

AD-A040 067

CORNELL UNIV ITHACA N Y SCHOOL OF OPERATIONS RESEARC--ETC F/6 12/1
A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE --ETC(U)
JAN 77 A C TAMHANE, R E BECHHOFFER

DAA029-73-C-0008

UNCLASSIFIED

TR-323

NL

| OF |

AD
A040067



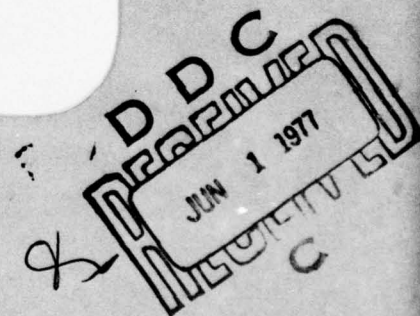
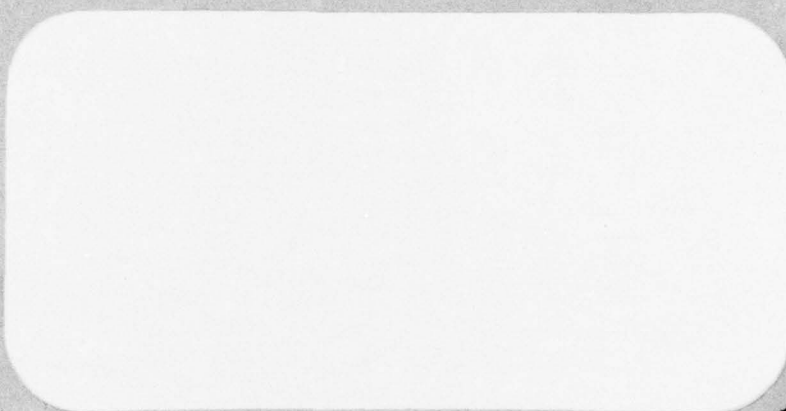
END

DATE
FILMED
6-77

ADA040067

SCHOOL
OF
OPERATIONS RESEARCH
AND
INDUSTRIAL ENGINEERING

SEE 1473
IN BACK



COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

AD No. _____
DDC FILE COPY

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

(1)

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

TECHNICAL REPORT NO. 323

January 1977

A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING
FOR SELECTING THE LARGEST NORMAL MEAN

by

Ajit C. Tamhane
Northwestern University

Robert E. Bechhofer
Cornell University

Prepared under contracts
DAAG29-73-C-0008 and DAAG29-77-C-0003,
U.S. Army Research Office-Durham
and
N00014-75-C-0586,
Office of Naval Research

Approved for Public Release; Distribution Unlimited.



ACCESSION for		White Section ☒
NTIS		Ref Section ☐
DDC		
UNANNOUNCED		
JUSTIFICATION		
BY DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL. and/or SPECIAL	
A		

THE FINDINGS IN THIS REPORT ARE NOT TO BE
CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE
ARMY POSITION, UNLESS SO DESIGNATED BY
OTHER AUTHORIZED DOCUMENTS.

Table of Contents

Abstract

1. Introduction and summary
2. Preliminaries
 - 2.1 Assumptions
 - 2.2 Goal and probability requirement
3. Background: Single-stage and sequential procedures
4. A two-stage procedure (P_2)
 - 4.1 R-minimax design criterion
 - 4.2 U-minimax design criterion
5. Probability of a correct selection for P_2
 - 5.1 General expression for $P_{\underline{\mu}}\{CS|P_2\}$
 - 5.2 LF-configuration for P_2
 - 5.3 Lower bound for $P_{\underline{\mu}}\{CS|P_2\}$
6. Expected total sample size for P_2
 - 6.1 General expression for $E_{\underline{\mu}}\{T|P_2\}$
 - 6.2 The supremum of $E_{\underline{\mu}}\{T|P_2\}$
7. Optimization problems yielding conservative solutions
 - 7.1 Discrete optimization problems
 - 7.1.1 R-minimax design criterion
 - 7.1.2 U-minimax design criterion
 - 7.2 Continuous optimization problems
8. Constants to implement P_2
 - 8.1 Constants to implement $P_2(U_E)$ and $P_2(R_E)$ for $k = 2$
 - 8.2 Constants to implement $P_2(U_C)$ for $k \geq 3$
9. The performance of P_2 relative to P_1
 - 9.1 $P_2(U_E)$ and $P_2(U_C)$ vs. P_1 for $k = 2$
 - 9.2 $P_2(U_C)$ vs. P_1 for $k \geq 3$
10. Directions of future research
11. Acknowledgments
12. References

Abstract

The problem of selecting the normal population with the largest population mean when the populations have a common known variance is considered. A two-stage procedure is proposed which guarantees the same probability requirement using the indifference-zone approach as does the single-stage procedure of Bechhofer [1954]. The two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than the total number of observations required by the corresponding single-stage procedure, regardless of the configuration of the population means. The saving in expected total number of observations can be substantial, particularly when the configuration of the population means is favorable to the experimenter. The saving is accomplished by screening out "non-contending" populations in the first stage, and concentrating sampling only on "contending" populations in the second stage. The two-stage procedure can be regarded as a composite one which uses a screening subset-type approach (Gupta [1956], [1965]) in the first stage, and an indifference-zone approach (Bechhofer [1954]) applied to all populations retained in the selected subset in the second stage. Constants to implement the procedure for various k and P^* are provided, as are calculations giving the saving in expected total sample size if the two-stage procedure is used in place of the corresponding single-stage procedure.

1. Introduction and Summary

In many practical situations a statistician is faced with the problem of designing an experiment to select one (or more) out of $k \geq 2$ possible competing categories. Typically the categories (populations) are characterized by a real-valued parameter, and the experimenter is interested in selecting the population having the largest (or smallest) parameter value. This population is referred to as the "best" population. Thus, for example, the medical research worker might be studying the response of patients to different types of analgesic drugs in which case his interest might lie in selecting that drug which produces, on the average, the largest period of time without pain, or the agronomist might be conducting field trials with different varieties of grain in which case his purpose might be to select that variety which produces, on the average, the largest yield per acre.

Procedures for achieving such objectives have received considerable attention in recent years. Various probability distributions have been postulated as being appropriate to model these and other real-life problems, and several statistical formulations of these problems have been proposed, and associated statistical selection procedures devised. Recent reviews of the literature with particular reference to the normal means problem appear in Wetherill and Ofosu [1974] and Bechhofer [1975]. The present paper continues the study of the normal means problem, and explores in depth a new approach which has highly desirable properties. This same approach is also applicable to the normal variances problem.

The statistical formulation of the problem is given in Section 2. In Section 3 we sketch the relevant history of the normal means problem, and indicate the virtues and drawbacks of the various procedures which have

been proposed to deal with the problem; the reader is thus enabled to understand the role that our proposed procedure plays. The procedure itself as well as the design criterion that we adopt are described in Section 4. The main analytical results are contained in Sections 5 and 6 which deal with the probability of a correct selection and the expected total sample size, respectively; Section 5.2 discusses a key unsolved problem (that of determining the so-called least favorable configuration of the population means) associated with the procedure, while Section 5.3 contains a strategem which permits us to bypass this difficulty (at the expense of some loss of efficiency of our procedure). In Section 7 we formulate the problem that we must solve to obtain design constants to implement our procedure; tables of these constants are provided in Section 8. The performance of our two-stage procedure relative to that of the best competing single stage procedure is studied in Section 9, and is shown to be highly satisfactory. We conclude in Section 10 with suggestions for future research in this area.

2. Preliminaries

2.1 Assumptions

Let Π_i ($1 \leq i \leq k$) denote a normal population with unknown mean μ_i and known variance σ^2 , and let $\Omega = \{\underline{\mu} = (\mu_1, \dots, \mu_k) | -\infty < \mu_i < \infty (1 \leq i \leq k)\}$ be the parameter space of the μ_i . Denote the ranked values of the μ_i by $\mu_{[1]} \leq \dots \leq \mu_{[k]}$, and let $\delta_{i,j} = \mu_{[i]} - \mu_{[j]}$. We assume that the experimenter has no prior knowledge concerning the pairing of the Π_i with the $\mu_{[j]}$ ($1 \leq i, j \leq k$). Let $\Pi_{(j)}$ denote the population associated with $\mu_{[j]}$. Suppose that $\mu_{[k-r]} < \mu_{[k-r+1]} = \mu_{[k]}$ for some r ($1 \leq r \leq k$) where we define $\mu_{[0]} = -\infty$; then any one of the r populations $\Pi_{(k-r+j)}$ ($1 \leq j \leq r$) is regarded as "best."

2.2 Goal and probability requirement

The goal of the experimenter is to select a best population. This event is referred to as a correct selection (CS). The experimenter restricts consideration to procedures (P) which guarantee the probability requirement

$$P_{\underline{\mu}} \{CS|P\} \geq P^* \quad \forall \underline{\mu} \in \Omega(\delta^*) \quad (2.1)$$

where $\{\delta^*, P^*\}$ $0 < \delta^* < \infty$, $1/k < P^* < 1$ are specified prior to the start of experimentation, and

$$\Omega(\delta^*) = \{\underline{\mu} \in \Omega \mid \delta_{k,k-1} \geq \delta^*\}. \quad (2.2)$$

We refer to $\Omega(\delta^*)$ as the preference zone for a CS and to $\Omega_0(\delta^*) = \Omega - \Omega(\delta^*)$ as the associated indifference zone. The formulation (2.1) is called the indifference-zone approach.

3. Background: Single-stage and sequential procedures

The indifference-zone approach as applied to the normal means (common known variance) problem has received considerable study. Bechhofer [1954] proposed a single-stage procedure which guarantees (2.1); Hall [1959] showed that among single-stage procedures this procedure is "most economical" and Eaton [1967] proved that it has additional desirable decision theoretic properties. Bechhofer, Kiefer, and Sobel [1968] (see also, Bechhofer and Sobel [1954]) proposed an open sequential procedure without elimination which guarantees (2.1). Paulson [1964] proposed a closed sequential procedure with permanent elimination which also guarantees (2.1); Fabian [1974a] (see also, Fabian [1974b] and Lawing and David [1966]) showed how Paulson's procedure could be modified to improve its performance characteristics.

The single-stage procedure (P_1) of Bechhofer requires a common sample size n per population which is chosen in such a way that (2.1) is guaranteed even when $\mu_{[k]} - \mu_{[i]} = \delta^*$ ($1 \leq i \leq k-1$), this being the so-called least favorable (LF) configuration of the population means. However, the procedure is conservative in that if, unknown to the experimenter, $\mu_{[k]} - \mu_{[i]} \geq \delta^*$ ($1 \leq i \leq k-1$) with strict inequality for at least one i -value--in particular if $\mu_{[k]} - \mu_{[k-1]} \gg \delta^*$, then $P_{\underline{\mu}}\{CS|P_1\} > P^*$ for the actual $\underline{\mu} \in \Omega(\delta^*)$ which the experimenter has encountered. If this is the situation he may have greatly overprotected himself, the overprotection having been purchased by the use of a much larger n -value than would have been necessary had the true μ -values been known.

Unlike single-stage procedures, multistage or sequential procedures provide information concerning the true but unknown μ -values as sampling proceeds.

The sequential procedure (P_{S_1}) of Bechhofer, Kiefer, and Sobel takes a single vector of observations at each stage of experimentation. Here the number of stages (N_{S_1}) to terminate experimentation is an unbounded r.v. (For P_{S_1} a vector consists of one observation from each of the k populations.) In addition to guaranteeing (2.1) when the population means are in the LF-configuration, it also reacts to favorable configurations of the population means and thereby tends to terminate experimentation early resulting in $E_{\underline{\mu}}\{N_{S_1}|P_{S_1}\}$ -values which are smaller than n (c.f., B-K-S [1968], Section 12.8.1). (Throughout this paper n will denote the single-stage sample size for P_1 .) However, if P^* is sufficiently close to unity and if $\mu_{[k]} - \mu_{[k-1]} < \delta^*$ --in particular, if $\mu_{[k]} = \mu_{[1]}$, then $E_{\underline{\mu}}\{N_{S_1}|P_{S_1}\} > n$ for the actual $\underline{\mu} \in \Omega_0(\delta^*)$ which the experimenter has

encountered. However, P_{S_1} does have a practical disadvantage: It is open-ended, i.e., although N_{S_1} is finite w.p. 1, it is unbounded; this latter fact may inhibit or prevent the use of the procedure in certain situations.

The sequential procedure (P_{S_2}) of Paulson which takes a single vector of observations at each stage of experimentation, and for which the number of stages (N_{S_2}) to terminate experimentation is a bounded r.v. ($\leq M$), is an adaptive procedure. (For P_{S_2} a vector consists of one observation from each of the k_j populations still retained "in contention" by the procedure at stage j ($1 \leq j \leq M$, $k = k_1 \geq k_2 \geq \dots \geq k_M \geq 2$), those not retained at stage j being permanently eliminated; here the k_j ($2 \leq j \leq M$) are r.v.'s.) In addition to guaranteeing (2.1) when the population means are in the LF-configuration, it also reacts to favorable configurations of the population means, eliminating from further sampling populations which are indicated as not being in contention, and in general terminating experimentation early resulting in $E_{\underline{\mu}}\{N_{S_2} | P_{S_2}\}$ -values which are less than n . (See Ramberg [1966].) Of considerable interest is the fact that if $\mu_{[1]} = \mu_{[k]}$ and P^* is close to unity, then $E_{\underline{\mu}}\{N_{S_2} | P_{S_2}\} < E_{\underline{\mu}}\{N_{S_1} | P_{S_1}\}$ for the actual $\underline{\mu} \in \Omega_0(\delta^*)$ which the experimenter has encountered. The quantity $E_{\underline{\mu}}\{T_{S_2} | P_{S_2}\}$, where $T = \text{total number of observations to terminate experimentation}$, behaves similarly w.r.t. $E_{\underline{\mu}}\{T_{S_1} | P_{S_1}\}$. Since $N_{S_2} \leq M$ we have $E_{\underline{\mu}}\{N_{S_2} | P_{S_2}\} < M$ and $E_{\underline{\mu}}\{T_{S_2} | P_{S_2}\} < kM$ for all $\underline{\mu} \in \Omega$; the bound M is a function of k , $\{\delta^*, P^*\}$ and also of a design parameter λ ($0 < \lambda \leq \delta^*/2$) which is fixed by the experimenter before the start of experimentation.

Even though P_{S_1} and P_{S_2} have certain highly desirable properties relative to P_1 , both being adaptive and therefore being able to capitalize

on favorable configurations of the population means, both have the drawback that they may require many stages to terminate experimentation. Such procedures are often very costly to implement, and in some experimental situations may be completely impractical, e.g., in agriculture where only one stage, i.e., vector of observations, can be obtained each growing season, the number of stages (years) to terminate experimentation would be prohibitively large.

Thus, in this present paper we study a two-stage procedure which takes a fixed number of vectors of observations at each stage of experimentation. The procedure guarantees (2.1) when the population means are in the LF-configuration. It is adaptive, eliminating from further sampling in the second stage populations which are indicated as not being in contention after the first stage, and in general terminating experimentation after the first stage if the configuration of the population means is very favorable, e.g., $\mu[k] - \mu[k-1] \gg \delta^*$. In addition this procedure is designed to be minimax within a certain class of two-stage procedures.

4. A two-stage procedure (P_2)

We propose a two-stage procedure $P_2 = P_2(n_1, n_2, h)$ which depends on non-negative integers n_1, n_2 and a real constant $h \geq 0$ which are determined prior to the start of experimentation. The constants (n_1, n_2, h) depend on k and $\{\delta^*, P^*\}$, and are chosen so that P_2 guarantees (2.1) and possesses a certain minimax property.

Procedure P_2 :

1. In the first stage take n_1 independent observations $x_{ij}^{(1)}$
 $(1 \leq j \leq n_1)$ from Π_i $(1 \leq i \leq k)$, and compute the k

sample means $\bar{X}_i^{(1)} = \sum_{j=1}^{n_1} X_{ij}^{(1)} / n_1$ ($1 \leq i \leq k$). Let

$$\bar{X}_{[k]}^{(1)} = \max_{1 \leq i \leq k} \bar{X}_i^{(1)}. \quad \text{Determine the subset } I \text{ of} \quad (4.1a)$$

$\{1, 2, \dots, k\}$ where $I = \{i | \bar{X}_i^{(1)} \geq \bar{X}_{[k]}^{(1)} - h\}$, and

let Π_I denote the associated subset of $\{\Pi_1, \Pi_2, \dots, \Pi_k\}$.

a) If Π_I consists of one population, stop sampling

and assert that the population associated with

$\bar{X}_{[k]}^{(1)}$ is best.

b) If Π_I consists of more than one population

proceed to the second stage.

2. In the second stage take n_2 additional independent observations $X_{ij}^{(2)}$ ($1 \leq j \leq n_2$) from each population

in Π_I , and compute the cumulative sample means (4.1b)

$\bar{X}_i = \left(\sum_{j=1}^{n_1} X_{ij}^{(1)} + \sum_{j=1}^{n_2} X_{ij}^{(2)} \right) / (n_1 + n_2)$ for $i \in I$. Assert

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

Remark 4.1: This procedure had been proposed previously by Cohen [1959] and Alam [1970]. Due to analytical and computational difficulties, most of their work was limited to the special case $k = 2$.

Remark 4.2: If $h = 0$ ($h = \infty$) the two-stage procedure P_2 reduces to Bechhofer's [1954] single-stage procedure P_1 with single-stage sample size $n = n_1$ ($n_1 + n_2$) per population. Also, the rule determining I in (4.1a) is of the type proposed by Gupta [1956, 1965] in his subset selection procedure.

There are an infinite number of combinations of (n_1, n_2, h) which for any k and $\{\delta^*, P^*\}$ will exactly guarantee (2.1), and different design criteria lead to different choices. In the next sections we consider two of these criteria.

Let S' denote the cardinality of the set I in (4.1a), and let

$$S = \begin{cases} 0 & \text{if } S' = 1 \\ S' & \text{if } S' > 1. \end{cases} \quad (4.2)$$

Then the total sample size required by $P_2(n_1, n_2, h)$ is

$$T = kn_1 + Sn_2. \quad (4.3)$$

Let $E_{\underline{\mu}}\{T|P_2\}$ denote the expected total sample size for $P_2(n_1, n_2, h)$ under $\underline{\mu}$.

4.1 R-minimax design criterion

The design criterion proposed by Alam [1970] is the following: For given k and specified $\{\delta^*, P^*\}$ choose (n_1, n_2, h) to

$$\begin{aligned} & \text{minimize } \sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}\{T|P_2\} \\ & \text{subject to } \inf_{\underline{\mu} \in \Omega(\delta^*)} P_{\underline{\mu}}\{CS|P_2\} \geq P^*, \end{aligned} \quad (4.4)$$

where n_1, n_2 are non-negative integers and $h \geq 0$.

We denote by $(n_1, n_2, h|R_E)$ the exact solution to (4.4), and by $P_2(R_E)$ the procedure using this solution. The R-minimax criterion in which minimization takes place over a restricted portion of Ω insures that $E_{\underline{\mu}}\{T|P_2(R_E)\} \leq kn$ $\forall \underline{\mu} \in \Omega(\delta^*)$ for any given k and specified $\{\delta^*, P^*\}$. However, it ignores what can happen to $E_{\underline{\mu}}\{T|P_2(R_E)\}$ if, unknown to the experimenter, $\underline{\mu} \in \Omega_0(\delta^*)$. Indeed, for $\mu_{[1]} = \mu_{[k]}$ it is possible to have

Let S' denote the cardinality of the set I in (4.1a), and let

$$S = \begin{cases} 0 & \text{if } S' = 1 \\ S' & \text{if } S' > 1. \end{cases} \quad (4.2)$$

Then the total sample size required by $P_2(n_1, n_2, h)$ is

$$T = kn_1 + Sn_2. \quad (4.3)$$

Let $E_{\underline{\mu}}\{T|P_2\}$ denote the expected total sample size for $P_2(n_1, n_2, h)$ under $\underline{\mu}$.

4.1 R-minimax design criterion

The design criterion proposed by Alam [1970] is the following: For given k and specified $\{\delta^*, P^*\}$ choose (n_1, n_2, h) to

$$\begin{aligned} & \text{minimize } \sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}\{T|P_2\} \\ & \text{subject to } \inf_{\underline{\mu} \in \Omega(\delta^*)} P\{CS|P_2\} \geq P^*, \end{aligned} \quad (4.4)$$

where n_1, n_2 are non-negative integers and $h \geq 0$.

We denote by $(n_1, n_2, h|R_E)$ the exact solution to (4.4), and by $P_2(R_E)$ the procedure using this solution. The R-minimax criterion in which minimization takes place over a restricted portion of Ω insures that $E_{\underline{\mu}}\{T|P_2(R_E)\} \leq kn$ $\forall \underline{\mu} \in \Omega(\delta^*)$ for any given k and specified $\{\delta^*, P^*\}$. However, it ignores what can happen to $E_{\underline{\mu}}\{T|P_2(R_E)\}$ if, unknown to the experimenter, $\underline{\mu} \in \Omega_0(\delta^*)$. Indeed, for $\mu_{[1]} = \mu_{[k]}$ it is possible to have

$E_{\underline{\mu}}\{T|P_2(R_E)\} \gg kn$ for P^* sufficiently close to unity (as happens when $E_{\underline{\mu}}\{T\}$ of the Wald-Girshick SPRT is compared to the total single-stage sample size which guarantees the same probability requirement. See Bechhofer [1960] and B-K-S [1968], Section 12.8.1). It is to guard against this latter undesirable possibility that we propose the design criterion described below.

4.2 U-minimax design criterion

Our design criterion is the following: For given k and specified $\{\delta^*, P^*\}$ choose (n_1, n_2, h) to

$$\begin{aligned} & \text{minimize } \sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}}\{T|P_2\} \\ & \text{subject to } \inf_{\underline{\mu} \in \Omega(\delta^*)} P_{\underline{\mu}}\{CS|P_2\} \geq P^*, \end{aligned} \quad (4.5)$$

where n_1, n_2 are non-negative integers and $h \geq 0$.

We denote by $(n_1, n_2, h|U_E)$ the exact solution to (4.5) and by $P_2(U_E)$ the procedure using this solution. Our U-minimax criterion (4.5) in which minimization takes place over the unrestricted parameter space Ω insures that $E_{\underline{\mu}}\{T|P_2(U_E)\} \leq kn \forall \underline{\mu} \in \Omega$ for any given k and specified $\{\delta^*, P^*\}$; in this sense $P_2(U_E)$ is uniformly (in $\underline{\mu}$) superior to P_1 .

As the first step in determining $(n_1, n_2, h|R_E)$ or $(n_1, n_2, h|U_E)$ we find an exact analytical expression for $P_{\underline{\mu}}\{CS|P_2\}$.

5. Probability of a correct selection for P_2

5.1 General expression for $P_{\underline{\mu}}\{CS|P_2\}$

Our result concerning a general expression for $P_{\underline{\mu}}\{CS|P_2\}$ is summarized in the following theorem:

Theorem 5.1: For any $\underline{\mu} \in \Omega$ we have

$$\begin{aligned}
 P_{\underline{\mu}}\{CS|P_2\} = & \sum_{s \in S} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i \in s} \left\{ \int_{x + (\delta_{ki} - h)\sqrt{n_1}/\sigma}^{x + \delta_{ki}\sqrt{n_1}/\sigma} \phi\left[y + (x-z)(p/q)^{1/2} + \frac{\delta_{ki}}{\sigma} \left(\frac{m}{q}\right)^{1/2}\right] d\phi(z) \right\} \\
 & \times \prod_{i \notin s} \phi\left[x + (\delta_{ki} - h)\sqrt{n_1}/\sigma\right] d\phi(y) d\phi(x) \quad (5.1) \\
 & + \sum_{j=1}^{k-1} \sum_{s \in S_j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{x - (\delta_{kj} + h)\sqrt{n_1}/\sigma}^{x - \delta_{kj}\sqrt{n_1}/\sigma} \int_{y + (x-u)(p/q)^{1/2} - \delta_{kj}(m/q)^{1/2}/\sigma}^{\infty} \right. \\
 & \times \left(\prod_{i \in s} \left\{ \int_{x - (\delta_{ij} + h)\sqrt{n_1}/\sigma}^{x - \delta_{ij}\sqrt{n_1}/\sigma} \phi\left[v + (u-z)(p/q)^{1/2} + \frac{\delta_{ki}}{\sigma} \left(\frac{m}{q}\right)^{1/2}\right] d\phi(z) \right\} \right) \\
 & \times \left(\prod_{i \notin s} \phi\left[x - (\delta_{ij} + h)\sqrt{n_1}/\sigma\right] \right) d\phi(v) d\phi(u) \Big] d\phi(y) d\phi(x),
 \end{aligned}$$

where $\phi(\cdot)$ is the standard univariate normal cdf, and

S = the collection of all subsets of $\{1, 2, \dots, k-1\}$,

S_j = the collection of all subsets of $\{1, 2, \dots, j-1, j+1, \dots, k-1\}$,

$m = n_1 + n_2$, $p = n_1/m$, $q = n_2/m$.

Proof: Let $\bar{X}_{(i)}^{(1)} = \sum_{a=1}^{n_1} X_{(i)a}^{(1)} / n_1$ and $\bar{X}_{(i)} = \sum_{j=1}^2 \sum_{a=1}^{n_j} X_{(i)a}^{(j)} / m$ where $X_{(i)a}^{(j)}$ is the a th observation in the j th stage from $\Pi_{(i)}$, all $X_{(i)a}^{(j)}$ ($1 \leq i \leq k$, $1 \leq a \leq n_j$, $j = 1, 2$) being independent. Then

$$P_{\underline{\mu}}\{CS|P_2\} = \sum_{s \in S} P_{\underline{\mu}}\{\bar{X}_{(k)}^{(1)} = \bar{X}_{[k]}^{(1)}, \bar{X}_{(i)}^{(1)} \geq \bar{X}_{[k]}^{(1)} - h \forall i \in s;$$

$$\bar{X}_{(i)}^{(1)} < \bar{X}_{[k]}^{(1)} - h \forall i \notin s; \bar{X}_{(k)} > \bar{X}_{(i)} \forall i \in s\}$$

$$+ \sum_{j=1}^{k-1} \sum_{s \in S_j} P_{\underline{\mu}}\{\bar{X}_{(j)}^{(1)} = \bar{X}_{[k]}^{(1)}, \bar{X}_{(k)}^{(1)} \geq \bar{X}_{[k]}^{(1)} - h, \quad (5.2)$$

$$\bar{X}_{(i)}^{(1)} \geq \bar{X}_{[k]}^{(1)} - h \forall i \in s; \bar{X}_{(i)}^{(1)} < \bar{X}_{[k]}^{(1)} - h \forall i \notin s;$$

$$\bar{X}_{(k)} > \bar{X}_{(j)}, \bar{X}_{(k)} > \bar{X}_{(i)} \forall i \in s\}$$

$$= \sum_{s \in S} A_s + \sum_{j=1}^{k-1} \sum_{s \in S_j} B_{s,j}.$$

Denoting $[(\bar{X}_{(i)}^{(1)} - \mu_{[i]})\sqrt{n_1}/\sigma, (\bar{X}_{(i)} - \mu_{[i]})\sqrt{m}/\sigma]$ by $[X_i, Y_i]$, we see that $[X_i, Y_i]$ has a standard bivariate normal distribution with correlation coefficient $= \sqrt{\rho}$ ($1 \leq i \leq k$). In what follows we shall use the equality

$$\Phi_2[a, b|\rho] = \int_{-\infty}^a \Phi\left[\frac{b - \rho z}{(1 - \rho^2)^{1/2}}\right] d\Phi(z) \quad (5.3)$$

where $\Phi_2[\cdot, \cdot|\rho]$ is the standard bivariate normal cdf with correlation coefficient ρ ($-1 < \rho < 1$).

We first consider

$$A_s = P_{\underline{u}} \{X_k + \delta_{ki} \sqrt{n_1}/\sigma > X_i > X_k + (\delta_{ki} - h) \sqrt{n_1}/\sigma \quad \forall i \in s,$$

$$X_k + (\delta_{ki} - h) \sqrt{n_1}/\sigma > X_i \quad \forall i \notin s,$$

$$Y_k + \delta_{ki} \sqrt{m}/\sigma > Y_i \quad \forall i \in s\}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i \in s} \{ \Phi_2[u + \delta_{ki} \sqrt{n_1}/\sigma, v + \delta_{ki} \sqrt{m}/\sigma | \sqrt{p}] - \Phi_2[u + (\delta_{ki} - h) \sqrt{n_1}/\sigma, v + \delta_{ki} \sqrt{m}/\sigma | \sqrt{p}] \} \\ \times \prod_{i \notin s} \Phi[u + (\delta_{ki} - h) \sqrt{n_1}/\sigma] d\Phi_2[u, v | \sqrt{p}] \quad (5.4)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i \in s} \left\{ \int_{u + (\delta_{ki} - h) \sqrt{n_1}/\sigma}^{u + \delta_{ki} \sqrt{n_1}/\sigma} \Phi[v/\sqrt{q} - z(p/q)^{1/2} + \delta_{ki}(m/q)^{1/2}/\sigma] d\Phi(z) \right\} \\ \times \prod_{i \notin s} \Phi[u + (\delta_{ki} - h) \sqrt{n_1}/\sigma] d\Phi_2[u, v | \sqrt{p}] \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i \in s} \left\{ \int_{x + (\delta_{ki} - h) \sqrt{n_1}/\sigma}^{x + \delta_{ki} \sqrt{n_1}/\sigma} \Phi[y + (x - z)(p/q)^{1/2} + \delta_{ki}(m/q)^{1/2}/\sigma] d\Phi(z) \right\} \\ \times \prod_{i \notin s} \Phi[x + (\delta_{ki} - h) \sqrt{n_1}/\sigma] d\Phi(y) d\Phi(x).$$

The next to last equality was obtained using (5.3), and the last was obtained

by making the transformation $x = u$, $y = (v - u\sqrt{p})/(1-p)^{1/2}$.

We next consider

$$\begin{aligned}
B_{s,j} &= P_{\underline{\mu}} \{X_j - \delta_{kj} \sqrt{n_1}/\sigma > X_k \geq X_j - (\delta_{kj} + h) \sqrt{n_1}/\sigma, \\
&\quad X_j - \delta_{ij} \sqrt{n_1}/\sigma > X_i \geq X_j - (\delta_{ij} + h) \sqrt{n_1}/\sigma \quad \forall i \in s; \\
&\quad X_j - (\delta_{ij} + h) \sqrt{n_1}/\sigma > X_i \quad \forall i \notin s; \\
&\quad Y_k + \delta_{kj} \sqrt{m}/\sigma > Y_j, \quad Y_k + \delta_{ki} \sqrt{m}/\sigma > Y_i \quad \forall i \in s\}
\end{aligned} \tag{5.5}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{x - (\delta_{kj} + h) \sqrt{n_1}/\sigma}^{x - \delta_{kj} \sqrt{n_1}/\sigma} \int_{t - \delta_{kj} \sqrt{m}/\sigma}^{\infty} \left(\prod_{i \in s} \{\phi_2[x - \delta_{ij} \sqrt{n_1}/\sigma, w + \delta_{ki} \sqrt{m}/\sigma | \sqrt{p}]\} \right. \right. \\
&\quad \left. \left. - \phi_2[x - (\delta_{ij} + h) \sqrt{n_1}/\sigma, w + \delta_{ki} \sqrt{m}/\sigma | \sqrt{p}]\} \right) \right. \\
&\quad \left. \times \prod_{i \notin s} \phi[x - (\delta_{ij} + h) \sqrt{n_1}/\sigma] \right] d\phi_2(u, w | \sqrt{p}) d\phi_2(x, t | \sqrt{p}).
\end{aligned}$$

Proceeding as with A_s , we apply (5.3) to (5.5) and then make the transformations $x = x$, $y = (t - x\sqrt{p})/(1-p)^{1/2}$ and $u = u$, $v = (w - u\sqrt{p})/(1-p)^{1/2}$.

Substituting the resulting expression and (5.4) in (5.2) we obtain (5.1).

Corollary 5.1: Let $\underline{\mu}(\delta)$ denote any $\underline{\mu} \in \Omega$ such that $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta$ where $\delta \geq 0$. ($\underline{\mu}(\delta)$ is known as a slippage configuration.) Then we have

$$\begin{aligned}
P_{\underline{\mu}(\delta)} \{CS | P_2\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \int_{x + (\delta - h) \sqrt{n_1}/\sigma}^{x + \delta \sqrt{n_1}/\sigma} \phi[y + (x - z)(p/q)^{1/2} + \delta(m/q)^{1/2}/\sigma] d\phi(z) \right. \\
&\quad \left. + \phi[x + (\delta - h) \sqrt{n_1}/\sigma] \right\}^{k-1} d\phi(y) d\phi(x) \\
&\quad + (k-1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{x - (\delta + h) \sqrt{n_1}/\sigma}^{x - \delta \sqrt{n_1}/\sigma} \int_{y + (x - u)(p/q)^{1/2} - \delta(m/q)^{1/2}/\sigma}^{\infty} \right. \\
&\quad \left. \left\{ \int_{x - h \sqrt{n_1}/\sigma}^x \phi[v + (u - z)(p/q)^{1/2} + \delta(m/q)^{1/2}/\sigma] d\phi(z) \right. \right. \\
&\quad \left. \left. + \phi(x - h \sqrt{n_1}/\sigma) \right\}^{k-2} d\phi(v) d\phi(u) \right] d\phi(y) d\phi(x).
\end{aligned} \tag{5.6}$$

Proof: The proof is straightforward.

Remark 5.1: For $k = 2$, (5.6) simplifies considerably; the resulting expression is given by Alam [1970] as his equation (3.1). Alam also gives an expression for $P_{\underline{\mu}(\delta)}\{CS|P_2\}$ (see his second equation (3.24)); however, we were not able to verify his expression.

5.2 LF-configuration for P_2

In order to solve (4.4) or (4.5) it first is necessary to determine the LF-configuration of the μ_i for P_2 , i.e., to determine any $\underline{\mu}_0 \in \Omega(\delta^*)$ such that

$$P_{\underline{\mu}_0}\{CS|P_2\} = \inf_{\underline{\mu} \in \Omega(\delta^*)} P_{\underline{\mu}}\{CS|P_2\}. \quad (5.7)$$

As a first step toward determining $\underline{\mu}_0$ we now study the monotonicity of $P_{\underline{\mu}}\{CS|P_2\}$ w.r.t. the $\mu_{[i]}$ ($1 \leq i \leq k$).

Lemma 5.1: For fixed k and $\mu_{[i]}$ ($1 \leq i \leq k-1$) and fixed (n_1, n_2, h) , $P_{\underline{\mu}}\{CS|P_2\}$ is non-decreasing in $\mu_{[k]}$.

Proof: From (5.4) and (5.5) we have

$$\begin{aligned} P_{\underline{\mu}}\{CS|P_2\} &= P_{\underline{\mu}} \left\{ \bigcup_{s \in S} \left[\left((X_k + \delta_{ki}\sqrt{n_1}/\sigma > X_i > X_k + (\delta_{ki}-h)\sqrt{n_1}/\sigma \quad \forall i \in s, \right. \right. \right. \\ &\quad \left. \left. X_k + (\delta_{ki}-h)\sqrt{n_1}/\sigma > X_i \quad \forall i \notin s \right) \right. \\ &\quad \left. \bigcup_{j \in S} \left(X_j - \delta_{kj}\sqrt{n_1}/\sigma > X_k > X_j - (\delta_{kj}+h)\sqrt{n_1}/\sigma, \right. \right. \\ &\quad \left. \left. X_j - \delta_{ij}\sqrt{n_1}/\sigma > X_i > X_j - (\delta_{ij}+h)\sqrt{n_1}/\sigma \quad \forall i \in s, i \neq j; \right. \right. \\ &\quad \left. \left. X_j - (\delta_{ij}+h)\sqrt{n_1}/\sigma > X_i \quad \forall i \notin s \right) \right] \cap \{Y_k + \delta_{ki}\sqrt{m}/\sigma > Y_i \quad \forall i \in s\} \} \\ &= P\{A(\underline{\mu})\} \end{aligned} \quad (5.8)$$

where $A(\underline{\mu})$ is in the sigma algebra generated by the r.v.'s $[X_i, Y_i]$ ($1 \leq i \leq k$).

Now consider a vector $\underline{\mu}' = (\mu'_1, \dots, \mu'_k)$ where $\mu'_{[i]} = \mu_{[i]}$ ($1 \leq i \leq k-1$) and $\mu'_{[k]} > \mu_{[k]}$. Then $P_{\underline{\mu}}, \{CS|P_2\} = P\{A(\underline{\mu}')\}$. We shall show that $A(\underline{\mu}') \supseteq A(\underline{\mu})$. We denote the value taken on by a r.v. X at a sample point ω by $X(\omega)$. Also let $\delta'_{ij} = \mu'_{[i]} - \mu'_{[j]}$ ($1 \leq i, j \leq k$). Then $\delta'_{ki} > \delta_{ki}$ ($1 \leq i \leq k-1$) and $\delta'_{ij} = \delta_{ij}$ ($1 \leq i, j \leq k-1$).

Fix $\omega \in A(\underline{\mu})$ which corresponds to some set $s \in S$.

Case 1: Suppose that ω belongs to the following event:

$$\{X_k(\omega) + \delta_{ki}\sqrt{n_1}/\sigma > X_i(\omega) > X_k(\omega) + (\delta_{ki}-h)\sqrt{n_1}/\sigma \quad \forall i \in s,$$

$$X_k(\omega) + (\delta_{ki}-h)\sqrt{n_1}/\sigma > X_i(\omega) \quad \forall i \notin s\}.$$

Then it also belongs to the event

$$\{X_k(\omega) + \delta'_{ki}\sqrt{n_1}/\sigma > X_i(\omega) > X_k(\omega) + (\delta'_{ki}-h)\sqrt{n_1}/\sigma \quad \forall i \in s',$$

$$X_k(\omega) + (\delta'_{ki}-h)\sqrt{n_1}/\sigma > X_i(\omega) \quad \forall i \notin s'\}$$

for some set $s' \in S$, $s' \subseteq s$.

Case 2: Suppose that s is non-empty and for some $j \in s$, ω belongs to the following event:

$$\{X_j(\omega) - \delta_{kj}\sqrt{n_1}/\sigma > X_k(\omega) > X_j(\omega) - (\delta_{kj}+h)\sqrt{n_1}/\sigma, \quad (5.9)$$

$$X_j(\omega) - \delta_{ij}\sqrt{n_1}/\sigma > X_i(\omega) > X_j(\omega) - (\delta_{ij}+h)\sqrt{n_1}/\sigma \quad \forall i \in s, i \neq j;$$

$$X_j(\omega) - (\delta_{ij}+h)\sqrt{n_1}/\sigma > X_i(\omega) \quad \forall i \notin s\}.$$

Now suppose that δ_{ij} in (5.9) are replaced by δ'_{ij} and that $X_j(\omega) - \delta'_{kj}\sqrt{n_1}/\sigma > X_k(\omega)$ is still satisfied. Then (5.9) holds with δ_{ij} replaced by δ'_{ij} ($1 \leq i, j \leq k$) and $s = s'$. On the other hand, if $X_j(\omega) - \delta'_{kj}\sqrt{n_1}/\sigma > X_k(\omega)$ is violated, then ω must belong to the following event:

$$\{X_k(\omega) + \delta'_{kj}\sqrt{n_1}/\sigma > X_i(\omega) > X_k(\omega) + (\delta'_{ki}-h)\sqrt{n_1}/\sigma \quad \forall i \in s',$$

$$X_k(\omega) + (\delta'_{ki}-h)\sqrt{n_1}/\sigma > X_i(\omega) \quad \forall i \notin s'\}$$

for some $s' \subseteq s$.

From Cases 1 and 2 we have $s' \subseteq s$, and hence we obtain

$$\{Y_k(\omega) + \delta_{ki}\sqrt{n_1}/\sigma > Y_i(\omega) \quad \forall i \in s\}$$

$$\Rightarrow \{Y_k(\omega) + \delta'_{ki}\sqrt{n_1}/\sigma > Y_i(\omega) \quad \forall i \in s'\}.$$

Therefore $\omega \in A(\underline{\mu}) \Rightarrow \omega \in A(\underline{\mu}')$ and $A(\underline{\mu}) \subseteq A(\underline{\mu}')$. Hence $P\{A(\underline{\mu}')\} \geq P\{A(\underline{\mu})\}$ and $P_{\underline{\mu}'}\{CS|P_2\} \geq P_{\underline{\mu}}\{CS|P_2\}$ which completes the proof of the lemma.

Corollary 5.2: $P_{\underline{\mu}(\delta)}\{CS|P_2\}$ is non-decreasing in $\delta \geq 0$ when $\mu_{[1]} = \mu_{[k-1]}$ is fixed. In particular, for $k = 2$ $P_{\underline{\mu}}\{CS|P_2\}$ achieves its infimum over $\Omega(\delta^*)$, i.e., satisfies (5.7), at any $\underline{\mu}$ satisfying $\mu_{[2]} - \mu_{[1]} = \delta^*$.

Proof: Since P_2 is translation invariant, $P_{\underline{\mu}}\{CS|P_2\}$ depends on $\underline{\mu}$ only through the δ_{ki} ($1 \leq i \leq k-1$). The result then follows from Lemma 5.1.

Remark 5.2: The method of proof used for Lemma 5.1 does not carry over to prove our intuitive conjecture that $P_{\underline{\mu}}\{CS|P_2\}$ is non-increasing in each $\mu_{[i]}$ ($i \neq k$) when the remaining μ -values remain fixed; nor have we been successful

in finding alternative methods of proof. However, Monte Carlo samplings that we have conducted have supported this conjecture. Nonetheless, the monotonicity of $P_{\underline{\mu}}\{CS|P_2\}$ in the δ_{ki} ($1 \leq i \leq k-1$) for $k > 2$ still remains an open question. We believe that the following conjecture made by Alam [1970] is correct:

Conjecture 5.1: For fixed $k > 2$ and (n_1, n_2, h) , the slippage configuration $\underline{\mu}(\delta^*)$ is a LF-configuration for P_2 .

5.3 Lower bound for $P_{\underline{\mu}}\{CS|P_2\}$

In this section we derive a lower bound for $P_{\underline{\mu}}\{CS|P_2\}$. This lower bound will prove to be particularly useful for $k > 2$ since we will prove that it achieves its infimum over $\Omega(\delta^*)$ at $\underline{\mu}(\delta^*)$, the conjectured LF-configuration for P_2 ; this result will permit us to construct a conservative 2-stage procedure (for $k > 2$) which will guarantee (2.1). The lower bound involves integrals the values of which can be easily calculated on a digital computer.

Theorem 5.2: For any $\underline{\mu} \in \Omega$ we have

$$P_{\underline{\mu}}\{CS|P_2\} \geq \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^{k-1} \Phi[x + (\delta_{ki} + h)\sqrt{n_1}/\sigma] + \prod_{i=1}^{k-1} \Phi[x + \delta_{ki}\sqrt{m}/\sigma] \right\} d\Phi(x) - 1. \quad (5.10)$$

Proof:

$$\begin{aligned} 1 - P_{\underline{\mu}}\{CS|P_2\} &= P_{\underline{\mu}}\{\text{Incorrect selection}|P_2\} \\ &\leq P_{\underline{\mu}}\{\bar{X}_{(k)}^{(1)} < \bar{X}_{(i)}^{(1)} - h \text{ for some } i \neq k\} + P_{\underline{\mu}}\{\bar{X}_{(k)} < \bar{X}_{(i)} \text{ for some } i \neq k\} \\ &= 1 - P_{\underline{\mu}}\{\bar{X}_{(k)}^{(1)} \geq \bar{X}_{(i)}^{(1)} - h \text{ } \forall i \neq k\} + 1 - P_{\underline{\mu}}\{\bar{X}_{(k)} \geq \bar{X}_{(i)} \text{ } \forall i \neq k\} \\ &= 2 - \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + (\delta_{ki} + h)\sqrt{n_1}/\sigma] d\Phi(x) - \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + \delta_{ki}\sqrt{m}/\sigma] d\Phi(x). \end{aligned}$$

A rearrangement of the terms gives the desired lower bound.

Corollary 5.3: For all $\underline{\mu} \in \Omega(\delta^*)$ we have

$$P_{\underline{\mu}}\{CS|P_2\} \geq \int_{-\infty}^{\infty} \{\Phi^{k-1}[x + (\delta^*+h)\sqrt{n_1}/\sigma] + \Phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]\}d\Phi(x) - 1. \quad (5.11)$$

Proof: The proof follows immediately on noting that the r.h.s. of (5.10) is non-decreasing in each δ_{ki} for $1 \leq i \leq k-1$.

Corollary 5.4: Since the r.h.s. of (5.11) is strictly increasing in each of n_1 , m , h , and $\rightarrow 1$ as n_1 or as n_2 and $h \rightarrow \infty$, we see that (2.1) can be guaranteed if all are chosen sufficiently large.

As a consequence of Corollary 5.4 it is clear that a conservative two-stage procedure which guarantees (2.1) and which employs either the R-minimax or the U-minimax design criterion can be constructed and implemented using the lower bound given by the r.h.s. of (5.11). We shall denote such procedures which employ these criteria by $P_2(R_C)$ and $P_2(U_C)$, respectively. $P_2(R_C)$ is conservative relative to $P_2(R_E)$ (as is $P_2(U_C)$ relative to $P_2(U_E)$) since it overprotects the experimenter with respect to (2.1), this overprotection being purchased at the expense of an increase in $E_{\underline{\mu}}\{T|P_2(R_C)\}$ and $E_{\underline{\mu}}\{T|P_2(U_C)\}$ at $\underline{\mu}(\delta^*)$ and $\underline{\mu}(0)$, respectively. We consider $P_2(R_C)$ and $P_2(U_C)$ in detail in Sections 7-9.

Remark 5.3: If we let $h \rightarrow \infty$ on the r.h.s. of (5.10) we obtain

$\int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + \delta_{ki}\sqrt{m}/\sigma]d\Phi(x)$ which is an expression for $P_{\underline{\mu}}\{CS|P_1\}$ where P_1 uses a common single-stage sample size m per population. Thus P_1 is a special case of any $P_2(\cdot)$ based on the conservative lower bound.

6. Expected total sample size for P_2

In order to solve either (4.4) or (4.5) we first find an analytical expression for $E_{\underline{\mu}}\{T|P_2\}$; this is done in Section 6.1. Then it is necessary to determine $\text{Sup } E_{\underline{\mu}}\{T|P_2\}$ for $\underline{\mu} \in \Omega(\delta^*)$ and for $\underline{\mu} \in \Omega$ for (4.4) and (4.5), respectively; the sets of μ -values at which these suprema occur are found in Section 6.2.

6.1 General expression for $E_{\underline{\mu}}\{T|P_2\}$

Our result concerning a general expression for $E_{\underline{\mu}}\{T|P_2\}$ is summarized in the following theorem:

Theorem 6.1: For any $\underline{\mu} \in \Omega$ we have

$$E_{\underline{\mu}}\{T|P_2\} = \quad (6.1)$$

$$kn_1 + n_2 \sum_{i=1}^k \int_{-\infty}^{\infty} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^k \phi[x + (\delta_{ij} + h)\sqrt{n_1}/\sigma] - \prod_{\substack{j=1 \\ j \neq i}}^k \phi[x + (\delta_{ij} - h)\sqrt{n_1}/\sigma] \right\} d\phi(x).$$

Proof:

$$\begin{aligned} E_{\underline{\mu}}\{T|P_2\} &= kn_1 + n_2 E_{\underline{\mu}}\{S|P_2\} \\ &= kn_1 + n_2 [E_{\underline{\mu}}\{S'|P_2\} - P_{\underline{\mu}}\{S' = 1|P_2\}] \\ &= kn_1 + n_2 \left[\sum_{i=1}^k P_{\underline{\mu}}\{\bar{X}_{(i)}^{(1)} \geq \bar{X}_{(j)}^{(1)} - h \mid j \neq i\} \right. \\ &\quad \left. - \sum_{i=1}^k P_{\underline{\mu}}\{\bar{X}_{(i)}^{(1)} \geq \bar{X}_{(j)}^{(1)} + h \mid j \neq i\} \right]. \end{aligned} \quad (6.2)$$

Theorem 6.1 follows immediately.

6.2 The supremum of $E_{\underline{\mu}}\{T|P_2\}$

Our results concerning the supremum of $E_{\underline{\mu}}\{T|P_2\}$ for $\underline{\mu} \in \Omega(\delta^*)$ and $\underline{\mu} \in \Omega$ are summarized in a) and b) of the following theorem.

Theorem 6.2: For fixed k and (n_1, n_2, h) we have that

$$\begin{aligned}
 \text{a) } \sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}\{T|P_2\} &= kn_1 \\
 &+ n_2 \left[\int_{-\infty}^{\infty} \{ \phi^{k-1}[x + (\delta^*+h)\sqrt{n_1}/\sigma] - \phi^{k-1}[x + (\delta^*-h)\sqrt{n_1}/\sigma] \} d\phi(x) \right. \\
 &+ (k-1) \int_{-\infty}^{\infty} \{ \phi^{k-2}[x + h\sqrt{n_1}/\sigma] \phi[x - (\delta^*-h)\sqrt{n_1}/\sigma] \\
 &\quad \left. - \phi^{k-2}[x - h\sqrt{n_1}/\sigma] \phi[x - (\delta^*+h)\sqrt{n_1}/\sigma] \} d\phi(x) \right] \quad (6.3)
 \end{aligned}$$

which occurs when $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$.

$$\text{b) } \sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}}\{T|P_2\} = kn_1 + kn_2 \int_{-\infty}^{\infty} \{ \phi^{k-1}[x + h\sqrt{n_1}/\sigma] - \phi^{k-1}[x - h\sqrt{n_1}/\sigma] \} d\phi(x) \quad (6.4)$$

which occurs when $\mu_{[1]} = \mu_{[k]}$.

We shall prove part b) of Theorem 6.2; the proof of part a) follows along the same lines.

Proof: Gupta [1965] has shown that $E_{\underline{\mu}}\{S'|P_2\}$ achieves its supremum for $\underline{\mu} \in \Omega$ when $\mu_{[1]} = \mu_{[k]}$. It only remains to show that $P_{\underline{\mu}}\{S' = 1|P_2\}$ achieves its infimum when $\mu_{[1]} = \mu_{[k]}$. We use the method of Gupta [1965].

Set $\mu_{[1]} = \dots = \mu_{[b]} = \mu < \mu_{[b+1]}$ for some b ($1 \leq b \leq k-1$) and define $\delta_i = \mu_{[i]} - \mu$ for $b+1 \leq i \leq k$. Define

$$\begin{aligned}
Q(\mu) &= P_{\underline{\mu}}\{S' = 1 | P_2; \mu_{[1]} = \dots = \mu_{[b]} = \mu\} \\
&= b \int_{-\infty}^{\infty} \phi^{b-1}[x - h\sqrt{n_1}/\sigma] \prod_{j=b+1}^k \phi[x - (\delta_j + h)\sqrt{n_1}/\sigma] d\phi(x) \\
&\quad + \sum_{i=b+1}^k \int_{-\infty}^{\infty} \phi^b[x + (\delta_i - h)\sqrt{n_1}/\sigma] \prod_{\substack{j=b+1 \\ j \neq i}}^k \phi[x + (\delta_{ij} - h)\sqrt{n_1}/\sigma] d\phi(x).
\end{aligned}$$

After differentiating w.r.t. μ , and then interchanging the order of integration and summation in the first term, and making appropriate substitutions, we obtain

$$\frac{dQ}{d\mu} = \frac{b\sqrt{n_1}}{\sigma} \sum_{i=b+1}^k \int_{-\infty}^{\infty} \phi^{b-1}[x + (\delta_i - h)\sqrt{n_1}/\sigma] \prod_{\substack{j=b+1 \\ j \neq i}}^k \phi[x + (\delta_{ij} - h)\sqrt{n_1}/\sigma] \quad (6.5)$$

$$\cdot \{\phi[x - h\sqrt{n_1}/\sigma]\phi[x + \delta_i\sqrt{n_1}/\sigma] - \phi[x + (\delta_i - h)\sqrt{n_1}/\sigma]\phi(x)\} dx \leq 0.$$

The last inequality is obtained by noting that the quantity inside $\{ \}$ in (6.5) is non-positive for every x and i for $b+1 \leq i \leq k$ due to the monotone likelihood ratio property of ϕ . It follows that Q is non-increasing in μ and is in fact strictly decreasing if $h, n_1 > 0$. Thus subject to $\mu_{[1]} = \dots = \mu_{[b]}$, we see that $P_{\underline{\mu}}\{S' = 1 | P_2\}$ is minimized by increasing the common value μ until $\mu = \mu_{[b+1]}$. Since this is true for each $b \leq k-1$, it follows that $P_{\underline{\mu}}\{S' = 1 | P_2\}$ is minimized and hence $E_{\underline{\mu}}\{T | P_2\}$ is maximized over Ω when $\mu_{[1]} = \mu_{[k]}$.

Using the results of Theorem 5.1 and Lemma 5.1 along with Theorem 6.2 we can now formulate our optimization problems (4.4) and (4.5) precisely.

7. Optimization problems yielding conservative solutions

In this section we consider the optimization problems (4.4) and (4.5) which one must solve in order to determine $(n_1, n_2, h | R_E)$ and $(n_1, n_2, h | U_E)$ which are necessary to implement $P_2(R_E)$ and $P_2(U_E)$. As noted in Section 5.2, we have not been successful in determining the LF-configuration of the $\mu[i]$ ($1 \leq i \leq k$), except for $k = 2$. Thus for $k > 2$ we replace the exact probability $\inf_{\mu \in \Omega(\delta^*)} P\{CS | P_2\}$ by the conservative lower bound on that probability given by the r.h.s. of (5.11), and consider the following optimization problems:

7.1 Discrete optimization problems

7.1.1 R-minimax design criterion

For given k and specified $\{\delta^*, P^*\}$ choose (n_1, n_2, h) to

$$\begin{aligned} &\text{minimize } \sup_{\mu \in \Omega(\delta^*)} E_{\mu}\{T | P_2\} \\ &\text{subject to } \int_{-\infty}^{\infty} \{\phi^{k-1}[x + (\delta^* + h)\sqrt{n_1}/\sigma] + \phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]\} d\phi(x) - 1 \geq P^*, \end{aligned} \quad (7.1)$$

where n_1, n_2 are non-negative integers and $h \geq 0$.

In (7.1), $\sup_{\mu \in \Omega(\delta^*)} E_{\mu}\{T | P_2\}$ is given by (6.3). We denote by $(n_1, n_2, h | R_C)$ the solution to (7.1), and regard it as a conservative solution to (4.4); we denote the corresponding procedure by $P_2(R_C)$.

7.1.2 U-minimax design criterion

For given k and specified $\{\delta^*, P^*\}$ choose (n_1, n_2, h) to

$$\begin{aligned}
 & \text{minimize } \sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}} \{T|P_2\} \\
 & \text{subject to } \int_{-\infty}^{\infty} \{\phi^{k-1}[x + (\delta^* + h)\sqrt{n_1}/\sigma] + \phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]\} d\phi(x) - 1 \geq P^*,
 \end{aligned} \tag{7.2}$$

where n_1, n_2 are non-negative integers and $h \geq 0$.

In (7.2), $\sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}} \{T|P_2\}$ is given by (6.4). We denote by $(n_1, n_2, h|U_C)$ the solution to (7.2), and regard it as a conservative solution to (4.5); we denote the corresponding procedure by $P_2(U_C)$.

7.2 Continuous optimization problems

The problems (7.1) and (7.2) are extremely complicated integer programming problems with non-linear constraints and objective functions. Although these problems can be solved in principle by enumeration, the search is likely to be a costly one because of the numerical evaluation of the integrals involved. Additionally, since the solution depends on δ^* , a separate solution is required not only for each k and P^* -value, but also for each δ^* . Hence we shall remove the restriction that n_1, n_2 must be integers; we reparametrize the problem and regard the new design constants (which are functions of n_1, n_2 , and h) as continuous. We use this continuous version as a large sample approximation to the discrete version.

We define the new design constants

$$c_1 = \delta^*\sqrt{n_1}/\sigma, \quad c_2 = \delta^*\sqrt{n_2}/\sigma, \quad d = h\sqrt{n_1}/\sigma. \tag{7.3}$$

We note that the exact expression for $P_{\underline{\mu}}\{CS|P_2\}$ and the conservative lower bound on it, as well as $E_{\underline{\mu}}\{T|P_2\}$ depend on (n_1, n_2, h) , δ^* , σ only through (c_1, c_2, d) and δ_{ki}/δ^* ($1 \leq i \leq k-1$).

Thus, for example, for given k and specified $\{\delta^*, P^*\}$ we can approximate the design constants $(n_1, n_2, h|U_C)$ necessary to implement $P_2(U_C)$ by solving the continuous optimization problem:

$$\text{minimize } kc_1^2 + kc_2^2 \int_{-\infty}^{\infty} \{\phi^{k-1}(x+d) - \phi^{k-1}(x-d)\} d\phi(x) \quad (7.4)$$

$$\text{subject to } \int_{-\infty}^{\infty} \{\phi^{k-1}(x+c_1+d) + \phi^{k-1}(x + (c_1^2+c_2^2)^{1/2})\} d\phi(x) - 1 \geq P^*$$

where $c_1, c_2, d \geq 0$.

We denote by $(\hat{c}_1, \hat{c}_2, \hat{d}|U_C)$ the solution to (7.4), and use the approximate design constants

$$\hat{n}_1 = \left[\left(\frac{\hat{c}_1^{\sigma}}{\delta^*} \right)^2 + 1 \right], \quad \hat{n}_2 = \left[\left(\frac{\hat{c}_2^{\sigma}}{\delta^*} \right)^2 + 1 \right], \quad \hat{h} = \frac{\hat{d}\delta^*}{\hat{c}_1} \quad (7.5)$$

where $[z]$ denotes the greatest integer $< z$, to implement $P_2(U_C)$.

Similarly, for $k = 2$ and specified $\{\delta^*, P^*\}$ we can approximate the design constants $(n_1, n_2, h|U_E)$ necessary to implement $P_2(U_E)$ by solving the continuous optimization problem:

$$\text{minimize } 2c_1^2 + 2c_2^2 \{\phi(d/\sqrt{2}) - \phi(-d/\sqrt{2})\} \quad (7.6)$$

$$\text{subject to } \phi[(c_1-d)/\sqrt{2}] + \int_{(c_1-d)/\sqrt{2}}^{(c_1+d)/\sqrt{2}} \phi[-x\sqrt{p/q} + \sqrt{(c_1^2+c_2^2)/2q}] d\phi(x) \geq P^*,$$

where $c_1, c_2, d \geq 0$.

We denote the solution by $(\hat{c}_1, \hat{c}_2, \hat{d}|U_E)$. Analogous expressions can be written

in order to approximate the design constants $(n_1, n_2, h|R_C)$ for $k \geq 2$, and $(n_1, n_2, h|R_E)$ for $k = 2$.

8. Constants to implement P_2

8.1 Constants to implement $P_2(U_E)$ and $P_2(R_E)$ for $k = 2$

Table 1 contains constants necessary to approximate $(n_1, n_2, h|U_E)$ and $(n_1, n_2, h|R_E)$ for $k = 2$ and selected P^* ; although we are primarily interested in the ones associated with $P_2(U_E)$, we have computed those associated with $P_2(R_E)$ for comparative purposes. (See Section 9.) The computations for $P_2(U_E)$ are the solutions of (7.5), while those for $P_2(R_E)$ are the solutions of the analogous problem wherein $\sup_{\underline{\mu}} E\{T|P_2\}$ over $\Omega(\delta^*)$ is minimized. The constants given here for $P_2(U_E)$ and $P_2(R_E)$ are exact since the LF-configuration for $P_{\underline{\mu}}\{CS\}$ is known for $k = 2$.

8.2 Constants to implement $P_2(U_C)$ for $k \geq 3$

Table 2 contains constants necessary to approximate $(n_1, n_2, h|U_C)$ for $k = 3, 4, 5, 10, 15, 25$ and selected P^* ; the computations for $P_2(U_C)$ are the solution of (7.4). The constants given here for $P_2(U_C)$ are conservative since the LF-configuration for $P_{\underline{\mu}}\{CS|P_2\}$ is unknown for $k \geq 3$. (We have not attempted to compute the constants $(\hat{c}_1, \hat{c}_2, \hat{d}|U_E)$ which would be used if the conjectured LF-configuration for $P_{\underline{\mu}}\{CS|P_2\}$ were indeed $\mu[k] - \mu[i] = \delta^*$ ($1 \leq i \leq k-1$) for $k \geq 3$; such computations, although of interest, would be difficult to carry out because it would be necessary to evaluate numerically very complicated iterated integrals.)

All of the computations for Tables 1 and 2 (as well as those described in Section 9) were carried out in double precision arithmetic on either Cornell's IBM 360/65 and IBM 370/168 or on Northwestern's CDC 6400. To solve the

Table 1

Constants to implement $P_2(U_E)$ and $P_2(R_E)$ for $k = 2$

p^*	$(\hat{c}_1, \hat{c}_2, \hat{d} U_E)$			$(\hat{c}_1, \hat{c}_2, \hat{d} R_E)$		
	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$
0.9999	4.5397	2.9087	0.97215	3.4801	4.4120	1.9162
0.9995	3.9742	2.6708	0.95623	3.1239	3.9034	1.6992
0.999	3.7062	2.5712	0.94824	2.9566	3.6506	1.6026
0.99	2.7189	2.0906	0.91913	2.2931	2.7371	1.2803
0.95	1.8621	1.6156	0.88072	1.6583	1.9347	1.0574
0.90	1.4270	1.3132	0.85278	1.3224	1.4996	0.92974
0.85	1.1391	1.0930	0.84174	1.0789	1.1890	0.90951
0.80	0.91577	0.90970	0.82702	0.88255	0.96174	0.87132
0.75	0.72801	0.74161	0.81999	0.71036	0.77072	0.84413
0.70	0.56240	0.58661	0.80783	0.55468	0.59769	0.82505
0.65	0.41227	0.43299	0.80202	0.40845	0.43957	0.81002
0.60	0.26982	0.28775	0.79714	0.26907	0.28914	0.79865
0.55	0.13374	0.14322	0.79140	0.13378	0.14318	0.79053

Table 2

Constants to implement $P_2(U_C)$ for $k \geq 3$

k	p*	$(\hat{c}_1, \hat{c}_2, \hat{d} U_C)$		
		$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$
3	0.99	2.9326	2.4083	1.2458
	0.95	2.0893	1.8974	1.6303
	0.90	1.6699	1.5491	2.1814
	0.75	1.0492	0.97980	3.9335
4	0.99	3.0432	2.5805	1.2596
	0.95	2.2400	2.1106	1.4806
	0.90	1.8262	1.7859	1.8245
	0.75	1.2203	1.1712	3.2365
5	0.99	3.1035	2.7301	1.2712
	0.95	2.3184	2.2622	1.4604
	0.90	1.9209	1.9786	1.6403
	0.75	1.3191	1.3304	2.7280
	0.60	0.96047	0.91856	4.3038
10	0.99	3.2364	3.1620	1.3453
	0.95	2.5094	2.7750	1.3529
	0.90	2.1466	2.5349	1.3830
	0.75	1.5712	1.9725	1.6980
	0.60	1.2248	1.3491	2.7648
	0.45	0.95459	0.92770	4.3433
15	0.99	3.2983	3.3883	1.3999
	0.95	2.5886	3.0259	1.3771
	0.90	2.2344	2.8212	1.3676
	0.75	1.6899	2.4268	1.3974
	0.60	1.3404	1.7600	2.0210
	0.45	1.0897	1.1270	3.5471
25	0.99	3.3634	3.6572	1.4783
	0.95	2.6646	3.3204	1.4401
	0.90	2.3270	3.1393	1.3972
	0.75	1.8000	2.8539	1.3234
	0.60	1.4909	2.7798	1.1769
	0.45	1.3026	3.1516	0.76886

continuous optimization problems, first a "reasonably good" discrete optimal solution was found by a search method. This solution was used as an initial guess in the computer program using a modified version of the steepest descent method to solve the continuous non-linear programming problem; Algorithm 304 of Hill and Joyce [1967] was used to evaluate $\Phi(\cdot)$; the integrals were evaluated using the Romberg method of integration. We do not claim that our solutions represent the absolute optima, but they are reasonably close to the optima. (The $E_{\underline{\mu}}\{T|P_2(U_C)\}$ -surface is very flat in the neighborhood of the maximum for $P^* \rightarrow 1/k$ since $P_2(U_C) \rightarrow P_1$.) The tabulated values should be correct to the number of significant figures given.

9. The performance of P_2 relative to P_1

As a measure of the efficiency of P_1 (Bechhofer [1954]) relative to that of P_2 when both guarantee the same probability requirement (2.1), we consider the ratio (termed relative efficiency (RE)) $E_{\underline{\mu}}\{T|P_2\}/kn$ where $n = [(\hat{\sigma}/\delta^*)^2 + 1]$, and $\hat{c}$ is the solution of

$$\int_{-\infty}^{\infty} \Phi^{k-1}(x+\hat{c})d\Phi(x) = P^*. \quad (9.1)$$

Clearly RE depends on $\underline{\mu}$ and $\{\delta^*, P^*\}$; values of RE less than unity favor P_2 over P_1 . For mathematical convenience we shall use the continuous approximations to $E_{\underline{\mu}}\{T|P_2\}$ and n (thereby ignoring the fact that the sample sizes must be integers). RE is then given by

$$\left[kc_1^2 + c_2^2 \sum_{i=1}^k \int_{-\infty}^{\infty} \left\{ \prod_{\substack{j=1 \\ j \neq i}}^k \Phi(x+d+\delta_{ij}c_1/\delta^*) - \prod_{\substack{j=1 \\ j \neq i}}^k \Phi(x-d+\delta_{ij}c_1/\delta^*) \right\} d\Phi(x) \right] / kc^2 \quad (9.2)$$

where we employ in (9.2) either $(\hat{c}_1, \hat{c}_2, \hat{d} | P_2(U_E))$ for $k = 2$ or $(\hat{c}_1, \hat{c}_2, \hat{d} | P_2(U_C))$ for $k \geq 3$. (In order to compare the performance of $P_2(U_C)$ with that of $P_2(U_E)$ for $k = 2$, we also will employ $(\hat{c}_1, \hat{c}_2, \hat{d} | P_2(U_C))$ for $k = 2$. See Table 3.) The value of $\hat{c}$ in (9.1) has been tabulated for selected k and P^* by Bechhofer [1954], Gupta [1963], and Milton [1963] (Bechhofer's $\lambda = \hat{c}$, Gupta's and Milton's $H = \hat{c}/\sqrt{2}$.)

Remark 9.1: For the equal means (EM) configuration $\mu_{[1]} = \mu_{[k]}$, and for the $\underline{\mu}(\delta^*)$ configuration $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ (known to be LF for $P_2(U_E)$ for $k = 2$ and for $P_2(U_C)$ for $k \geq 2$, and conjectured to be LF for $P_2(U_E)$ for $k > 3$), we note that RE depends only on k and P^* for given $P_2(\cdot)$ and $\underline{\mu} \in \Omega$.

Table 3 which concerns $P_2(U_E)$ and $P_2(U_C)$ for $k = 2$, and Table 4 which concerns $P_2(U_C)$ for $k \geq 3$, give computed RE-values for selected P^* and $\underline{\mu} \in \Omega$ to indicate the magnitude of the saving in $E_{\underline{\mu}}\{T|P\}$ achieved by the screening property of $P_2(U_C)$ ($P_2(U_E)$) when $P_2(\cdot)$ is used in place of P_1 for $k \geq 2$ ($k = 2$); the computations for Tables 3 and 4 are based on $(\hat{c}_1, \hat{c}_2, \hat{d})$ listed in Tables 1 and 2, respectively.

9.1 $P_2(U_E)$ and $P_2(U_C)$ vs. P_1 for $k = 2$

For all P^* we note that $RE_{\underline{\mu}(0)}$ is less for $P_2(U_E)$ than for $P_2(U_C)$, but $RE_{\underline{\mu}(\infty)}$ is greater for $P_2(U_E)$ than for $P_2(U_C)$ (since $\hat{n}_1$ for $P_2(U_E)$ turns out to be greater than $\hat{n}_1$ for $P_2(U_C)$). The range of $(\mu_{[2]} - \mu_{[1]})/\delta^* = \delta/\delta^*$ values over which $RE_{\underline{\mu}(\delta)}$ is less for $P_2(U_E)$ than for $P_2(U_C)$ for given P^* appears to depend critically on P^* being small for P^* close to unity and large for $P^* \rightarrow 1/2$ (since in this latter situation $P_2(U_C) \rightarrow P_1$). However, of greatest importance, is the fact that for either $P_2(U_E)$ or $P_2(U_C)$ used at any P^* ($1/2 < P^* < 1$) we have

Table 3
Efficiency of P_1 relative to $P_2(u_E)$ and $P_2(u_C)$ for $k = 2$ when $\mu_{[2]} - \mu_{[1]} = \delta$

P_1^*	$\delta = 0$ (EM)		$\delta = 0.5\delta^*$		$\delta = \delta^*$ (LF)		$\delta = 2\delta^*$		$\delta = 4\delta^*$		$\delta = \infty$	
	$P_2(u_E)$	$P_2(u_C)$	$P_2(u_E)$	$P_2(u_C)$	$P_2(u_E)$	$P_2(u_C)$	$P_2(u_E)$	$P_2(u_C)$	$P_2(u_E)$	$P_2(u_C)$	$P_2(u_E)$	$P_2(u_C)$
0.9999	0.9005	0.9190	0.7966	0.7951	0.7468	0.7285	0.7450	0.7256	0.7450	0.7256	0.7450	0.7256
0.9995	0.8944	0.9202	0.8000	0.8046	0.7347	0.7147	0.7294	0.7055	0.7294	0.7055	0.7294	0.7055
0.999	0.8914	0.9216	0.8013	0.8109	0.7279	0.7089	0.7192	0.6937	0.7192	0.6937	0.7192	0.6937
0.99	0.8785	0.9370	0.8139	0.8536	0.7220	0.7164	0.6833	0.6369	0.6830	0.6356	0.6830	0.6356
0.95	0.8654	0.9725	0.8268	0.9312	0.7457	0.8271	0.6512	0.6223	0.6408	0.5591	0.6408	0.5590
0.90	0.8580	0.9925	0.8326	0.9778	0.7716	0.9362	0.6588	0.7709	0.6201	0.5364	0.6199	0.5206
0.85	0.8532		0.8359		0.7908		0.6824		0.5063		0.6039	
0.80	0.8498		0.8381		0.8059		0.7139		0.6047		0.5920	
0.75	0.8472		0.8395		0.8176		0.7473		0.6220		0.5825	
0.70	0.8452		0.8408		0.8270		0.7787		0.6619		0.5751	
0.65	0.8436		0.8410		0.8334		0.8052		0.7196		0.5724	
0.60	0.8425		0.8414		0.8380		0.8251		0.7791		0.5671	
0.55	0.8419		0.8417		0.8409		0.8376		0.8248		0.5664	

Table 4

Efficiency of P_1 relative to $P_2(U_C)$ for $k \geq 3$
 when the $\mu_{[i]}$ ($1 \leq i \leq k$) are in various configurations

k	F*	(EM) $\mu_{[k]} = \mu_{[1]}$	$\mu_{[k]} - \mu_{[k-1]} = \delta^*$, $\mu_{[i]} - \mu_{[i-1]} = \delta$ ($2 \leq i \leq k-1$)					$\mu_{[k]} - \mu_{[k-1]} = \infty$
			(LF)					
			$\delta/\delta^*=0$	$\delta/\delta^*=0.5$	$\delta/\delta^*=1.0$	$\delta/\delta^*=2.0$	$\delta/\delta^*=4.0$	
3	0.99	0.9328	0.7184	0.6938	0.6913	0.6912	0.6912	0.6572
	0.95	0.9624	0.7965	0.7430	0.7206	0.7148	0.7147	0.5943
	0.90	0.9838	0.8887	0.8394	0.7990	0.7684	0.7659	0.5606
	0.75	0.9999	0.9947	0.9899	0.9808	0.9462	0.8603	0.5354
4	0.99	0.9139	0.7020	0.6676	0.6661	0.6660	0.6660	0.6424
	0.95	0.9392	0.7650	0.6832	0.6685	0.6663	0.6663	0.5900
	0.90	0.9631	0.8389	0.7377	0.6995	0.6866	0.6862	0.5549
	0.75	0.9967	0.9772	0.9380	0.8768	0.7926	0.7511	0.5262
5	0.99	0.8954	0.6866	0.6468	0.6457	0.6457	0.6457	0.6269
	0.95	0.9165	0.7431	0.6465	0.6361	0.6347	0.6347	0.5759
	0.90	0.9403	0.8040	0.6755	0.6493	0.6424	0.6422	0.5460
	0.75	0.9888	0.9473	0.8433	0.7573	0.7014	0.6847	0.5104
	0.60	0.9984	0.9959	0.9818	0.9360	0.8211	0.7405	0.5225
10	0.99	0.8300	0.6409	0.5916	0.5911	0.5911	0.5911	0.5811
	0.95	0.8366	0.6778	0.5698	0.5661	0.5658	0.5658	0.5390
	0.90	0.8518	0.7174	0.5695	0.5610	0.5595	0.5595	0.5179
	0.75	0.9171	0.8323	0.6032	0.5740	0.5624	0.5616	0.4818
	0.60	0.9306	0.9496	0.6968	0.6278	0.5918	0.5789	0.4788
	0.45	0.9994	0.9970	0.8879	0.7497	0.6699	0.6281	0.5160
15	0.99	0.7920	0.6178	0.5662	0.5659	0.5659	0.5659	0.5588
	0.95	0.7906	0.6463	0.5376	0.5353	0.5352	0.5352	0.5169
	0.90	0.7993	0.6769	0.5296	0.5244	0.5236	0.5236	0.4958
	0.75	0.8504	0.7653	0.5423	0.5263	0.5211	0.5209	0.4639
	0.60	0.9389	0.8864	0.5820	0.5477	0.5309	0.5278	0.4568
	0.45	0.9934	0.9831	0.6905	0.6195	0.5826	0.5835	0.4940
25	0.99	0.7472	0.5913	0.5381	0.5379	0.5379	0.5379	0.5333
	0.95	0.7372	0.6109	0.5023	0.5009	0.5008	0.5008	0.4892
	0.90	0.7388	0.6333	0.4914	0.4885	0.4881	0.4881	0.4709
	0.75	0.7683	0.6954	0.4878	0.4790	0.4766	0.4766	0.4449
	0.60	0.8326	0.7769	0.5204	0.5042	0.4980	0.4976	0.4497
	0.45	0.9531	0.9088	0.6306	0.6058	0.5951	0.5940	0.5249

$RE_{\underline{\mu}(\delta)} < 1$ for all $\underline{\mu} \in \Omega$ and $RE_{\underline{\mu}(\infty)} \ll 1$. Thus both $P_2(U_E)$ and $P_2(U_C)$ are highly effective as screening procedures.

9.2 $P_2(U_C)$ vs. P_1 for $k \geq 3$

For given $k \geq 3$ the performance of $P_2(U_C)$ relative to that of P_1 as measured by RE for $1/k < P^* < 1$ and $\underline{\mu} \in \Omega$ is similar to that noted for $k = 2$. In addition, if we regard RE as a function of k for fixed P^* and configuration of the μ_i ($1 \leq i \leq k$), specifically for the configurations $\underline{\mu}(0)$; $\mu[k] - \mu[k-1] = \delta^*$, $\mu[i] - \mu[i-1] = \delta$ ($2 \leq i \leq k-1$, $0 \leq \delta < \infty$); $\mu[k] - \mu[k-1] = \infty$, our computations indicate that RE is decreasing in k (although this has not been established analytically). Thus the effectiveness of $P_2(U_C)$ as a screening procedure appears to be increasing with increasing k .

10. Directions of future research

The most important unsolved problem associated with P_2 is that of determining the LF-configuration of the μ_i (see (5.7)) for $k > 2$; as noted earlier, we conjecture the answer to be the slippage configuration $\underline{\mu}(\delta^*)$. If this conjecture can be shown to be true, it will be necessary to find efficient algorithms for evaluating $P_{\underline{\mu}(\delta^*)}\{CS|P_2\}$ (as given by (5.6)) numerically before the design constants $(\hat{c}_1, \hat{c}_2, \hat{d}|U_E)$ for use with $P_2(U_E)$ can be determined.

A two-stage minimax screening procedure, analogous to P_2 , can be devised for selecting the smallest (say) normal variance; this was done in Tamhane [1975b]. However, design constants to implement the procedure must still be computed.

Procedure P_2 given by (4.1) permanently eliminates populations for which $\bar{X}_i^{(1)} < \bar{X}_{[k]}^{(1)} - h$. However, P_2 can be modified in such a way that populations from which a total of only n_1 observations are taken, are eligible for selection as "best" along with those from which a total of $n_1 + n_2$ observations are taken; in this modification we assert that the population associated with $\max\{\max_{i \in I} \bar{X}_i, \max_{i \notin I} \bar{X}_i^{(1)}\}$ is best. Such a procedure was considered in Tamhane [1975a]; an exact analytical expression for the PCS was derived, and the PCS performance was studied by Monte Carlo sampling methods. Aside from the analytical and computational difficulties (as well as the problem of determining the LF-configuration of the μ_i), procedures of this type would appear to represent a fruitful direction of generalization.

11. Acknowledgments

We are indebted to Professor Thomas J. Santner for a critical reading of a draft of this paper, and for several useful suggestions. We also thank Dr. Yosef Rinott for suggesting the present proof of Lemma 5.1, and Professor Donald Bartel, School of Mechanical and Aerospace Engineering, Cornell University, for providing us with a computer program for solving nonlinear programming problems.

This research was supported by the U.S. Army Research Office-Durham under contracts DAAG29-73-C-0008 and DAAG29-77-C-0003 and by the Office of Naval Research under contract N00014-67-A-0077-0020.

This paper is a revision of a part of the first author's Ph.D. dissertation submitted to Cornell University, 1975. A slightly revised version of this paper will appear in Communications in Statistics.

12. References

- Alam, K. (1970). A two-sample procedure for selecting the population with the largest mean from k normal populations. Ann. Inst. Statist. Math. (Tokyo), 22, 127-136.
- 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. (1960). A note on the limiting relative efficiency of the Wald sequential probability ratio test. J. Amer. Statist. Assoc., 55, 660-663.
- Bechhofer, R.E. (1975). Ranking and selection procedures. Proceedings of the Twentieth Conference on the Design of Experiments in Army Research Development and Testing, Part 2, 929-949.
- Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures, Chicago, Ill.: The University of Chicago Press.
- Bechhofer, R.E. and Sobel, M. (1954). On a sequential ranking procedure (preliminary report). Abstract, Bull. Am. Math. Soc., 60, 34-35.
- Cohen, D.S. (1959). A two sample decision procedure for ranking means of normal populations with a common known variance. Unpublished M.S. thesis, Dept. of Ind. Eng., Cornell Univ., Ithaca, N.Y.
- Eaton, M.L. (1967). Some optimum properties of ranking procedures. Ann. Math. Statist., 38, 124-137.
- Fabian, V. (1974a). Note on Anderson's sequential procedures with triangular boundary. Ann. Statist., 2, 170-176.
- Fabian, V. (1974b). Acknowledgment of Priority to "Note on Anderson's sequential procedures with triangular boundary." Ann. Statist., 2, 1063.
- Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Inst. Statist. Mimeo Ser. No. 150, Inst. Statist., U. of North Carolina, Chapel Hill, N.C.
- Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules. Technometrics, 7, 225-245.
- Hall, W.J. (1959). The most-economical character of some Bechhofer and Sobel decision rules. Ann. Math. Statist., 30, 964-969.
- Hill, I.D. and Joyce, S.A. (1967). Algorithm 304, normal curve integral. Commun. A.C.M., 10, 374.
- Lawing, W.D. and David, H.T. (1966). Likelihood ratio computations of operating characteristics. Ann. Math. Statist., 37, 1704-1716.

- Milton, R.C. (1963). Tables of the equally correlated multivariate normal probability integral. Tech. Rep. No. 27, Dept. of Statistics,, U. of Minnesota, Minneapolis, Minn.
- Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. Ann. Math. Statist., 35, 174-180.
- Ramberg, J. (1966). A comparison of the performance characteristics of two sequential procedures for ranking the means of normal populations. Unpublished M.S. thesis. Tech. Rep. No. 4, Dept. of Industrial Engineering and Operations Research, Cornell Univ., Ithaca, N.Y.
- Tamhane, A.C. (1975a). A minimax multistage elimination type rules for selecting the largest normal mean. Unpublished Ph.D. dissertation. Tech. Rep. No. 259, Dept. of Operations Research, Cornell Univ., Ithaca, N.Y.
- Tamhane, A.C. (1975b). A minimax two-stage permanent elimination type procedure for selecting the smallest normal variance. Tech. Rep. No. 260, Dept. of Operations Research, Cornell Univ., Ithaca, N.Y.
- Wetherill, G.B. and Ofosu, J.B. (1974). Selection of the best of k normal populations. J.R. Statist. Soc. C, 23, 253-277.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR 323	2. GOVT ACCESSION NO.	3. REPORT'S CATALOG NUMBER
4. TITLE (and Subtitle) A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE LARGEST NORMAL MEAN.		5. TYPE OF REPORT & PERIOD COVERED Technical Report.
6. AUTHOR(s) Ajit C. Tamhane and Robert E. Bechhofer		7. PERFORMING ORG. REPORT NUMBER
8. PERFORMING ORGANIZATION NAME AND ADDRESS School of Operations Research and Industrial Engineering, College of Engineering Cornell University, Ithaca, NY 14853		9. CONTRACT OR GRANT NUMBER(s) DAAG29-73-C-0008 DAAG29-77-C-0003 N00014-75-C-0586
10. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activity: U.S. Army Research Office, P.O. Box 12211 Research Triangle Park, North Carolina 27709		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
12. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activity: Statistics and Probability Program Office of Naval Research Arlington, Virginia 22217		13. REPORT DATE January 1977
14. DISTRIBUTION STATEMENT (of this Report) Approved for public release, distribution unlimited.		15. NUMBER OF PAGES 34
16. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		17. SECURITY CLASS. (of this report) Unclassified
18. SUPPLEMENTARY NOTES		19. DECLASSIFICATION/DOWNGRADING SCHEDULE
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Ranking procedure, Selection procedure, Minimax procedure, Two-stage procedure, Indifference-zone approach, Screening		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The problem of selecting the normal population with the largest population mean when the populations have a common known variance is considered. A two-stage procedure is proposed which guarantees the same probability		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE
5/5 0102-014-66011

code 409 369

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont

→ requirement using the indifference-zone approach as does the single-stage procedure of Bechhofer [1954]. The two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than the total number of observations required by the corresponding single-stage procedure, regardless of the configuration of the population means. The saving in expected total number of observations can be substantial, particularly when the configuration of the population means is favorable to the experimenter. The saving is accomplished by screening out "non-contending" populations in the first stage, and concentrating sampling only on "contending" populations in the second stage. The two-stage procedure can be regarded as a composite one which uses a screening subset-type approach (Gupta [1956], [1965]) in the first stage, and an indifference-zone approach (Bechhofer [1954]) applied to all populations retained in the selected subset in the second stage. Constants to implement the procedure for various k and P^* are provided, as are calculations giving the saving in expected total sample size if the two-stage procedure is used in place of the corresponding single-stage procedure.

↑