miescke, K. J. (1980). On two-stage procedures for finding a population better than a control. Mimeo. Ser. #80-28, Dept. of Statist., Purdue University, W. Lafayette, IN.

Miescke, K. J. (1982). Recent results on multi-stage selection procedures. Technical Report, #82-25, Dept. of Statist., Purdue University, W. Lafayette, IN.

Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. Ann. Math. Statist. 35, 174-180.

Tamhane, A. C. (1975). On minimax multistage elimination type rules for selecting the largest normal mean. Technical Report No. 259, Dept. of Operations Research, Cornell University, Ithaca, New York.

Tamhane, A. C. (1976). A three-stage elimination type procedure for selecting the largest normal mean (common unknown variance). Sankhyā Ser. B 38, 339-349.

Tamhane, A. C., and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist.-Theor. Meth. A6, 1003-1033.

Tamhane, A. C., and Bechhofer, R. E. (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): an improved PCS lower bound and associated tables. Commun. Statist.-Theor. Meth. A8, 337-358.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report #82-27	2. GOVT ACCESSION NO.: AD-A118501
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): A TWO-STAGE ELIMINATION TYPE PROCEDURE FOR SELECTING THE LARGEST OF SEVERAL NORMAL MEANS WITH A COMMON UNKNOWN VARIANCE	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: Technical Report #82-27
7. AUTHOR(s): Shanti S. Gupta and Woo-Chul Kim	8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, West Lafayette, IN 47907	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 042-243
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, D.C.	12. REPORT DATE: August 1982
	13. NUMBER OF PAGES: 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Selection procedures; common unknown variance; two-stage elimination type procedures; indifference-zone approach; subset selection; the expected total sample size; relative savings

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

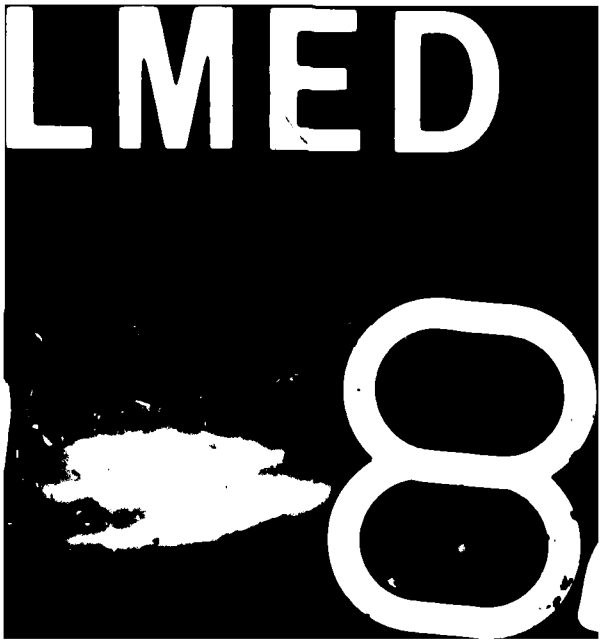
The problem of selecting the normal population with the largest mean from several populations with common unknown variance is considered. A new simple two-stage elimination type procedure is proposed which guarantees the same probability requirement using the indifference-zone approach as does the two-stage non-elimination type procedure of Bechhofer, Dunnett and Sobel (1954). The observations in the first stage are used not only to estimate the common unknown variance but also to screen out non-contending populations. The

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

proposed procedure has some similarities with the procedure of Tamhane (1975), but uses a new design criterion and a sharper lower bound on the probability of a correct selection. A table of constants necessary to implement the procedure is provided. The results of a Monte Carlo study of comparison of the elimination type procedure with the non-elimination type procedure are also given. The results strongly indicate that the proposed elimination type procedure performs much better than the non-elimination type procedure of Bechhofer, Dunnett and Sobel (1954) in terms of the expected total sample size.


LMED
8