SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

TECHNICAL REPORT NO. 894

April 1990

A COMPARISON OF THE PERFORMANCES
OF PROCEDURES FOR SELECTING THE
NORMAL POPULATION HAVING THE LARGEST
MEAN WHEN THE POPULATIONS HAVE A
COMMON UNKNOWN VARIANCE

by

Robert E. Bechhofer¹
Charles W. Dunnett²
David M. Goldsman³
Mark Hartmann⁴

Approved for public release; distribution unlimited

¹Department of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853. Research partially supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University and by U.S. Army Research Contract DAAL03-86-K-0046 at Cornell University.

²Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S-4K1. Research partially supported by the Natural Sciences and Engineering Research Council of Canada.

³School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

⁴Department of Operations Research, University of North Carolina, Chapel Hill, NC 27599. Research partially supported by National Science Foundation grant DMS-8905654.

TABLE OF CONTENTS

Abstract	1
1. Introduction and summary	1
2. Statement of the problem	2
3. Procedures considered	3
3.1 The Bechhofer-Dunnett-Sobel two-stage procedure	3
3.2 The Gupta-Kim elimination procedure	5
3.3 The Paulson fully sequential elimination procedure with the Hartmann improvement	7
4. The planning of the Monte Carlo sampling experiments	10
5. The performances of the procedures	11
6. Discussion of Monte Carlo sampling results	24
6.1 Discussion of B-D-S	24
6.2 Discussion of G-K	24
6.3 Discussion of P-H	25
6.4 Relative performances of the procedures	26
7. Concluding remarks	33
Acknowledgments	34
Bibliography	35

A COMPARISON OF THE PERFORMANCES OF PROCEDURES FOR SELECTING
THE NORMAL POPULATION HAVING THE LARGEST MEAN WHEN THE
POPULATIONS HAVE A COMMON UNKNOWN VARIANCE

Robert E. Bechhofer
School of Operations Research
and Industrial Engineering
Upson Hall
Cornell University
Ithaca, NY 14853

Charles W. Dunnett
Department of Mathematics
and Statistics
McMaster University
Hamilton, Ontario
Canada L8S-4K1

David M. Goldsman
School of Industrial
and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Mark Hartmann
Department of Operations Research
University of North Carolina
Chapel Hill, NC 27599

Key Words and Phrases: Normal means selection; indifference-zone approach; two-stage and multistage open procedures without and with elimination; achieved $P(CS)$; expected number of stages and expected total number of observations; Monte Carlo sampling.

ABSTRACT

We study the performances of the two-stage non-eliminating procedure of Bechhofer, Dunnett and Sobel (1954), the two-stage eliminating procedure of Gupta and Kim (1984) and the multi-stage eliminating procedure of Paulson (1964) as modified by Hartmann (1990) for selecting the normal population which has the largest mean when the common variance is unknown. Constants to implement the procedures are provided. All of these procedures are open and guarantee the indifference-zone probability requirement of Bechhofer (1954). The following performance characteristics are estimated by Monte Carlo sampling: the achieved probability of a correct selection, the expected number of vector-observations and the expected total number of observations to termination of experimentation. The critical role of the common initial sample size per population is discussed. Guidance is provided as to which procedure to use in different environments.

1. INTRODUCTION AND SUMMARY

In an earlier article (Bechhofer and Goldsman (1989)), the authors compared the performances of procedures for selecting the normal population having the largest mean when the variances are known and equal. Ten procedures (several of which were variants of the same procedures) were considered; these included a single-stage procedure,

two-stage and multi-stage closed procedures (with permanent elimination of populations indicated as being non-contending), as well as a fully sequential open procedure without elimination and a truncated version of the latter. All of these procedures guaranteed the same indifference-zone probability requirement of Bechhofer (1954). Each, at the time that it was proposed, had a different virtue and in some sense represented an improvement over the procedures that had been proposed earlier. Based on their observed performance characteristics, recommendations were made concerning which procedure to use in particular circumstances.

In the present article we carry out a similar study for normal populations which have a common unknown variance. Since the variance is unknown it is not possible, prior to the start of experimentation, to place an upper bound on the total number of observations required to terminate sampling and still satisfy the probability requirement; this fact had been established earlier by Dudewicz (1971) for selection problems using the indifference-zone approach. In particular, it is not possible to use truncated procedures with predetermined truncation numbers.

We consider here the two-stage non-eliminating procedure of Bechhofer, Dunnett and Sobel (1954), and a variant of this procedure by Gupta and Kim (1984) which after the first stage permits permanent elimination of populations indicated as being non-contending. We also study the performance of a fully sequential procedure of Paulson (1964) as improved by Hartmann (1990) which after the first stage permits permanent elimination of populations indicated as being non-contending. For all of these procedures the number of vector-observations (V) and the total number of observations (T) to terminate sampling are random variables. We shall be interested in the expected values and variances of these quantities as well as the associated achieved probabilities of a correct selection. As with our known variances study it is our objective to analyze the performances of these procedures, and to recommend which ones it is preferable to use in particular circumstances.

2. STATEMENT OF THE PROBLEM

We assume that we have $k \geq 2$ normal populations, $\Pi_1, \Pi_2, \dots, \Pi_k$ with unknown population means μ_i ($1 \leq i \leq k$) and a common unknown variance σ^2 . The ordered values of the μ_i are denoted by $\mu_{[1]} \leq \mu_{[2]} \leq \dots \leq \mu_{[k]}$. We further assume that the values of the $\mu_{[j]}$ ($1 \leq j \leq k$) are unknown and that the pairing of the $\mu_{[j]}$ with the Π_i ($1 \leq i, j \leq k$) is completely unknown. The objective of the experiment, and the associated probability requirement, are stated below in (2.1) and (2.2), respectively.

GOAL: To select the population associated with $\mu_{[k]}$. (2.1)

If the population selected by a procedure has the mean $\mu_{[k]}$, then we say that a correct selection (CS) has been made. We limit consideration to procedures that guarantee the following indifference-zone requirement on the $P\{CS\}$:

PROBABILITY REQUIREMENT:

$$P\{CS\} \geq P^* \text{ whenever } \mu_{[k]} - \mu_{[k-1]} \geq \delta^* \text{ for all } \sigma^2 > 0. \quad (2.2)$$

The quantities $\{\delta^*, P^*\}$ with $0 < \delta^* < \infty$, $1/k < P^* < 1$ are specified prior to the start of experimentation. All of the procedures under consideration in the present article guarantee (2.2). These procedures are described in Section 3.

3. PROCEDURES CONSIDERED

In order to make the present article self-contained, we describe and critique in this section the procedures under consideration. The reader is referred to the cited articles for more detailed information concerning these procedures and their performance characteristics.

3.1 The Bechhofer-Dunnett-Sobel two-stage procedure

This procedure (referred to herein as B-D-S) has the following form:

a) In the first stage an arbitrary common number $n_1 > 1$ of independent observations X_{ij} ($1 \leq i \leq k$, $1 \leq j \leq n_1$) is taken from Π_i ($1 \leq i \leq k$). The sample variance $s^2 = \sum_{i=1}^k \sum_{j=1}^{n_1} (x_{ij} - \bar{x}_i)^2 / \nu$ based on $\nu = k(n_1 - 1)$ d.f. is calculated. Here $\bar{x}_i = \sum_{j=1}^{n_1} x_{ij} / n_1$ ($1 \leq i \leq k$).

b) In the second stage the same number $N - n_1$ of additional independent observations X_{ij} is taken from every population Π_i ($1 \leq i \leq k$) where

$$N = \max\{n_1, [2(gs/\delta^*)^2]^+\}, \quad (3.1)$$

$[b]^+$ denotes the smallest integer equal to or greater than b , and $g = g(p, \nu, \alpha)$ with $p = k - 1$, and $\alpha = 1 - P^*$ is a constant tabulated in Table 3.1 for $1 - \alpha = 0.75$ and 0.90 . (A more extensive tabulation of these constants is given in Tables A.1–A.4 of Bechhofer and Dunnett (1988).) Next the k overall sample means $\bar{x}_i^* = \sum_{j=1}^N x_{ij} / N$ ($1 \leq i \leq k$) based on the combined first- and second-stage observations x_{ij} ($1 \leq i \leq k$, $1 \leq j \leq N$) are calculated, and it is asserted that the population associated with $\max_{1 \leq i \leq k} \bar{x}_i^*$ is the one having mean $\mu_{[k]}$.

TABLE 3.1

Constants $g = g(p, \nu, \alpha)$ to implement the Bechhofer-Dunnett-Sobel procedure for selected values of k , n_1 and α ($p = k-1$ and $\nu = k(n_1-1)$)

$$1-\alpha = 0.75$$

n_1	k							
	2	3	4	5	10	15	20	25
5	0.706	1.052	1.226	1.340	1.624	1.763	1.854	1.921
10	0.688	1.031	1.206	1.320	1.611	1.753	1.846	1.914
15	0.683	1.025	1.200	1.315	1.607	1.750	1.843	1.912
20	0.681	1.022	1.197	1.313	1.606	1.749	1.842	1.911
25	0.680	1.020	1.195	1.311	1.605	1.748	1.842	1.910
30	0.679	1.019	1.194	1.310	1.604	1.747	1.841	1.910

$$1-\alpha = 0.90$$

n_1	k							
	2	3	4	5	10	15	20	25
5	1.397	1.689	1.833	1.927	2.167	2.287	2.367	2.427
10	1.330	1.625	1.777	1.877	2.134	2.263	2.348	2.411
15	1.313	1.608	1.761	1.863	2.125	2.256	2.342	2.406
20	1.304	1.599	1.754	1.856	2.121	2.253	2.340	2.404
25	1.299	1.595	1.749	1.853	2.119	2.251	2.338	2.403
30	1.296	1.592	1.747	1.850	2.117	2.250	2.337	2.402

Remark 3.1: The entries in Tables A.1–A.4 of B-D (1988) in the columns for $\rho = 1/2$ give $g(p, \nu, \alpha)$ for $p = k-1 = 2(1)16(2)20$; $1-\alpha = P^* = 0.80, 0.90, 0.95, 0.99$; and $\nu = 2(1)30(5)50, 60(20)100, 200, \infty$. The B-D-S procedure is also described in Section 6.3.1 of B-D (1988). Gupta, Panchapakesan and Sohn (1985) give extensive although somewhat less accurate tables of $g(p, \nu, \alpha)$, including values for $1-\alpha = P^* = 0.75$.

Remark 3.2: Gupta and Kim (1984) report Monte Carlo (MC) simulation results obtained with B-D-S in their Table 4.1, pp. 86 and 87, under the heading $\mathcal{P}(\text{BDS})$. However, we were not able to verify their results. (See our Table 5.4a for $k = 5$, $P^* = 0.90$). We are convinced that all of the MC results reported in their paper for $\mathcal{P}(\text{BDS})$ are erroneous.

The B-D-S procedure cannot capitalize on favorable configurations of the population means. That is, the distribution of the number of vector-observations $(X_{1j}, \dots, X_{kj})$ ($j = 1, 2, \dots$) to terminate sampling is the same for all configurations of the population means even if, e.g., $\mu_{[k]} - \mu_{[k-1]} \gg \sigma$. Also, the same number of observations is taken from all populations even if, after the first stage, one or more populations have sample means which differ by many standard errors from the largest sample mean (and thus are indicated as being non-contending for selection). Reacting to these drawbacks of the B-D-S procedure, Gupta and Kim (1984) adopted the strategy of Tamhane and Bechhofer (1977, 1979) and proposed a procedure which selects a subset of the populations in the first stage, retaining in this subset for second stage experimentation only those populations indicated as being in contention for selection. The intent of this strategy is to reduce the expected total number of observations required in the experiment. Their procedure is described in the next subsection.

3.2 The Gupta-Kim elimination procedure

This procedure (referred to herein as G-K) has the following form:

a) In the first stage an arbitrary common number $n_1 > 1$ of independent observations X_{ij} ($1 \leq i \leq k, 1 \leq j \leq n_1$) is taken from Π_i ($1 \leq i \leq k$). The k observed first-stage sample means $\bar{x}_i^{(1)} = \sum_{j=1}^{n_1} x_{ij}/n_1$ ($1 \leq i \leq k$) and sample variance $s^2 = \sum_{i=1}^k \sum_{j=1}^{n_1} (x_{ij} - \bar{x}_i^{(1)})^2/\nu$ based on $\nu = k(n_1 - 1)$ d.f. are calculated. A subset I of $\{1, 2, \dots, k\}$ is determined by

$$I = \{i \mid \bar{x}_i^{(1)} \geq \max_{1 \leq j \leq k} \bar{x}_j^{(1)} - (\sqrt{2} ds/\sqrt{n_1} - \delta^*)^\# \}, \quad (3.2)$$

where $b^\# = \max(b, 0)$, and $d = d(p, \nu, \alpha)$ is a constant described in Remark 3.3, below. If $|I| = 1$ then sampling terminates and it is asserted that the population associated with $\max_{1 \leq i \leq k} \bar{x}_i^{(1)}$ is the one having mean $\mu_{[k]}$; if $|I| > 1$ then a second stage of sampling is required.

b) In the second stage the same number $N - n_1$ of additional independent observations X_{ij} is taken from each Π_i for which $i \in I$, where N is defined in (3.1) but with g taken from Table 3.2. Next the overall sample means $\bar{x}_i = \sum_{j=1}^N x_{ij}/N$ for all $i \in I$ based on the combined first and second-stage observations x_{ij} for $i \in I, 1 \leq j \leq N$ are calculated, and it is asserted that the population associated with $\max_{i \in I} \bar{x}_i$ is the one having mean $\mu_{[k]}$.

Remark 3.3: In its most general form the G-K procedure requires the specification of two constants, $g = g(p, \nu, \alpha)$ of Remark 3.1 and $d = d(p, \nu, \alpha)$ employed in (3.2). The

TABLE 3.2

Constants $g = g(p, \nu, \alpha)$ to implement the Gupta-Kim procedure
for selected values of k , n_1 and α ($p = k-1$ and $\nu = k(n_1-1)$)

$$1-\alpha = \sqrt{0.75}$$

n_1	k							
	2	3	4	5	10	15	20	25
5	1.191	1.500	1.654	1.754	2.008	2.135	2.218	2.280
10	1.143	1.452	1.611	1.715	1.983	2.116	2.203	2.268
15	1.130	1.438	1.599	1.704	1.976	2.110	2.198	2.264
20	1.124	1.433	1.593	1.699	1.972	2.108	2.196	2.262
25	1.121	1.429	1.590	1.696	1.970	2.106	2.195	2.261
30	1.119	1.426	1.587	1.694	1.969	2.105	2.194	2.261

$$1-\alpha = \sqrt{0.90}$$

n_1	k							
	2	3	4	5	10	15	20	25
5	1.843	2.093	2.213	2.291	2.494	2.600	2.672	2.727
10	1.720	1.984	2.120	2.210	2.444	2.564	2.643	2.703
15	1.687	1.955	2.094	2.188	2.431	2.553	2.635	2.696
20	1.672	1.942	2.083	2.178	2.425	2.549	2.631	2.693
25	1.664	1.934	2.076	2.172	2.420	2.546	2.629	2.691
30	1.658	1.928	2.072	2.168	2.418	2.544	2.628	2.690

constant d controls the probability that the subset selected in the first stage will contain the population associated with $\mu_{[k]}$; the constant g controls the probability that the population ultimately selected after the first or second stage is indeed the one associated with $\mu_{[k]}$. It is convenient (although not optimal) to set $g = d$ in which case a conservative g -value is obtained, which for given (p, ν) guarantees a $P\{CS\}$ of at least P^* . This was done in Table 3.3 of G-K (1984) the entries of which are the solution of their equation (3.3) and yield their procedure $\mathcal{P}_3$. (These new g -values have the property that they would guarantee $\sqrt{P^*} = \sqrt{1-\alpha}$ in (2.2) if used in (3.1) for the B-D-S procedure; they guarantee at least P^* in (2.2) if used in (3.1) and (3.2) with $d = g$ for the G-K procedure.) Selected new g -values are given in our Table 3.2 for $1-\alpha = \sqrt{0.75}$ and $\sqrt{0.90}$; here (following G-K (1984)) $g = g(p, \nu, \alpha)$ is tabled as a function of k and n_1 where $k = p+1$ and $\nu = k(n_1-1)$. Table 3.2 can be regarded

as a supplement to Tables A.1, A.2 and A.3 given in Bechhofer and Dunnett (1988) for $1-\alpha = 0.80, 0.90$ and 0.95 , respectively, with $\rho = 1/2$.

A drawback of the G-K procedure is that populations indicated as being non-contending for selection can be permanently eliminated from sampling only after the first stage of sampling with no further elimination being possible since the second stage is the last one. The main purpose of the first stage is to provide an estimate of σ^2 . With this estimate of σ^2 in hand it clearly would be desirable to employ a multi-stage procedure which permits permanent elimination of populations indicated as being non-contending, not only after the first stage but also after each subsequent stage where such apparently non-contending populations can be identified.

Paulson (1964), Section 5, proposed such a procedure. It takes observations one vector-at-a-time after the first stage which mainly serves to provide an estimate of σ^2 ; however, populations can be eliminated at every stage after that first stage. These subsequent vectors, each of which constitutes a stage, contain one observation from all populations not eliminated at an earlier stage. Thus Paulson's procedure has the highly desirable feature of reacting to favorable configurations of the population means by eliminating populations indicated as not being in contention for selection, and then focusing sampling on populations still in contention. We now describe that procedure as improved by Hartmann (1989).

3.3 The Paulson fully sequential elimination procedure with the Hartmann improvement

This procedure (referred to herein as P-H) is actually a family of procedures which depend on a quantity $\lambda = \delta^*/2t$ where t is a positive integer (usually 1 or 2) specified prior to the start of experimentation. The procedure has the following form:

a) In the first stage an arbitrary common number $n_1 > 1$ of independent observations X_{ij} ($1 \leq i \leq k, 1 \leq j \leq n_1$) is taken from Π_i ($1 \leq i \leq k$). The k observed first-stage sample means $\bar{x}_i^{(1)} = \sum_{j=1}^{n_1} x_{ij}/n_1$ ($1 \leq i \leq k$) and sample variance $s^2 = \sum_{i=1}^k \sum_{j=1}^{n_1} (x_{ij} - \bar{x}_i^{(1)})^2/\nu$ based on $\nu = k(n_1 - 1)$ d.f. are calculated. For $\lambda = \delta^*/2$ (where $t = 1$) define

$$\gamma = (2[1 - (P^*)^{1/(k-1)}])^{-2/\nu} - 1, \quad (3.3a)$$

and for $\lambda = \delta^*/4$ (where $t = 2$) define γ as the solution of

$$(1+3\gamma)^{-\nu/2} - (1/2)(1+4\gamma)^{-\nu/2} = 1 - (P^*)^{1/(k-1)}. \quad (3.3b)$$

Next let $a_\lambda^* = t\gamma\nu s^2/\delta^*$ and let W_λ^* denote the largest integer less than a_λ^*/λ . Tables 3.3 and 3.4 give values of the product $\gamma\nu$ for $\lambda = \delta^*/4$ and $\delta^*/2$, respectively, for selected k , $\nu = k(n_1-1)$ and P^* . If $n_1 > W_\lambda^*$ the experiment is terminated and the population associated with $\max_{1 \leq i \leq k} \sum_{j=1}^{n_1} x_{ij}$ is selected as the population associated with $\mu_{[k]}$. If $n_1 \leq W_\lambda^*$ then any population Π_j for which

$$\sum_{s=1}^{n_1} x_{js} < \max_{1 \leq i \leq k} \sum_{s=1}^{n_1} x_{is} - a_\lambda^* + n_1\lambda, \quad (3.4)$$

is permanently eliminated, and sampling proceeds to the next stage.

TABLE 3.3

Constants $\gamma\nu$ to implement the Paulson-Hartmann procedure for $\lambda = \delta^*/4$ and selected values of k , n_1 and P^* ($\nu = k(n_1-1)$)

$P^* = 0.75$

n_1	k							
	2	3	4	5	10	15	20	25
5	0.672	1.249	1.571	1.788	2.352	2.642	2.838	2.987
10	0.642	1.172	1.476	1.686	2.257	2.558	2.764	2.921
15	0.634	1.151	1.451	1.659	2.230	2.534	2.743	2.902
20	0.630	1.142	1.438	1.645	2.218	2.524	2.733	2.893
25	0.628	1.136	1.431	1.638	2.211	2.518	2.727	2.888
30	0.626	1.132	1.426	1.633	2.206	2.513	2.724	2.885

$P^* = 0.90$

n_1	k							
	2	3	4	5	10	15	20	25
5	1.656	2.250	2.541	2.726	3.197	3.444	3.616	3.750
10	1.469	2.004	2.294	2.491	3.019	3.300	3.495	3.644
15	1.421	1.940	2.229	2.428	2.970	3.261	3.461	3.614
20	1.398	1.910	2.199	2.399	2.948	3.242	3.445	3.601
25	1.385	1.893	2.181	2.382	2.935	3.231	3.436	3.592
30	1.377	1.882	2.170	2.371	2.925	3.224	3.429	3.587

TABLE 3.4

Constants $\gamma\nu$ to implement the Paulson-Hartmann procedure
for $\lambda = \delta^*/2$ and selected values of k , n_1 and P^* ($\nu = k(n_1-1)$)

$P^* = 0.75$

n_1	k							
	2	3	4	5	10	15	20	25
5	1.514	2.945	3.786	4.367	5.933	6.758	7.326	7.760
10	1.441	2.767	3.563	4.128	5.705	6.558	7.148	7.599
15	1.421	2.718	3.503	4.063	5.643	6.503	7.098	7.554
20	1.412	2.696	3.475	4.033	5.613	6.477	7.075	7.533
25	1.407	2.683	3.459	4.015	5.596	6.461	7.061	7.520
30	1.403	2.674	3.448	4.004	5.585	6.452	7.052	7.512

$P^* = 0.90$

n_1	k							
	2	3	4	5	10	15	20	25
5	3.963	5.537	6.348	6.880	8.274	9.017	9.536	9.938
10	3.525	4.960	5.764	6.319	7.844	8.667	9.238	9.677
15	3.411	4.809	5.610	6.170	7.726	8.570	9.155	9.604
20	3.359	4.740	5.539	6.101	7.671	8.525	9.116	9.570
25	3.329	4.700	5.498	6.061	7.640	8.499	9.094	9.550
30	3.310	4.674	5.472	6.036	7.619	8.482	9.079	9.537

b) Continuing for $r = n_1+1, \dots, W_\lambda^*$, the procedure takes observations $\{X_{ir}\}$ a vector-at-a-time from populations Π_i which have not already been eliminated, and eliminates any remaining population Π_j for which

$$\sum_{s=1}^r x_{js} < \max_i \sum_{s=1}^r x_{is} - a_\lambda^* + r\lambda,$$

where the maximum is taken over the remaining populations. If at some stage only one population remains, then that population is selected as the one associated with $\mu_{[k]}$. If two or more populations remain after stage W_λ^* , then one more observation is taken from each of these populations, and the one with the largest cumulative sum is selected as the population associated with $\mu_{[k]}$.

Remark 3.4. The Hartmann (1990) improvement consists of replacing Paulson's a^* by a smaller number a_λ^* which has the effect of decreasing the expected number of stages and expected total number of observations to termination uniformly in the μ_i ($1 \leq i \leq k$) and σ^2 .

Remark 3.5: Conditional on the value of s^2 calculated from the first-stage observations, the choice $\lambda = \delta^*/2$ minimizes the maximum number of stages to termination of sampling for P-H. Paulson (1964) recommended the choice $\lambda = \delta^*/4$ for the common known variance case. In B-G (1989), where P-H for this case (as described in Hartmann (1988)) was one of the procedures studied, MC sampling results for $\lambda = \delta^*/4$ and $\delta^*/2$ were reported.

4. THE PLANNING OF THE MC SAMPLING EXPERIMENTS

Although only three basically different procedures were to be compared, namely, B-D-S (1954), G-K (1984) and H (1990); the latter two are each actually families of procedures. For the reasons given in Remarks 3.3 and 3.5, only one procedure in the G-K family and two procedures in the P-H family were studied. Thus we focused on four procedures.

The experiments necessary to evaluate the relative performances of these selection procedures for the common unknown variance case presented design issues that we had not encountered for the common known variance case as reported in B-G (1989). There are many factors that affect the performances of all of these procedures for the unknown variance case. The new underlying factor is the value of the unknown variance (which plays its role in terms of the ratio σ/δ^* where $\delta^* > 0$ is specified). The new design factor is the common number of observations (n_1) per population taken in the first stage, mainly to estimate the value of σ^2 ; these same observations also contribute to the estimates of the population means. Additional factors are the number of populations (k), the specified value of P^* , and the spacing of the unknown μ_i ($1 \leq i \leq k$).

It was anticipated that certain procedures would outperform others in different environments. Thus we varied the levels of the several factors of the problem over reasonable ranges in order to study the effects of changes in the levels of these factors on the performances of the procedures. For each procedure the number of populations studied was $k = 3, 5, 10$; the specifications considered were $P^* = 0.75, 0.90$ along with ratios $\sigma/\delta^* = 2, 3$. The first-stage sample sizes were set at $n_1 = 5, 10, 15, 20$. Two configurations of the population means were studied, namely the so-called least-favorable (LF), i.e., $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ and equal means (EM), i.e., $\mu_{[1]} = \mu_{[k]}$. Thus

for each of the procedures we conducted a large factorial experiment. Each experiment involved the five factors at levels 3, 2, 2, 4, 2, respectively, yielding 96 factor-level combinations. We would have liked to include additional levels for certain of the factors but felt that the incremental knowledge which might be gained could not be justified in terms of the computing time that would be required.

Remark 4.1: We actually did conduct certain supplementary incomplete factorial experiments for additional values of the factors and for other configurations, namely, $k = 2$ and 4 , $\sigma/\delta^* = 1$, $P^* = 0.95$, $n_1 = 25, 30$, and some experiments with the population means equally-spaced δ^* apart. However, these results are not reported herein. The supplementary experiments helped us indicate the extent to which our findings can be extrapolated to additional $(k, \sigma/\delta^*, P^*, n_1)$ -values and other configurations of the population means.

Each of the 24 simulation runs for $P^* = 0.90$, $\sigma/\delta^* = 3$ is based on 40,000 independent replications; all other runs are based on at least 10,000 independent replications. Based on preliminary estimates of the various variances, the number of replications used was deemed sufficient to pin down the quantities being estimated. Every replication was independent of all other replications. For each of the 48 LF experiments the primary results recorded were the observed proportion of correct selections ($\bar{p}$), the observed average number of vector-observations ($\bar{v}$) to termination of sampling and the observed average total number of observations ($\bar{t}$) to termination of sampling, each with its estimated standard error. For each of the 48 EM experiments we recorded only $\bar{v}$ and $\bar{t}$ along with their standard errors.

5. THE PERFORMANCES OF THE PROCEDURES

Tables 5.1, 5.2, 5.3, 5.4, 5.5, and 5.6 give the MC sampling results for $(k, P^*) = (3, 0.75), (3, 0.90), (5, 0.75), (5, 0.90), (10, 0.75)$ and $(10, 0.90)$, respectively; LF (EM) results are given for $(3, 0.75)$ in 5.1a (5.1b). The same pattern was followed for each (k, P^*) combination. For the B-D-S procedure, the same $\bar{v}$ - and $\bar{t}$ -results for the LF-configuration are repeated for the EM-configuration, i.e., separate simulations were not conducted since $\bar{v}$ - and $\bar{t}$ -results are independent of the configuration.

In the first column of each of the six tables are two numbers; the upper number is the value of σ/δ^* while the lower one in parentheses (n_{SS}) is the smallest common

TABLE 5.1A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 3$, $P^* = 0.75$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (9)	5	0.7654 (0.0042)	0.8232 (0.0038)	0.8469 (0.0036)	0.8261 (0.0038)	9.4 (<0.1)	13.1 (0.1)	10.6 (0.1)	10.0 (0.1)	28.3 (0.1)	34.5 (0.2)	28.2 (0.2)	27.3 (0.1)
	10	0.7934 (0.0040)	0.8431 (0.0036)	0.8659 (0.0034)	0.8525 (0.0035)	10.5 (<0.1)	12.2 (<0.1)	12.4 (<0.1)	11.7 (<0.1)	31.6 (<0.1)	34.8 (0.1)	35.3 (0.1)	33.8 (0.1)
	15	0.8551 (0.0035)	0.8703 (0.0034)	0.9011 (0.0030)	0.8817 (0.0032)	15.0 (<0.1)	15.3 (<0.1)	16.1 (<0.1)	15.4 (<0.1)	45.0 (<0.1)	45.6 (<0.1)	47.2 (0.1)	45.9 (<0.1)
	20	0.9006 (0.0030)	0.8997 (0.0030)	0.9243 (0.0026)	0.9115 (0.0028)	20.0 (<0.1)	20.0 (<0.1)	20.4 (<0.1)	20.1 (<0.1)	60.0 (<0.1)	60.0 (<0.1)	60.8 (<0.1)	60.1 (<0.1)
3 (19)	5	0.7499 (0.0043)	0.7963 (0.0040)	0.8184 (0.0039)	0.8137 (0.0039)	20.3 (0.1)	33.0 (0.2)	21.2 (0.1)	20.4 (0.1)	61.0 (0.2)	86.3 (0.6)	55.2 (0.4)	53.9 (0.3)
	10	0.7518 (0.0043)	0.8044 (0.0040)	0.8334 (0.0037)	0.8159 (0.0039)	19.7 (0.1)	26.5 (0.2)	20.3 (0.1)	19.5 (0.1)	59.0 (0.2)	68.9 (0.4)	54.1 (0.3)	52.7 (0.2)
	15	0.7600 (0.0043)	0.8148 (0.0039)	0.8437 (0.0036)	0.8261 (0.0038)	19.6 (<0.1)	24.9 (0.1)	22.0 (0.1)	20.8 (0.1)	58.9 (0.1)	67.1 (0.3)	60.7 (0.2)	58.2 (0.2)
	20	0.7779 (0.0042)	0.8262 (0.0038)	0.8543 (0.0035)	0.8328 (0.0037)	21.1 (<0.1)	25.1 (0.1)	24.9 (0.1)	23.5 (0.1)	63.2 (0.1)	71.1 (0.2)	70.8 (0.2)	67.5 (0.1)

TABLE 5.1B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 3$, $P^* = 0.75$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (9)	5					9.4 (<0.1)	14.4 (0.1)	13.7 (0.1)	12.1 (0.1)	28.3 (0.1)	38.1 (0.3)	35.9 (0.2)	32.5 (0.2)
	10					10.5 (<0.1)	13.2 (0.1)	14.5 (0.1)	12.9 (<0.1)	31.6 (<0.1)	37.3 (0.1)	40.2 (0.1)	36.7 (0.1)
	15					15.0 (<0.1)	15.7 (<0.1)	17.4 (<0.1)	16.0 (<0.1)	45.0 (<0.1)	46.4 (<0.1)	50.2 (0.1)	47.2 (<0.1)
	20					20.0 (<0.1)	20.0 (<0.1)	21.3 (<0.1)	20.2 (<0.1)	60.0 (<0.1)	60.1 (<0.1)	62.8 (0.1)	60.5 (<0.1)
3 (19)	5					20.3 (0.1)	34.5 (0.2)	27.4 (0.2)	25.0 (0.2)	61.0 (0.2)	91.0 (0.6)	70.4 (0.5)	65.8 (0.4)
	10					19.7 (0.1)	29.3 (0.2)	25.2 (0.1)	23.2 (0.1)	59.0 (0.2)	76.5 (0.4)	65.8 (0.3)	61.8 (0.3)
	15					19.6 (<0.1)	27.5 (0.1)	26.4 (0.1)	23.9 (0.1)	58.9 (0.1)	74.1 (0.3)	71.1 (0.3)	65.6 (0.2)
	20					21.1 (<0.1)	27.7 (0.1)	28.5 (0.1)	25.8 (0.1)	63.2 (0.1)	77.1 (0.2)	78.9 (0.2)	73.0 (0.2)

TABLE 5.2A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 3$, $P^* = 0.90$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (20)	5	0.9077 (0.0029)	0.9212 (0.0027)	0.9365 (0.0024)	0.9321 (0.0025)	23.3 (0.1)	30.5 (0.2)	18.4 (0.1)	18.2 (0.1)	69.9 (0.3)	82.3 (0.5)	49.0 (0.3)	49.5 (0.3)
	10	0.9027 (0.0030)	0.9187 (0.0027)	0.9384 (0.0024)	0.9366 (0.0024)	21.6 (0.1)	22.5 (0.1)	17.2 (0.1)	17.0 (0.1)	64.8 (0.2)	59.9 (0.3)	47.0 (0.2)	46.8 (0.2)
	15	0.9071 (0.0029)	0.9240 (0.0027)	0.9439 (0.0023)	0.9399 (0.0024)	21.3 (<0.1)	20.2 (0.1)	19.1 (0.1)	18.3 (<0.1)	64.0 (0.1)	56.7 (0.2)	54.0 (0.1)	52.5 (0.1)
	20	0.9146 (0.0028)	0.9340 (0.0025)	0.9574 (0.0020)	0.9442 (0.0023)	22.0 (<0.1)	22.0 (0.1)	22.1 (<0.1)	21.5 (<0.1)	66.0 (0.1)	64.3 (0.1)	64.5 (0.1)	63.2 (0.1)
3 (45)	5	0.8969 (0.0030)	0.9119 (0.0028)	0.9248 (0.0013)	0.9212 (0.0013)	52.0 (0.2)	74.4 (0.4)	39.5 (0.1)	39.3 (0.1)	155.9 (0.6)	206.8 (1.2)	104.5 (0.3)	106.1 (0.3)
	10	0.8994 (0.0030)	0.9153 (0.0028)	0.9266 (0.0013)	0.9225 (0.0013)	48.0 (0.1)	59.8 (0.3)	35.5 (0.1)	35.3 (0.1)	143.9 (0.4)	158.7 (0.9)	93.5 (0.2)	95.0 (0.2)
	15	0.9023 (0.0030)	0.9178 (0.0027)	0.9294 (0.0013)	0.9243 (0.0013)	46.9 (0.1)	52.6 (0.3)	34.8 (0.1)	34.7 (0.1)	140.7 (0.3)	137.6 (0.8)	92.7 (0.2)	93.8 (0.2)
	20	0.9053 (0.0029)	0.9222 (0.0027)	0.9325 (0.0013)	0.9244 (0.0013)	46.6 (0.1)	47.9 (0.3)	35.4 (0.1)	34.8 (0.1)	139.8 (0.3)	126.4 (0.7)	95.8 (0.2)	95.3 (0.2)

TABLE 5.2B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when

$k = 3$, $P^* = 0.90$, $\sigma/\delta^* = 2.3$; $n = 5(5)20$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (20)	5					23.3 (0.1)	32.3 (0.2)	29.4 (0.2)	25.7 (0.2)	69.9 (0.3)	89.1 (0.5)	77.0 (0.5)	69.0 (0.4)
	10					21.6 (0.1)	26.9 (0.1)	25.4 (0.1)	22.5 (0.1)	64.8 (0.2)	72.3 (0.4)	67.1 (0.3)	61.0 (0.3)
	15					21.3 (<0.1)	24.9 (0.1)	26.0 (0.1)	22.8 (0.1)	64.0 (0.1)	68.4 (0.3)	70.5 (0.3)	63.4 (0.2)
	20					22.0 (<0.1)	24.7 (0.1)	27.4 (0.1)	24.5 (0.1)	66.0 (0.1)	70.6 (0.2)	76.7 (0.2)	70.2 (0.1)
3 (45)	5					52.0 (0.2)	75.8 (0.4)	63.2 (0.2)	55.8 (0.2)	155.9 (0.6)	213.2 (1.2)	164.6 (0.6)	149.7 (0.5)
	10					48.0 (0.1)	64.2 (0.3)	52.6 (0.2)	48.1 (0.1)	143.9 (0.4)	174.5 (0.8)	136.4 (0.4)	128.2 (0.3)
	15					46.9 (0.1)	60.3 (0.3)	50.4 (0.1)	46.1 (0.1)	140.7 (0.3)	161.0 (0.7)	131.2 (0.3)	123.1 (0.3)
	20					46.6 (0.1)	57.4 (0.2)	50.0 (0.1)	45.5 (0.1)	139.8 (0.3)	153.3 (0.7)	131.4 (0.3)	122.4 (0.2)

TABLE 5.3A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 5$, $P^* = 0.75$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (14)	5	0.7632 (0.0043)	0.7978 (0.0040)	0.8600 (0.0035)	0.8490 (0.0036)	14.9 (<0.1)	21.7 (0.1)	16.9 (0.1)	16.3 (0.1)	74.3 (0.2)	80.4 (0.5)	63.5 (0.3)	63.7 (0.3)
	10	0.7615 (0.0043)	0.8158 (0.0039)	0.8663 (0.0034)	0.8446 (0.0036)	14.5 (<0.1)	17.6 (0.1)	17.1 (0.1)	16.0 (0.1)	72.4 (0.1)	70.6 (0.3)	69.4 (0.2)	67.1 (0.2)
	15	0.7775 (0.0042)	0.8375 (0.0037)	0.8861 (0.0032)	0.8620 (0.0034)	15.6 (<0.1)	17.6 (<0.1)	19.0 (0.1)	17.9 (<0.1)	78.2 (0.1)	81.0 (0.1)	84.7 (0.1)	82.0 (0.1)
	20	0.8381 (0.0037)	0.8585 (0.0035)	0.9087 (0.0029)	0.8836 (0.0032)	20.0 (<0.1)	20.5 (<0.1)	22.1 (<0.1)	21.1 (<0.1)	100.0 (<0.1)	101.1 (<0.1)	104.7 (0.1)	102.4 (0.1)
3 (31)	5	0.7524 (0.0043)	0.7853 (0.0041)	0.8437 (0.0036)	0.8375 (0.0037)	32.7 (0.1)	52.1 (0.2)	36.1 (0.2)	34.8 (0.2)	163.4 (0.5)	202.3 (1.1)	132.6 (0.7)	134.2 (0.6)
	10	0.7494 (0.0043)	0.7926 (0.0041)	0.8460 (0.0036)	0.8305 (0.0038)	31.9 (0.1)	46.3 (0.2)	34.5 (0.2)	33.1 (0.1)	159.3 (0.3)	168.3 (0.8)	126.8 (0.5)	127.3 (0.5)
	15	0.7597 (0.0043)	0.8059 (0.0040)	0.8458 (0.0018)	0.8363 (0.0019)	31.6 (0.1)	41.5 (0.2)	34.4 (0.1)	33.0 (0.1)	158.2 (0.3)	152.5 (0.7)	131.0 (0.2)	129.9 (0.2)
	20	0.7532 (0.0043)	0.8100 (0.0039)	0.8516 (0.0018)	0.8371 (0.0018)	31.5 (<0.1)	38.6 (0.2)	35.2 (0.1)	33.5 (0.1)	157.6 (0.2)	150.2 (0.5)	141.0 (0.2)	137.9 (0.2)

TABLE 5.3B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when

$$k = 5, P^* = 0.75, \sigma/\delta^* = 2, 3; n = 5(5)20$$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (14)	5					14.9 (<0.1)	22.8 (0.1)	23.6 (0.1)	20.7 (0.1)	74.3 (0.2)	86.8 (0.5)	85.1 (0.5)	79.2 (0.4)
	10					14.5 (<0.1)	20.1 (0.1)	22.2 (0.1)	19.7 (0.1)	72.4 (0.1)	79.2 (0.3)	84.6 (0.3)	78.9 (0.2)
	15					15.6 (<0.1)	19.4 (0.1)	23.3 (0.1)	20.6 (0.1)	78.2 (0.1)	86.0 (0.2)	96.6 (0.2)	89.8 (0.2)
	20					20.0 (<0.1)	21.2 (<0.1)	25.7 (0.1)	22.8 (<0.1)	100.0 (<0.1)	102.6 (0.1)	114.0 (0.2)	106.7 (0.1)
3 (31)	5					32.7 (0.1)	53.0 (0.2)	49.7 (0.3)	44.4 (0.2)	163.4 (0.5)	209.8 (1.1)	176.6 (1.0)	168.7 (0.8)
	10					31.9 (0.1)	48.5 (0.2)	45.3 (0.2)	40.7 (0.2)	159.3 (0.3)	181.6 (0.8)	160.5 (0.7)	154.2 (0.6)
	15					31.6 (0.1)	45.8 (0.2)	44.5 (0.2)	40.2 (0.1)	158.2 (0.3)	171.6 (0.7)	162.6 (0.6)	155.1 (0.5)
	20					31.5 (<0.1)	43.8 (0.2)	44.9 (0.2)	40.4 (0.1)	157.6 (0.2)	168.5 (0.6)	169.9 (0.5)	160.2 (0.4)

TABLE 5.4A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (27)	5	0.9059 (0.0029)	0.9120 (0.0028)	0.9402 (0.0012)	0.9355 (0.0012)	30.3 (0.1)	40.2 (0.2)	25.4 (0.1)	25.0 (0.1)	151.7 (0.5)	165.8 (0.9)	97.6 (0.2)	100.4 (0.2)
	10	0.9025 (0.0030)	0.9148 (0.0028)	0.9418 (0.0023)	0.9326 (0.0025)	28.7 (0.1)	33.1 (0.1)	23.7 (0.1)	23.3 (0.1)	143.6 (0.3)	125.4 (0.6)	92.7 (0.3)	94.1 (0.3)
	15	0.9024 (0.0030)	0.9252 (0.0026)	0.9437 (0.0023)	0.9387 (0.0024)	28.3 (<0.1)	28.5 (0.1)	24.1 (0.1)	23.3 (0.1)	141.6 (0.2)	112.9 (0.4)	100.1 (0.3)	98.9 (0.2)
	20	0.8954 (0.0031)	0.9211 (0.0027)	0.9514 (0.0022)	0.9396 (0.0024)	28.1 (<0.1)	26.9 (0.1)	25.7 (0.1)	24.8 (0.1)	140.6 (0.2)	117.2 (0.3)	114.5 (0.2)	112.3 (0.2)
3 (61)	5	0.8974 (0.0030)	0.9093 (0.0029)	0.9346 (0.0012)	0.9329 (0.0013)	67.3 (0.2)	93.2 (0.3)	55.2 (0.1)	54.9 (0.1)	336.4 (1.1)	411.2 (1.8)	211.0 (0.5)	220.0 (0.5)
	10	0.9092 (0.0029)	0.9121 (0.0028)	0.9337 (0.0012)	0.9315 (0.0013)	64.0 (0.1)	83.6 (0.3)	51.1 (0.1)	50.8 (0.1)	320.2 (0.7)	338.1 (1.4)	193.0 (0.4)	201.5 (0.4)
	15	0.8958 (0.0031)	0.9131 (0.0028)	0.9352 (0.0012)	0.9311 (0.0013)	63.0 (0.1)	78.3 (0.3)	50.0 (0.1)	49.9 (0.1)	315.0 (0.5)	300.2 (1.3)	189.5 (0.4)	197.6 (0.3)
	20	0.9022 (0.0030)	0.9167 (0.0028)	0.9313 (0.0013)	0.9291 (0.0013)	62.7 (0.1)	73.1 (0.3)	49.7 (0.1)	49.6 (0.1)	313.3 (0.4)	273.8 (1.2)	191.4 (0.3)	198.0 (0.3)

TABLE 5.4B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
$^2_{(27)}$	5					30.3 (0.1)	41.3 (0.2)	41.5 (0.2)	35.6 (0.2)	151.7 (0.5)	178.0 (0.8)	152.3 (0.8)	140.4 (0.6)
	10					28.7 (0.1)	37.1 (0.1)	36.5 (0.2)	32.1 (0.1)	143.6 (0.3)	148.8 (0.6)	133.9 (0.5)	126.4 (0.4)
	15					28.3 (<0.1)	34.7 (0.1)	35.5 (0.1)	31.5 (0.1)	141.6 (0.2)	138.1 (0.5)	135.3 (0.4)	127.0 (0.3)
	20					28.1 (<0.1)	33.1 (0.1)	36.0 (0.1)	31.7 (0.1)	140.6 (0.2)	137.6 (0.3)	144.2 (0.4)	133.6 (0.3)
$^3_{(61)}$	5					67.3 (0.2)	93.2 (0.3)	89.7 (0.2)	78.7 (0.2)	336.4 (1.1)	420.2 (1.8)	327.6 (0.9)	309.2 (0.7)
	10					64.0 (0.1)	86.1 (0.2)	78.3 (0.2)	70.1 (0.1)	320.2 (0.7)	363.5 (1.3)	283.5 (0.6)	272.7 (0.5)
	15					63.0 (0.1)	83.1 (0.2)	75.3 (0.2)	68.0 (0.1)	315.0 (0.5)	337.2 (1.2)	272.8 (0.5)	264.2 (0.4)
	20					62.7 (0.1)	80.4 (0.2)	74.1 (0.1)	66.9 (0.1)	313.3 (0.4)	319.3 (1.1)	270.2 (0.5)	260.7 (0.4)

TABLE 5.5A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 10$, $P^* = 0.75$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (21)	5	0.7563 (0.0043)	0.7857 (0.0041)	0.8766 (0.0033)	0.8684 (0.0034)	21.7 (<0.1)	31.2 (0.1)	25.6 (0.1)	24.4 (0.1)	216.6 (0.5)	201.0 (0.9)	158.0 (0.6)	164.3 (0.6)
	10	0.7574 (0.0043)	0.8095 (0.0039)	0.8753 (0.0033)	0.8703 (0.0034)	21.3 (<0.1)	27.6 (0.1)	24.8 (0.1)	23.8 (0.1)	212.6 (0.3)	171.1 (0.6)	161.0 (0.4)	164.2 (0.4)
	15	0.7577 (0.0043)	0.8154 (0.0039)	0.8874 (0.0016)	0.8678 (0.0017)	21.1 (<0.1)	24.7 (0.1)	25.3 (<0.1)	24.0 (<0.1)	211.5 (0.2)	180.3 (0.3)	184.7 (0.2)	182.5 (0.1)
	20	0.7636 (0.0042)	0.8292 (0.0038)	0.9041 (0.0015)	0.8757 (0.0017)	21.5 (<0.1)	24.2 (0.1)	26.8 (<0.1)	25.2 (<0.1)	215.2 (0.2)	211.0 (0.2)	219.8 (0.1)	215.6 (0.1)
3 (47)	5	0.7577 (0.0043)	0.7731 (0.0042)	0.8688 (0.0034)	0.8539 (0.0035)	48.0 (0.1)	72.1 (0.2)	55.4 (0.2)	53.5 (0.2)	479.6 (1.1)	519.7 (2.3)	337.6 (1.3)	356.8 (1.2)
	10	0.7515 (0.0043)	0.7844 (0.0041)	0.8619 (0.0035)	0.8599 (0.0035)	47.2 (0.1)	68.7 (0.2)	53.6 (0.2)	51.8 (0.2)	472.2 (0.7)	438.6 (1.8)	323.5 (1.1)	342.6 (1.0)
	15	0.7537 (0.0043)	0.7951 (0.0040)	0.8630 (0.0034)	0.8512 (0.0036)	47.0 (0.1)	65.7 (0.2)	53.0 (0.2)	51.4 (0.2)	470.2 (0.6)	396.9 (1.5)	324.7 (0.1)	340.2 (0.9)
	20	0.7521 (0.0043)	0.7914 (0.0041)	0.8670 (0.0034)	0.8539 (0.0035)	46.8 (<0.1)	62.5 (0.2)	53.0 (0.2)	51.2 (0.2)	468.4 (0.5)	379.5 (1.3)	336.3 (0.9)	345.2 (0.8)

TABLE 5.5B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when
 $k = 10$, $P^* = 0.75$, $\sigma/\delta^* = 2, 3$; $n = 5(5)20$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (21)	5					21.7 (<0.1)	31.8 (0.1)	36.2 (0.2)	31.6 (0.1)	216.6 (0.5)	215.2 (0.9)	206.8 (0.8)	206.2 (0.7)
	10					21.3 (<0.1)	29.9 (0.1)	34.4 (0.1)	30.1 (<0.1)	212.6 (0.3)	191.7 (0.6)	202.9 (0.6)	197.8 (0.2)
	15					21.1 (<0.1)	28.2 (0.1)	34.4 (0.1)	29.9 (<0.1)	211.5 (0.2)	197.8 (0.4)	220.6 (0.5)	210.2 (0.2)
	20					21.5 (<0.1)	27.3 (0.1)	35.2 (0.1)	30.4 (0.1)	215.2 (0.2)	221.6 (0.2)	249.4 (0.4)	235.2 (0.3)
3 (47)	5					48.0 (0.1)	72.3 (0.2)	79.0 (0.3)	68.4 (0.2)	479.6 (1.1)	531.7 (2.2)	448.1 (1.9)	442.0 (1.6)
	10					47.2 (0.1)	69.9 (0.1)	73.8 (0.3)	65.3 (0.2)	472.2 (0.7)	466.4 (1.7)	416.2 (1.4)	418.0 (1.2)
	15					47.0 (0.1)	67.9 (0.2)	72.6 (0.3)	64.6 (0.2)	470.2 (0.6)	434.4 (1.5)	412.0 (1.2)	414.8 (1.0)
	20					46.8 (<0.1)	66.7 (0.2)	72.1 (0.2)	64.1 (0.2)	468.4 (0.5)	421.3 (1.2)	419.3 (1.1)	414.9 (0.9)

TABLE 5.6A

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when

$$k = 10, P^* = 0.90, \sigma/\delta^* = 2.3; n = 5(5)20$$

Least-favorable (LF) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (36)	5	0.9065 (0.0029)	0.9088 (0.0029)	0.9515 (0.0021)	0.9424 (0.0023)	38.1 (0.1)	49.8 (0.1)	33.6 (0.1)	33.1 (0.1)	380.7 (0.8)	387.0 (1.6)	219.4 (0.8)	234.2 (0.8)
	10	0.8978 (0.0030)	0.9145 (0.0028)	0.9453 (0.0011)	0.9405 (0.0012)	37.0 (0.1)	45.6 (0.1)	32.2 (0.1)	31.7 (<0.1)	370.0 (0.5)	299.7 (1.2)	210.5 (0.3)	223.2 (0.3)
	15	0.9037 (0.0030)	0.9136 (0.0028)	0.9499 (0.0022)	0.9465 (0.0023)	36.7 (<0.1)	41.3 (0.1)	31.9 (0.1)	31.3 (0.1)	366.7 (0.4)	266.2 (0.9)	219.9 (0.5)	226.3 (0.5)
	20	0.9087 (0.0029)	0.9170 (0.0028)	0.9537 (0.0021)	0.9444 (0.0023)	36.5 (<0.1)	37.6 (0.1)	32.3 (0.1)	31.6 (0.1)	365.0 (0.4)	262.0 (0.6)	243.1 (0.4)	244.7 (0.4)
3 (81)	5	0.8952 (0.0031)	0.9089 (0.0029)	0.9424 (0.0012)	0.9396 (0.0012)	85.0 (0.2)	112.3 (0.3)	73.8 (0.2)	72.9 (0.1)	850.0 (1.9)	966.0 (3.4)	478.3 (0.9)	514.0 (0.9)
	10	0.9003 (0.0030)	0.9097 (0.0029)	0.9423 (0.0012)	0.9383 (0.0012)	82.5 (0.1)	106.9 (0.2)	70.5 (0.1)	70.2 (0.1)	825.1 (1.2)	827.0 (2.7)	451.2 (0.7)	488.4 (0.7)
	15	0.9046 (0.0029)	0.9104 (0.0029)	0.9406 (0.0012)	0.9370 (0.0012)	81.8 (0.1)	104.8 (0.2)	69.6 (0.1)	69.1 (0.1)	818.0 (1.0)	744.2 (2.5)	444.1 (0.7)	480.3 (0.6)
	20	0.9037 (0.0030)	0.9081 (0.0029)	0.9420 (0.0012)	0.9369 (0.0012)	81.6 (0.1)	101.9 (0.2)	69.2 (0.1)	68.5 (0.1)	815.9 (0.8)	678.2 (2.4)	444.8 (0.7)	476.5 (0.6)

TABLE 5.6B

Estimates of achieved probability of a correct selection ($\bar{p}$), expected number of vector-observations ($\bar{v}$) and expected total number of observations ($\bar{t}$) to terminate sampling, for four procedures when

$$k = 10, P^* = 0.90, \sigma/\delta^* = 2, 3; n = 5(5)20$$

Equal-means (EM) configuration

σ/δ^* (n_{SS})	n_1	$\bar{p}$				$\bar{v}$				$\bar{t}$			
		B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)	B-D-S	G-K	P-H($\delta^*/4$)	P-H($\delta^*/2$)
2 (36)	5					38.1 (0.1)	49.9 (0.1)	54.1 (0.2)	46.8 (0.2)	380.7 (0.8)	407.1 (1.5)	322.0 (1.3)	317.2 (1.1)
	10					37.0 (0.1)	47.5 (0.1)	50.5 (0.2)	44.0 (0.1)	370.0 (0.5)	343.5 (1.1)	298.7 (0.9)	296.3 (0.8)
	15					36.7 (<0.1)	46.2 (0.1)	49.4 (0.2)	43.1 (0.1)	366.7 (0.4)	318.7 (0.9)	298.9 (0.8)	292.6 (0.6)
	20					36.5 (<0.1)	44.7 (0.1)	49.3 (0.1)	42.9 (0.1)	365.0 (0.4)	310.8 (0.7)	313.9 (0.7)	301.6 (0.5)
3 (81)	5					85.0 (0.2)	112.4 (0.3)	118.5 (0.2)	102.8 (0.2)	850.0 (1.9)	982.6 (3.2)	701.3 (1.4)	693.9 (1.2)
	10					82.5 (0.1)	107.3 (0.2)	109.7 (0.2)	96.4 (0.1)	825.1 (1.2)	867.6 (2.6)	641.7 (1.0)	644.8 (0.9)
	15					81.8 (0.1)	106.3 (0.2)	106.8 (0.2)	94.7 (0.1)	818.0 (1.0)	813.6 (2.4)	622.7 (0.9)	631.2 (0.7)
	20					81.6 (0.1)	105.0 (0.2)	105.6 (0.2)	93.9 (0.1)	815.9 (0.8)	764.8 (2.2)	617.9 (0.8)	625.7 (0.7)

number of observations per population necessary to guarantee the $(\sigma/\delta^*, P^*)$ —requirement using the Bechhofer (1954) single-stage procedure; the constants c_{k,P^*} used to calculate these single-stage sample sizes are provided in Table I of Bechhofer (1954), which is for the common known variance problem, i.e., $n_{SS} = [(c_{k,P^*}\sigma/\delta^*)^2]^+$. It will be seen that these n_{SS} -values play a crucial role in the interpretation of the $(\bar{p}, \bar{v}, \bar{t})$ values. Each of the other numbers in parentheses in the body of the table is the estimated standard error of the estimate above it.

6. DISCUSSION OF MC SAMPLING RESULTS

We first discuss the performances separately for each procedure. Then in Section 6.4 we compare the procedures with respect to the various performance characteristics of interest. The key performance summaries are given in Tables 6.1 and 6.2.

6.1 Discussion of B-D-S

It will be helpful to use the single-stage procedure of Bechhofer (1954) as a reference point when studying the performance of B-D-S. If $k(n_1 - 1)$, the first-stage d.f. available for estimating σ^2 , is sufficiently large (greater than 50, “say”) then σ^2 can be estimated with high precision. If, in addition, $n_1 \leq n_{SS}$, then the performance of B-D-S mimics that of the single-stage procedure in that the achieved $P\{CS|LF\} \sim P^*$ and $E\{V|LF\} \sim n_{SS}$ where V is the total number of vector-observations required by B-D-S. This can be seen, e.g., in Table 5.3a for $k = 5$, $P^* = 0.75$, $\sigma/\delta^* = 3$ where $n_{SS} = 31$ (150 d.f.); here for $n_1 = 5, 10, 15, 20$ the estimated achieved $P\{CS|LF\}$ is $\bar{p} = 0.7524, 0.7494, 0.7597, 0.7532$, respectively, with corresponding $\bar{v}$ -values of 32.7, 31.9, 31.6, 31.5. On the other hand, if $n_1 > n_{SS}$ which is the case in Table 5.3a for $k = 5$, $P^* = 0.75$, $\sigma/\delta^* = 2$ where $n_{SS} = 14$ (65 d.f.), then for $n_1 = 15$ we have $\bar{p} = 0.7775$ with $\bar{v} = 15.6$ while for $n_1 = 20$ we have $\bar{p} = 0.8381$ with $\bar{v} = 20.0$, i.e., $n_1 = 20$ is many more vector-observations than necessary to guarantee the probability requirement using the single-stage procedure, and the higher-than-specified probability achieved by B-D-S in the LF-configuration was purchased at the expense of the excess vector-observations.

6.2 Discussion of G-K

As with B-D-S, the single-stage procedure can serve as a reference point for G-K. It was mentioned in Remark 3.3 that the G-K procedure $\mathcal{P}_3$ (which is studied herein) is conservative, i.e., the $P\{CS|LF\} > P^*$ (usually by quite a bit) even when $n_1 \leq n_{SS}$; this is true for all $k, P^*, \sigma/\delta^*$ in our study. See, e.g., the $\bar{p}$ -values for $\sigma/\delta^* = 3$ in

Tables 5.1a–5.6a. This overprotection is accompanied by larger than necessary associated values of $\bar{v}$ and $\bar{t}$. Since neither we nor G-K simulated the performances of the G-K procedures, $\mathcal{P}_1$ and $\mathcal{P}_2$, to estimate their achieved $P\{\text{CS}|\text{LF}\}$, we do not know whether they too overprotect by quite a bit and have large associated values of $\bar{v}$ and $\bar{t}$. In Table 4.1 of G-K (1984) it can be seen that $\mathcal{P}_3$ is not uniformly dominated by $\mathcal{P}_1$ or $\mathcal{P}_2$ in terms of $\bar{t}$; in fact, in several regions of the table, $\mathcal{P}_3$ dominates both $\mathcal{P}_1$ and $\mathcal{P}_2$. However, the performances of $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$ in terms of $\bar{v}$ or $\bar{t}$ are actually difficult to compare in the LF- or EM-configurations because, almost certainly, they achieve different $P\{\text{CS}|\text{LF}\}$ -values.

Remark 6.1: In G-K (1984), p. 92, it is stated that their MC sampling results strongly indicate that the G-K elimination-type procedure performs much better than B-D-S in terms of expected total sample size. However, this conclusion cannot be deduced from the reported MC results since, as we mentioned in Remark 3.2, we believe that the G-K MC results for B-D-S are erroneous. (Their results for $\mathcal{P}_3$ when $k = 5$, $P^* = 0.90$, LF and EM given on their p.86 are consistent with ours given in Tables 5.4a and 5.4b. They simulated $k = 15$ but we did not.) Whether or not the G-K elimination-type procedure will perform better than B-D-S in the LF-configuration will depend in large measure on the extent to which their lower bound on $P\{\text{CS}|\text{LF}\}$ can be improved for the procedure under consideration.

6.3 Discussion of P-H

The P-H procedure with $\lambda = \delta^*/4$ achieved a $P\{\text{CS}|\text{LF}\}$ much higher than P^* for all of the 48 cases studied. For all of these cases $P\text{-H}(\delta^*/4)$ achieved a higher $P\{\text{CS}|\text{LF}\}$ than did $P\text{-H}(\delta^*/2)$ (and indeed higher than B-D-S or G-K). This is partially explained by the fact that $P\text{-H}(\delta^*/4)$ tends to eliminate “non-contending” populations earlier than does $P\text{-H}(\delta^*/2)$, and hence reduces the number of populations from among which the “best” one must be selected. Also, the lower bound on the $P\{\text{CS}|\text{LF}\}$ which is used to determine a_λ^* does not take into account the fact that inferior populations can eliminate each other; so the more likely this becomes, the higher the value of the achieved $P\{\text{CS}\}$. On the other hand the choice $\lambda = \delta^*/2$ yields a P-H procedure which conditional on s^2 has the smallest maximum number of stages; this property is highly desirable, particularly in the EM-configuration, when k is large, and/or P^* is high, and/or σ/δ^* is large since all of these conditions lead to large numbers of vector-observations and total number of observations.

By comparing equations (3.2) and (3.4), it can be shown that at the start of the second stage, G-K tends to eliminate more populations for any observed s^2 than either P-H($\delta^*/2$) or P-H($\delta^*/4$); however, G-K must take a common number of observations per stage from each remaining population on all subsequent vector-observations while P-H($\delta^*/4$) and P-H($\delta^*/2$) take a non-increasing number of observations per vector-observation in succeeding stages. This usually leads to smaller $E\{V\}$ - and $E\{T\}$ -values for P-H($\delta^*/4$) and P-H($\delta^*/2$) than for G-K. If $\mu_{[k]} - \mu_{[k-1]} \gg \sigma$, then G-K can be expected to yield the smallest $E\{V\}$ - and $E\{T\}$ -values.

6.4 Relative performances of the procedures

We have summarized the estimated “winners” among the four procedures B-D-S, G-K, P-H($\delta^*/4$) and P-H($\delta^*/2$) with respect to $E\{V\}$ and $E\{T\}$. These are given in Tables 6.1 and 6.2 for $\sigma/\delta^* = 2$ and 3, respectively, for the LF- and EM-configurations when $P^* = 0.75, 0.90$; $k = 3, 5, 10$; $n_1 = 5, 10, 15, 20$. Although not exhibited, P-H($\delta^*/4$) achieved the highest estimated $P\{CS|LF\}$ in all 48 cases.

B-D-S is the estimated winner for $E\{V|LF\}$ almost uniformly in k, n_1 and σ/δ^* for $P^* = 0.75$ while P-H($\delta^*/2$) is the estimated winner for $E\{V|LF\}$ uniformly in k, n_1 and σ/δ^* for $P^* = 0.90$. There is no consistent estimated winner for $E\{T|LF\}$ when $P^* = 0.75$ although when $P^* = 0.90$, P-H($\delta^*/4$) tends to be superior for $\sigma/\delta^* = 3$.

B-D-S is the estimated winner for $E\{V|EM\}$ almost uniformly in k, n_1 and σ/δ^* for both $P^* = 0.75$ and 0.90. B-D-S tends to be the estimated winner for $E\{T|EM\}$ when k is small and $P^* = 0.75$ both for $\sigma/\delta^* = 2$ and 3, while P-H($\delta^*/2$) or P-H($\delta^*/4$) tends to be superior when $k = 5$ or 10 and $P^* = 0.75$ for $\sigma/\delta^* = 3$. P-H($\delta^*/2$) is the estimated winner for $E\{T|EM\}$ almost uniformly in k, n_1 and σ/δ^* for $P^* = 0.90$ except for $k = 3, h_1 = 20, \sigma/\delta^* = 2$ and $k = 10, h_1 = 10, 15, 20, \sigma/\delta^* = 3$.

This latter case might be regarded as an anomaly along with the estimated $E\{T|LF\}$ -values for $k = 5, P^* = 0.90, n_1 = 15$ and 20, $\sigma/\delta^* = 2$, and $k = 3, P^* = 0.90, n_1 = 20, \sigma/\delta^* = 3$, and also the estimated $E\{V|EM\}$ -values for $k = 3, P^* = 0.90, n_1 = 15$ and 20, $\sigma/\delta^* = 3$. It is to be emphasized that these so-called anomalies are not attributable to sampling errors.

The common number of observations (n_1) taken from every population in the first stage plays a crucial role when the common variance is unknown. Most importantly, it provides the total number of d.f. $\nu = k(n_1 - 1)$ on which the estimate of σ^2 is based;

TABLE 6.1

Procedure having minimum $\bar{v}$ or minimum $\bar{t}$ for the
 LF- and EM-configurations of the population means
 when $P^* = 0.75, 0.90$; $k = 3, 5, 10$; $n_1 = 5, 10, 15, 20$
 ($\sigma/\delta^* = 2$)

P^*	k	n_1	LF		EM	
			$\bar{v}_{\min}$	$\bar{t}_{\min}$	$\bar{v}_{\min}$	$\bar{t}_{\min}$
0.75	3	5	B-D-S	$\delta^*/2$	B-D-S	B-D-S
		10	B-D-S	B-D-S	B-D-S	B-D-S
		15	B-D-S	B-D-S	B-D-S	B-D-S
		20	B-D-S	B-D-S, G-K	B-D-S, G-K	B-D-S
	5	5	B-D-S	$\delta^*/4$	B-D-S	B-D-S
		10	B-D-S	$\delta^*/2$	B-D-S	B-D-S
		15	B-D-S	B-D-S	B-D-S	B-D-S
		20	B-D-S	B-D-S	B-D-S	B-D-S
	10	5	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	B-D-S	$\delta^*/4$	B-D-S	G-K
		15	B-D-S	G-K	B-D-S	G-K
		20	B-D-S	G-K	B-D-S	B-D-S
0.90	3	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/2$	B-D-S	$\delta^*/2$
		15	$\delta^*/2$	$\delta^*/2$	B-D-S	$\delta^*/2$
		20	$\delta^*/2$	$\delta^*/2$	B-D-S	B-D-S
	5	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		15	$\delta^*/2$	$\delta^*/2$	B-D-S	$\delta^*/2$
		20	$\delta^*/2$	$\delta^*/2$	B-D-S	$\delta^*/2$
	10	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		15	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		20	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$

additionally, it provides the observations on which the first estimates of the μ_i ($1 \leq i \leq k$) are based. As pointed out in Sections 6.1-6.3, if $n_1 > n_{SS}$, the procedures B-D-S, G-K, P-H($\delta^*/4$) and P-H($\delta^*/2$) tend to overprotect, i.e., $P\{CS|LF\} > P^*$ (sometimes by a large amount), with associated larger-than-required values of $E\{V|LF\}$ and $E\{T|LF\}$. If $n_1 < n_{SS}$ the overprotection is less pronounced.

TABLE 6.2

Procedure having minimum $\bar{v}$ or minimum $\bar{t}$ for the
 LF- and EM-configurations of the population means
 when $P^* = 0.75, 0.90$; $k = 3, 5, 10$; $n_1 = 5, 10, 15, 20$
 ($\sigma/\delta^* = 3$)

P^*	k	n_1	LF		EM	
			$\bar{v}_{\min}$	$\bar{t}_{\min}$	$\bar{v}_{\min}$	$\bar{t}_{\min}$
0.75	3	5	B-D-S	$\delta^*/2$	B-D-S	B-D-S
		10	$\delta^*/2$	$\delta^*/2$	B-D-S	B-D-S
		15	B-D-S	$\delta^*/2$	B-D-S	B-D-S
		20	B-D-S	B-D-S	B-D-S	B-D-S
	5	5	B-D-S	$\delta^*/4$	B-D-S	B-D-S
		10	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/2$
		15	B-D-S	$\delta^*/2$	B-D-S	$\delta^*/2$
		20	B-D-S	$\delta^*/2$	B-D-S	B-D-S
	10	5	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/4$
		15	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/4$
		20	B-D-S	$\delta^*/4$	B-D-S	$\delta^*/2$
0.90	3	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		15	$\delta^*/2$	$\delta^*/4$	$\delta^*/2$	$\delta^*/2$
		20	$\delta^*/2$	$\delta^*/2$	$\delta^*/2$	$\delta^*/2$
	5	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		15	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		20	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
	10	5	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/2$
		10	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/4$
		15	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/4$
		20	$\delta^*/2$	$\delta^*/4$	B-D-S	$\delta^*/4$

Often overlooked is the role that n_1 plays with respect to the variance of V ($\text{Var}\{V\}$), variance of T ($\text{Var}\{T\}$) and the maximum observed V ($\max v$) and maximum observed T ($\max t$)—values obtained in a finite set of experiments. Using MC sampling we have estimated these quantities as well as several of the upper percentiles of V and T for the four procedures B-D-S, G-K, P-H($\delta^*/2$) and

P-H($\delta^*/4$) when $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 3$, LF. The results for V are given in Tables 6.3 and 6.4 while those for T are given in Tables 6.5 and 6.6. (We felt that this one case was sufficient to illustrate the issues involved.)

TABLE 6.3

Number of vector-observations (v)
Results for 10,000 MC experiments
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 3$, LF

Procedure	n_1	$\bar{v}^a/$	s_v	Percentiles for v			max v
				90	95	99	
B-D-S	5	67.5 (0.2)	21.4	96	106	129	182
	10	63.7 (0.1)	13.3	81	87	99	127
	15	62.9 (0.1)	10.5	77	81	90	114
	20	62.5 (0.1)	9.0	75	78	85	106
	25	62.3 (0.1)	8.0	73	76	82	97
	30	62.2 (0.1)	7.1	71	74	80	92
G-K	5	93.8 (0.3)	32.9	135	150	179	247
	10	84.3 (0.3)	26.2	112	121	139	181
	15	78.7 (0.3)	27.3	105	112	124	148
	20	73.5 (0.3)	29.0	101	106	117	137
	25	68.3 (0.3)	29.2	98	103	112	137
	30	64.5 (0.3)	28.4	96	100	108	131

^a/Corresponding single-stage procedure requires 61 vector-observations

TABLE 6.4

Number of vector-observations (v)
 Results for 10,000 MC experiments
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 3$, LF

Procedure	n_1	$\frac{\bar{a}}{\bar{v}}$	s_v	Percentiles for v			max v
				90	95	99	
P-H($\delta^*/2$)	5	55.0 (0.2)	24.4	88	100	124	189
	10	50.9 (0.2)	19.8	78	88	106	156
	15	49.8 (0.2)	18.2	75	83	99	140
	20	49.4 (0.2)	17.6	74	82	97	127
	25	49.5 (0.2)	17.0	73	81	96	125
	30	49.7 (0.2)	16.0	73	80	93	117
P-H($\delta^*/4$)	5	55.8 (0.3)	28.1	93	109	146	257
	10	51.1 (0.2)	23.9	84	96	124	165
	15	50.4 (0.2)	22.2	81	92	116	180
	20	49.7 (0.2)	21.3	79	90	112	172
	25	49.7 (0.2)	20.5	78	89	112	162
	30	50.6 (0.2)	19.5	78	89	110	166

^{a/} Corresponding single-stage procedure requires 61 vector-observations

TABLE 6.5

Total number of observations (t)
 Results for 10,000 MC experiments
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 3$, LF

Procedure	n_1	$\bar{t}^a/$	s_t	Percentiles for t			max t
				90	95	99	
B-D-S	5	337.4 (1.1)	107.0	480	530	645	910
	10	318.2 (0.7)	66.3	405	435	495	635
	15	314.7 (0.5)	52.5	385	405	450	570
	20	312.7 (0.5)	45.1	375	390	425	530
	25	311.6 (0.4)	39.9	365	380	410	485
	30	310.9 (0.4)	35.5	355	370	400	460
G-K	5	414.0 (1.8)	184.3	655	735	890	1235
	10	342.0 (1.4)	141.7	525	570	680	905
	15	303.0 (1.3)	128.9	470	515	590	740
	20	275.5 (1.2)	118.2	432	475	545	670
	25	256.9 (1.1)	105.3	397	437	510	640
	30	248.0 (0.9)	92.0	374	410	475	655

^{a/} Corresponding single-stage procedure requires a total of 305 observations

TABLE 6.6

Total number of observations (t)
 Results for 10,000 MC experiments
 $k = 5$, $P^* = 0.90$, $\sigma/\delta^* = 3$, LF

Procedure	n_1	$\bar{t}^a/$	s_t	Percentiles for t			max t
				90	95	99	
P-H($\delta^*/2$)	5	219.7 (1.0)	95.3	347	395	493	757
	10	201.6 (0.7)	73.5	299	335	416	643
	15	197.7 (0.7)	65.5	287	317	378	572
	20	197.0 (0.6)	60.3	279	308	363	442
	25	201.0 (0.6)	55.7	279	306	358	466
	30	207.6 (0.5)	49.9	277	305	352	472
P-H($\delta^*/4$)	5	213.4 (1.0)	101.9	347	404	526	808
	10	192.5 (0.8)	81.2	301	344	440	647
	15	190.4 (0.7)	72.7	289	323	403	593
	20	191.8 (0.7)	67.2	282	317	386	587
	25	196.5 (0.6)	61.4	281	317	385	554
	30	206.8 (0.6)	56.3	286	318	382	580

^{a/} Corresponding single-stage procedure requires a total of 305 observations.

We note for all four procedures that $\bar{v}$ and $\bar{t}$ decrease monotonically as n_1 increases from 5 to 15 but change little for $n_1 \geq 20$; however, as n_1 increases from 5 to 30 the sample standard deviations s_v and s_t decrease monotonically (except for s_v with G-K) as does the estimated 99th percentile of both V and T . Thus values of n_1 larger than 15 play only a modest role in stabilizing the estimates of $E\{V|LF\}$ and $E\{T|LF\}$ for all procedures but a major role in decreasing $\text{Var}\{V|LF\}$ and $\text{Var}\{T|LF\}$ for all procedures.

7. CONCLUDING REMARKS

As might have been anticipated, it is very difficult (if not impossible) to give general rules which will serve the experimenter well in all situations when the value of the common variance is unknown and the spacing of the population means is also completely unknown. However, several general guidelines can be given. These can be improved upon if certain partial information becomes available.

a) We have pointed out the critical role that the common initial sample (n_1) per population plays in the performances of all of the procedures. If n_1 is “too large,” then on the average many more observations (and stages) will be taken to guarantee the probability requirement than would be the case if the value of the common σ^2 were known; this is true for all of the procedures under consideration. Hence one might conclude that it is more advisable to start with “very small” n_1 and, particularly for the P-H procedures, to continue observation-taking a vector-at-a-time until the rule calls for termination of sampling. But this policy itself has serious drawbacks if $\nu = k(n_1 - 1)$ is not sufficiently large (say > 50); for then $S_\nu^2 \gg \sigma^2$ with high probability resulting in large values of $E\{N\}$ and $E\{T\}$ and excessively large values of N and T . Clearly, if k is large, then a moderate n_1 will protect the experimenter against unnecessarily large values of S_ν^2 and hence lead to moderate values of $E\{N\}$ and $E\{T\}$. If the experimenter has an order-of-magnitude idea of the value of σ^2 then n_1 could be chosen as (say) $(2/3)n_{SS}$ subject, of course, to the recommendation that $k(n_1 - 1) > 50$.

b) Of course the key question that must be answered is which procedure should be used in a given environment. The answer will depend in large part on which measure, $E\{V\}$ or $E\{T\}$, the experimenter wishes to minimize; and this in turn will depend on whether the cost of observation taking is more closely related to the number of vector-observations (which might be the case if only one vector-observation can be taken per unit time, e.g., per day) or the total number of observations (which would be the case if individual observations are very expensive). Additionally, the choice will depend on the

number of populations (k) under consideration, the specified value of P^* (whether it is “high,” “moderate” or “low”) and the first-stage sample size--all of which are under the control of the experimenter. Finally, the choice will depend on any (even partial) information which may be available concerning the value of σ/δ^* (where σ is unknown and δ^* is specified) and the unknown spacing of the population means.

For example, to minimize $E\{V|LF\}$ or $E\{V|EM\}$ when P^* is low or moderate then B-D-S would appear to be appropriate; to minimize $E\{V|LF\} - [E\{V|EM\}]$ when P^* is high then $P-H(\delta^*/2)$ [B-D-S] should be used. In general, to minimize $E\{T\}$ either $P-H(\delta^*/4)$ or $P-H(\delta^*/2)$ is preferable; in particular, to minimize $E\{T|LF\} - [E\{T|EM\}]$ when P^* is high then $P-H(\delta^*/4)$ [$P-H(\delta^*/2)$] should be used. In many situations the performances of the procedures in terms of $\bar{v}$ and/or $\bar{t}$ are very close, and the reader is referred to Tables 5.1-5.6 for finer assessments. Of course, the P-H procedures should not be used when it is not practical to take observations a vector-at-a-time. And one should not use G-K for high P^* unless it is strongly believed that $\mu_{[k]} - \mu_{[k-1]} \gg \sigma$ for in this latter situation most of the non-contending populations will be eliminated after the first stage. Finally we suggest that future research effort be focused on devising elimination-type procedures which are less conservative than P-H and G-K since we have shown that these latter procedures overprotect, often by a sizable amount, in the LF-configuration.

Remark 7.1: After all of the simulation runs had been completed and almost all of the text had been written we learned of the article of Kao and Lai (1980) which described a procedure which is a direct competitor to the ones described in the present paper. However, because of time and space considerations we opted not to study its performance characteristics at this time.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the efforts of Ms. Kathy King who typed the many versions of this manuscript with patience, good humor, and her usual expertise. The research was partially supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University, U.S. Army Research Contract DAAL03-86-K-0046 at Cornell University, the Natural Sciences and Engineering Research Council of Canada, and by the National Science Foundation grant DMS-8905654 at the University of North Carolina at Chapel Hill.

BIBLIOGRAPHY

- Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances, *Ann. Math. Statist.* 25, 16-39.
- Bechhofer, R.E. and Dunnett, C.W. (1988). Percentage Points of Multivariate Student t Distributions, *Selected Tables in Mathematical Statistics*, Vol. 11, American Mathematical Society, Providence, Rhode Island.
- 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.
- Bechhofer, R.E. and Goldsman, D.M. (1989). A comparison of the performances of procedures for selecting the normal population having the largest mean when the variances are known and equal. *Contributions to Probability and Statistics--Essays in Honor of Ingram Olkin* (Eds. L.J. Gleser, M.D. Perlman, S.J. Press, A.R. Sampson), Springer-Verlag, New York, 303-317.
- Dudewicz, E.J. (1971). Non-existence of a single-sample selection procedure whose $P(CS)$ is independent of the variances, *S. Afr. Statist. J.* 5, 37-39.
- Gupta, S.S. and Kim, W-C. (1984). A two-stage elimination-type procedure for selecting the largest of several normal means with a common unknown variance. *Design of Experiments: Ranking and Selection* (Eds. T.J. Santner and A.C. Tamhane), Marcel Dekker, NY, 77-94.
- Gupta, S.S., Panchapakesan, S. and Sohn, J.K. (1985). On the distribution of the Studentized maximum of equally correlated normal random variables. *Comm. Statist.--Sim. Comput.* 14(1), 103-135.
- Hartmann, M. (1988). An improvement on Paulson's sequential ranking procedure. *Sequential Analysis* 7(4), 363-372.
- Hartmann, M. (1990). An improvement on Paulson's procedure for selecting the population with the largest mean from k normal populations with a common unknown variance. Submitted for publication to Sequential Analysis.
- Kao, S.C. and Lai, T.L. (1980). Sequential selection procedures based on confidence sequences for normal populations. *Comm. Statist.--Theor. Meth.* A9, 1657-1676.
- 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. and Bechhofer, R.E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. *Comm. 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. *Comm. Statist.--Theor. Meth.* A8, 337-358.