PURDUE UNIVERSITY

purdue university

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

ISOTONIC PROCEDURES FOR SELECTING POPULATIONS
BETTER THAN A CONTROL UNDER ORDERING PRIOR*

by

Shanti S. Gupta
Purdue University

and

Hwa-Ming Yang
The University of Toledo

Mimeograph Series #81-24

AD-A103891

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #81-24

July 1981

(Revised November 1981)

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Isotonic Procedures for Selecting Populations
Better Than a Control under Ordering Prior**

by

Shanti S. Gupta and Hwa-Ming Yang

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

University of Toledo
Toledo, Ohio, U.S.A.

Abstract*

The problem of selecting a subset containing all populations better than a control under an ordering prior is considered. Three new selection procedures which satisfy a desirable basic requirement on the probability of a correct selection are proposed and studied. Two of the three procedures use the isotonic regression over the sample means of the k-treatments with respect to the given ordering prior. Tables of constants which are necessary to carry out the selection procedures with isotonic approach for the selection of unknown means of normal populations are given. The results including Monte Carlo studies indicate that, in general, the stepwise procedure δ_1 based on isotonic estimators is the best.

*This paper will be presented by Professor Shanti S. Gupta at the Golden Jubilee Conference of the Indian Statistical Institute, Calcutta, December 15-20, 1981.

**Research supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University.

Isotonic Procedures for Selecting Populations Better Than a Control under Ordering Prior*

by

Shanti S. Gupta
Purdue University

and

Hwa-Ming Yang
The University of Toledo

1. Introduction

In this paper, three new selection procedures are given for the problem of selecting a subset which contains all populations better than a standard or control under simple or partial ordering prior. Here by simple or partial ordering prior we mean that there exist known simple or partial order relationships (defined more specifically later in Section 2) among unknown parameters. The procedures described do meet the usual requirement that the probability of a correct selection is greater than or equal to a predetermined number P*, the so-called P*-condition.

Many authors have considered the problem of comparing populations with a control under different types of formulations (see Gupta and Panchapakesan (1979)). Dunnett (1955) considered the problem of separating those treatments which are better than the control from those that are worse. Gupta and Sobel (1958), Gupta (1965), Naik (1975), Broström (1977) studied the problem of selecting a subset containing all populations better than the control. Lehmann (1961) discussed similar problems with emphasis on the derivation of a restricted minimax procedure. Gupta and Kim (1980), Gupta and Hsiao (1980) studied the problem of

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University.

selecting populations close to a control. In all these papers it is assumed that all populations are independent and that there is no information about the ordering of unknown parameters. However, in many situations, we may know something about the unknown parameters. What we know is always not the prior distributions but some partial or incomplete prior information, such as the simple or partial order relationship among the unknown parameters. This type of information about the ordering prior may come from the past experiences; or it may arise in the experiments where, for example, higher dose level of a drug always has larger effect on the patients.

In Section 2 definitions and notations used in this paper are introduced. In Section 3 we consider the problem for location parameters. We propose three types of selection procedures for the cases when the control parameter is known or not known (the scale parameter may or may not be assumed known). Some equivalent forms of the procedures are given, and their properties are discussed. In Section 3 simple ordering priors are assumed and some theorems in the theory of random walks are used. A selection procedure for the problem of selecting all populations better than the control under partial ordering prior is given in Section 4. Section 5 deals with the use of Monte Carlo techniques to make comparisons among the selection procedures proposed in Section 3 and those in Section 4, respectively.

2. Notations and Definitions

Suppose we have k + 1 populations $\pi_0, \pi_1, \ldots \pi_k$. The population treatment π_0 is called the control or standard population. Assume that the random variable X_{ij} is associated with $F(\cdot;\theta_i)$ and $X_{i1},\ldots,X_{in_i}$, $i = 1,\ldots,k$, are independent samples from $\pi_1,\ldots,\pi_k$. Assume that we have an ordering prior of $\theta_1,\ldots,\theta_k$. First we assume that the ordering prior is the simple order, so that without loss of generality, we may assume that, $\theta_1 \leq \cdots \leq \theta_k$. In Section 4 we will consider the partial ordering prior case. Note that the values of θ_i's are unknown.

Suppose our goal is to select a non-trivial (small) subset which contains all populations with parameters larger (smaller) than the control θ_0 (known or unknown) with probability not less than a given value P*.

The action space $\mathcal{G}$ is the class of all subsets of the set $\{1,2,\ldots,k\}$. An action A is the selection of some subset of the k populations. $i \in A$ means that π_i is included in the selected subset.

Let $\underline{\theta} = (\theta_0, \theta_1,\ldots,\theta_k)$. Then the parameter space is denoted by Ω, where $\Omega = \{\underline{\theta} \in \mathbb{R}^{k+1} \mid \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k; -\infty < \theta_0 < \infty\}$ is a subset of k + 1 dimensional Euclidean space $\mathbb{R}^{k+1}$.

The sample space is denoted by $\mathcal{X}$ where

$$\mathcal{X} = \{\underline{x} \in \mathbb{R}^{n_1+\cdots+n_k} \mid \underline{x} = (x_{11},\ldots x_{1n_1},\ldots,x_{k1},\ldots x_{kn_k})\}.$$

(Here θ_0 is assumed to be known).

Definition 2.1. A (non-randomized) selection procedure (rule) $\delta(\underline{x})$ is a mapping from $\mathcal{X}$ to $\mathcal{G}$.

A population π_i $(i = 1,\ldots,k)$ is called a good population if $\theta_i \geq \theta_0$. A correct selection (CS) is the selection of a subset which contains all good populations. A selection procedure δ satisfies the P*-condition if

$$\inf_{\underset{\sim}{\theta}\in\Omega} P_{\underset{\sim}{\theta}}(CS|\delta) \geq P^*. \tag{2.1}$$

Let $\mathcal{D} = \{\delta | \inf_{\underset{\sim}{\theta}\in\Omega} P_{\underset{\sim}{\theta}}(CS|\delta) \geq P^*\}$ be the collection of all selection procedures satisfying the P*-condition.

In the sequel we will use the isotonic estimators (see Barlow, Bartholomew, Bremner and Brunk (1972)). Hence we give the following definitions and theorems.

<u>Definition 2.2</u>. Let the set $\mathcal{J}$ be a finite set. A binary relation "$\lesssim$" on $\mathcal{J}$ is called a <u>simple order</u> if it is

(1) reflexive: $x \lesssim x$ for $x \in \mathcal{J}$

(2) transitive: $x, y, z \in \mathcal{J}$ and $x \lesssim y$, $y \lesssim z$ imply $x \lesssim z$

(3) antisymmetric: $x, y \in \mathcal{J}$ and $x \lesssim y$, $y \lesssim x$ imply $x = y$

(4) every two elements are comparable: $x, y \in \mathcal{J}$ imply either $x \lesssim y$ or $y \lesssim x$.

<u>Definition 2.3</u>. A <u>partial order</u> on $\mathcal{J}$ is a binary relation "$\lesssim$" on $\mathcal{J}$, such that it is (1) reflexive, (2) transitive, and (3) antisymmetric. Thus every simple order is a partial order. We use poset $(\mathcal{J},\lesssim)$ to denote the set $\mathcal{J}$ that has a partial order binary relation "$\lesssim$" on it.

Definition 2.4. A real-valued function f is called isotonic on poset $(\mathcal{J},\leq)$ if and only if (1) f is defined on $\mathcal{J}$, (2) if $x, y \in \mathcal{J}$, $x \leq y$ imply $f(x) \leq f(y)$.

Definition 2.5. Let g be a real-valued function on $\mathcal{J}$ and let W be a given positive function on $\mathcal{J}$. A function g^* on $\mathcal{J}$ is called an isotonic regression of g with weights W if and only if:

(1) g^* is an isotonic function on poset $(\mathcal{J},\leq)$

(2) $\sum_{x\in\mathcal{J}} [g(x) - g^*(x)]^2 W(x) = \min_{f\in\mathfrak{I}} \sum_{x\in\mathcal{J}} [g(x) - f(x)]^2 W(x)$,

where $\mathfrak{I}$ is the class of all isotonic functions on poset $(\mathcal{J},\leq)$.

From Barlow, et. al. (1972), (see their Theorems 1.3, 1.6 and the corollary there), we have the following theorems.

Theorem 2.1. There exists one and only one isotonic regression g^* of g with weight W on poset $(\mathcal{J},\leq)$.

There are some known algorithms, such as the "pool-adjacent-violators" algorithm (see page 13 of Barlow, et. al. (1972)) or Ayer, Brunk, Ewing, Reid and Silverman (1955) or the "up-and-down blocks" algorithm, Kruskal (1964), which show how to calculate the isotonic regression under simple order.

The following max-min formulas were given by Ayer et. al. (1955).

Theorem 2.2. (max-min formulas)

Assume that we have poset $(\mathcal{J},\leq)$ where $\mathcal{J} = \{\theta_1,\ldots,\theta_k\}$, $\theta_1 \leq \cdots \leq \theta_k$, and that function $g: \mathcal{J} \to \mathbb{R}$, then the isotonic regression g^* of g with weight W has the following formulas:

$$
\begin{aligned}
g^*(\theta_i) &= \max_{s\le i}\ \min_{t\ge i}\ Av(s,t) \\
&= \max_{s\le i}\ \min_{t\ge s}\ Av(s,t) \\
&= \min_{t\ge i}\ \max_{s\le i}\ Av(s,t) \\
&= \min_{t\ge i}\ \max_{s\le t}\ Av(s,t)
\end{aligned}
$$

where

$$
Av(s,t) = \frac{\sum_{r=s}^{t} g(\theta_r)W(\theta_r)}{\sum_{r=s}^{t} W(\theta_r)} .
$$

Corollary 2.1. $(g + c)^* = g^* + c$, $(ag)^* = ag^*$, if $a > 0$, $c \in \mathbb{R}$.

Corollary 2.2. $[\rho(g^*)g + \varphi(g^*)]^* = \rho(g^*) + \varphi(g^*)$, where ρ is a nonnegative function and φ is an arbitrary function.

3. Proposed Selection Procedures for the Normal Means Problem

We are interested in the (subset) selection problem of the unknown means of k normal populations in comparison with a standard or control normal with its mean known or unknown. Thus observations are taken on X_{ij} which are independently distributed normal random variables $N(\mu_i,\sigma^2)$, $j = 1,\ldots,n_i$; $i = 1,\ldots,k$. The values of $\mu_1,\mu_2,\ldots,\mu_k$ are unknown, but their ordering, say, $\mu_1 \le \mu_2 \le\cdots\le \mu_k$ is known. Note that in this case we replace $\underline{\theta}$ in the parameter space Ω by $\underline{\mu}$, all other quantities remaining the same.

Let us define the subspace $\Omega_i = \{\underline{\mu} \in \Omega \mid \mu_{k-i} < \mu_0 \leq \mu_{k-i+1}\}$ for $i = 1,\ldots,k-1$, the subspace $\Omega_k = \{\underline{\mu} \in \Omega \mid \mu_0 \leq \mu_1\}$, and the subspace $\Omega_0 = \{\underline{\mu} \in \Omega \mid \mu_k < \mu_0\}$; then we have $\Omega = \bigcup_{i=0}^{k} \Omega_i$. Note that the control μ_0 could be known or unknown. If μ_0 is unknown, we assume that the distribution of population π_0 is $N(\mu_0, \sigma^2)$ and we take independent observations $X_{01},\ldots,X_{0n_0}$ from π_0 and the sample space $\mathcal{X}$ becomes $\{\underline{X} \in \mathbb{R}^{n_0+\ldots+n_k} \mid \underline{X} = (X_{01},\ldots,X_{0n_0}, X_{11},\ldots,X_{1n_1},\ldots,X_{k1},\ldots,X_{kn_k})\}$. Using the partition $\{\Omega_0,\ldots,\Omega_k\}$ of parameter space Ω, we have

$$\inf_{\underline{\mu}\in\Omega} P_{\underline{\mu}}(CS|\delta) = \inf_{1\leq i\leq k} \{\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta)\},$$

for any selection procedure $\delta \in \mathcal{D}$. Hence the P*-condition is equivalent to

$$\inf_{\mu\in\Omega_i} P_{\underline{\mu}}(CS|\delta) \geq P^*, \quad \text{for } i = 1,\ldots,k.$$

Note that $\inf_{\mu\in\Omega_0} P_{\underline{\mu}}(CS|\delta) = 1$ for any selection procedure δ since there exists no good population in this case.

Let $X_i = x_i$ be the observed sample mean from population π_i, $i = 1,\ldots,k$. Let $\mathcal{J}$ denote the set $\{\mu_1, \mu_2,\ldots,\mu_k\}$ where $\mu_1 \leq \cdots \leq \mu_k$, and let $W(\mu_i) = n_i\sigma^{-2} \equiv w_i$, $g(\mu_i) = x_i$, $i = 1,\ldots,k$. Then by the max-min formulas, the isotonic regression of g is g*, where

$$g^*(\mu_i) = \max_{1\leq s\leq i} \min_{s\leq t\leq k} \frac{\sum_{j=s}^{t} x_j w_j}{\sum_{j=s}^{t} w_j}, \quad i = 1,\ldots,k.$$

The isotonic estimator of μ_i is denoted by $\hat{X}_{i:k}$, $i = 1,\ldots,k$ where

$$\hat{X}_{i:k} = \max_{1\le s\le i}\ \min_{s\le t\le k} \frac{\sum_{j=s}^{t} X_j w_j}{\sum_{j=s}^{t} w_j}$$

$$= \max_{1\le s\le i} \{\hat{\hat{X}}_{s:k}\} \tag{3.1}$$

where

$$\hat{\hat{X}}_{s:k} = \min\{X_s, \frac{X_s w_s + X_{s+1} w_{s+1}}{w_s + w_{s+1}}, \ldots, \frac{X_s w_s + \ldots + X_k w_k}{w_s + \ldots + w_k}\}. \tag{3.2}$$

It is known that the isotonic estimators $\hat{X}_{i:k}$, $i = 1,\ldots,k$ are also the maximum likelihood estimators of μ_i, $i = 1,\ldots,k$.

3.1. Proposed Selection Procedure δ_1

Case I. μ_0 known, common variance σ^2 known, and common sample size n.

Definition 3.1. We define the procedure δ_1 as follows:

Step 1. Select π_i, $i = 1,\ldots,k$ and stop, if

$$\hat{X}_{1:k} \ge \mu_0 - d^{(1)}_{1:k} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_1 and go to Step 2.

Step 2. Select π_i, $i = 2,\ldots,k$ and stop, if

$$\hat{X}_{2:k} \ge \mu_0 - d^{(1)}_{2:k} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_2 and go to Step 3.

⋮

Step k-1. Select π_i, $i = k-1, k$ and stop, if

$$\hat{X}_{k-1:k} \ge \mu_0 - d^{(1)}_{k-1:k} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_{k-1} and go to Step k.

Step K. Select π_k and stop, if

$$\hat{X}_{k:k} \ge \mu_0 - d^{(1)}_{k:k} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_k.

Here $d^{(1)}_{i:k}$'s are the smallest values such that $\delta_1 \in \mathcal{D}$, that is δ_1 satisfies the P*-condition.

3.2. On the Evaluation of $\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1)$ and the Values of the Constants $d_{1:k}^{(1)},\ldots,d_{k:k}^{(1)}$

For any $\underline{\mu}\in\Omega_i$, $1 \leq i \leq k$, let Z_i's be i.i.d. N(0,1) and let $\hat{\hat{Z}}_{r:k} = \min\{Z_r, \frac{Z_r+Z_{r+1}}{2},\ldots,\frac{Z_r+Z_{r+1}+\ldots+Z_k}{k-r+1}\}$. Then

$$= P_{\underline{\mu}}(\bigcup_{j=1}^{k-i+1}\{\hat{X}_{j:k} \geq \mu_0 - d_{j:k}^{(1)}\frac{\sigma}{\sqrt{n}}\})$$

$$= P_{\underline{\mu}}(\bigcup_{j=1}^{k-i+1}\bigcup_{r=1}^{j}\{\hat{\hat{X}}_{r:k} \geq \mu_0 - d_{j:k}^{(1)}\frac{\sigma}{\sqrt{n}}\})$$

$$\geq P(\bigcup_{j=1}^{k-i+1}\bigcup_{r=1}^{j}\{\hat{\hat{Z}}_{r:k} + \frac{\mu_r-\mu_0}{\sigma/\sqrt{n}} \geq - d_{j:k}^{(1)}\})$$

which is increasing in μ_r, $r = 1,\ldots,k-i+1$.

Hence

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1) \geq P(\hat{\hat{Z}}_{k-i+1:k} \geq - d_{k-i+1:k}^{(1)}).$$

On the other hand,

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1)$$

$$\leq P_{\underline{\mu}^*}(\bigcup_{j=1}^{k-i+1}\{\hat{X}_{j:k} \geq \mu_0 - d_{j:k}^{(1)}\frac{\sigma}{\sqrt{n}}\})$$

$$= P(\hat{\hat{Z}}_{k-i+1:k} \geq - d_{k-i+1:k}^{(1)})$$

whenever $\underline{\mu}^* = (\mu_0, -\infty,\ldots,-\infty, \underbrace{\mu_0,\ldots,\mu_0}_{i}) \in \bar{\Omega}_i$.

Thus, we have

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1) = P(\hat{\hat{Z}}_{k-i+1:k} \geq - d_{k-i+1:k}^{(1)}).$$

Since $\hat{\hat{Z}}_{k-i+1:k}$ has the same distributions as $\hat{\hat{Z}}_{1:i}$

letting

$$V_i = \hat{\hat{Z}}_{1:i} \tag{3.3}$$

we have

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1) = P(V_i \geq - d^{(1)}_{k-i+1:k}), \quad i = 1,\ldots,k. \tag{3.4}$$

It is clear from the above that $d^{(1)}_{k-i+1:k} = d^{(1)}_{1:i}$ for all $i = 1, 2,\ldots,k$, and $d^{(1)}_{1:i}$ is increasing in i.

Theorem 3.1. In case I, (μ_0 known, common known σ^2 and common sample size n), if $d^{(1)}_{k-i+1:k}$ is the solution of equation

$$P(V_i \geq - x) = P^* \tag{3.5}$$

where

$$V_i = \min_{1\leq r\leq i} \frac{1}{r}\sum_{j=1}^{r} Z_j \quad \text{and} \quad Z_i \text{ are i.i.d. } N(0,1),$$

$i = 1,\ldots,k$, then δ_1 satisfies the P*-condition.

Proof. For any i, $1 \leq i \leq k$,

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_1) = P(V_i \geq - d^{(1)}_{i-i+1:k}) = P^*,$$

so δ_1 satisfies the P*-condition.

Therefore, the problem of finding the $d^{(1)}_{i:k}$'s reduces to finding the distributions of $V_1,\ldots,V_k$. This is achieved by using some results in the theory of random walk.

3.3. Some Theorems in the Theory of Random Walk

Suppose Y_1, Y_2,... are independent random variables with a common distribution H not concentrated on a half-axis, i.e. $0 < P(Y_1 < 0)$, $P(Y_1 > 0) < 1$. The induced random walk is the sequence of random variables

$$S_0 = 0, \; S_n = Y_1 + \ldots + Y_n, \quad n = 1, 2, \ldots .$$

Let

$$\tau_n = P(S_1 \leq 0, \ldots, S_{n-1} \leq 0, S_n > 0) \tag{3.6}$$

and

$$\tau(s) = \sum_{n=1}^{\infty} \tau_n s^n, \quad 0 \leq s \leq 1. \tag{3.7}$$

Then we have the following theorem which was discovered by Andersen (1953). Feller (1971) gave an elegant short proof.

Theorem 3.2. (Feller (1971))

Let

$$p_n = P(S_1 > 0, \ldots, S_n > 0),$$

then

$$p(s) = \sum_{n=1}^{\infty} p_n s^n = \frac{1}{1-\tau(s)}, \tag{3.8}$$

hence

$$\log p(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n > 0). \tag{3.9}$$

By symmetry, the probabilities

$$q_n = P(S_1 \leq 0,\ldots,S_n \leq 0) \tag{3.10}$$

have the generating function q given by

$$\log q(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n \leq 0). \tag{3.11}$$

Note: The above theorem remains valid if the signs > and $\leq$ are replaced by $\geq$ and <, respectively.

<u>Theorem 3.3</u>. The generating function $\rho(s)$ of $P(V_j \geq x)$, $j = 1,2,\ldots$ is

$$\sum_{j=1}^{\infty} s^j P(V_j \geq x) = \exp\{\sum_{n=1}^{\infty} \frac{1}{n} s^n P(S_n \geq 0)\} \tag{3.12}$$

where

$$S_n = \sum_{i=1}^{n} (Z_i - x), \quad n = 1, 2,\ldots.$$

Proof. Since the distribution of random variable $Y_i = Z_i - x$ is not concentrated on a half-axis, and Y_i's are i.i.d. let $S_r = \sum_{i=1}^{r}(Z_i - x)$, $r = 1,\ldots,k$. Then

$$\{V_j \geq x\} = \{\min_{1\leq r\leq j} \frac{1}{r} S_r \geq 0\} = \{S_1 \geq 0,\ldots,S_j \geq 0\}.$$

By Feller's Theorem 3.2 , we complete the proof.

Now let

$$\Delta_j(x) \equiv \Delta_j = P(S_j \geq 0) = \phi(-x\sqrt{j}),\ j = 1,2,\ldots,$$

$$a(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} \Delta_n,$$

then we have

$$p(s) \equiv \sum_{j=1}^{\infty} s^j\ P(V_j \geq x) = \exp\ (a(s)).$$

Lemma 3.1. $p^{(n+1)}(s) = \sum_{j=0}^{n} \binom{n}{j} p^{(j)}(s)\ a^{(n+1-j)}(s)$, for all $n \geq 1$.

Proof. Since $p'(s) = p(s) \cdot a'(s)$, the result can be proved by induction on n.

Theorem 3.4. Under the assumption of Theorem 3.3

$$P(V_{n+1} \geq x) = \frac{1}{(n+1)!} \lim_{s \to 0^+} \frac{d^{n+1}p(s)}{ds^{n+1}}$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} P(V_j \geq x)\Delta_{n-j+1},\quad n = 0,\ 1,\ 2,\ldots \tag{3.13}$$

where

$$P(V_0 \geq x) \equiv 1, \quad \text{for all } x.$$

Proof. By Lemma 3.1, we have

$$P(V_{n+1} \geq x) = \frac{1}{(n+1)!} \lim_{s \to 0^+} p^{(n+1)}(s)$$

$$= \sum_{j=0}^{n} \frac{1}{(n+1)!} \frac{n!}{j!(n-j)!} p^{(j)}(0)\ [(n-j)!\ \Delta_{n+1-j}]$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} \frac{p^{(j)}(0)}{j!} \Delta_{n+1-j}$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} P(V_j \geq x)\Delta_{n+1-j}.$$

Let $G_n(x) = P(V_n \geq x)$ and $1-G_\infty(x)$ denote the limiting distribution function as $n \to \infty$ of V_n. Suppose the distribution of random variable $Y_1 = Z_1 - x$ is not concentrated on a half axis, then we have from Andersen-Feller Theorem

$$G_\infty(x) = \exp\{-\sum_{r=1}^{\infty} \frac{1}{r} P(S_r \leq 0)\} .$$

Now, let

$$G_\infty(-d^{(1)}_{1:\infty}) = P^*. \tag{3.14}$$

Now we can use the recurrence formula of Theorem 3.4 to solve the equations $P(V_i \geq -d^{(1)}_{k-i+1:k}) = P^*$, $i = 1,\ldots,k$.

Remark 3.1. From Section 3.2 we know that $d^{(1)}_{k-i+1:k} = d^{(1)}_{1:i}$ $(i = 1,\ldots,k)$. The values of $d^{(1)}_{1:k}$, for $k = 1\ (1)\ 6,\ 10,\ \infty$ and $P^* = .99, .975, .95, .925, .90$ are tabulated in Table I.

Definition 3.2. We define a selection procedure δ_1' by replacing the inequality in the ith step of procedure δ_1 by the inequality

$$\hat{X}_{i:k} \geq \mu_0 - d'_{i:k} \frac{\sigma}{\sqrt{n}}, \quad i = 1,\ldots,k$$

where $d'_{1:k},\ldots,d'_{k:k}$ are the smallest values such that δ_1' satisfies the P^*-condition.

Then it can easily be shown that the selection procedure δ_1 and δ_1' are identical and $d^{(1)}_{i:k} = d'_{i:k}$, $i = 1,2,\ldots,k$.

3.4. Some Other Proposed Selection Procedures δ_2, δ_3, δ_4

In Case I, we propose some other selection procedures:

<u>Definition 3.3.</u> We define a selection procedure δ_2 by

$$\delta_2: \text{ Select } \pi_i \text{ if and only if } \hat{X}_{i:k} \geq \mu_0 - d\frac{\sigma}{\sqrt{n}} \quad i = 1,\dots,k$$

where d is the smallest value such that δ_2 satisfies the P*-condition.

Note that under assumptions of Case I, and selection procedure δ_2, if we select population π_i, then we will select populations π_j, for all $j \geq i$, since $\hat{X}_{i:k} \leq \hat{X}_{j:k}$.

<u>Evaluation of the d-Values of δ_2</u>

For any i, $1 \leq i \leq k$, we have from a similar argument as for δ_1 that

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_2) = \inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(\hat{X}_{k-i+1:k} \geq \mu_0 - d\frac{\sigma}{\sqrt{n}})$$

$$= P(V_i \geq -d).$$

We need the constant d such that $P(V_i \geq -d) \geq P^*$ holds for all i, $1 \leq i \leq k$. By Theorem 3.1 we have $d = d^{(1)}_{1:k}$. It also follows that if S_1 and S_2 are the selected subsets associated with selection procedures δ_1 and δ_2, respectively, then $S_1 \subseteq S_2$. Thus δ_1 is better than δ_2.

<u>Definition 3.4.</u> The procedure δ_3 is defined as follows: Let $\tilde{X}_j = \max(X_1,\dots,X_j)$.

Step 1. Select π_i, $i \geq 1$ and stop, if

$$\tilde{X}_1 \geq \mu_0 - d_1\frac{\sigma}{\sqrt{n}},$$

otherwise reject π_1 and go to Step 2.

Step 2. Select π_i, $i \geq 2$ and stop, if

$$\tilde{X}_2 \geq \mu_0 - d_2 \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_2 and go to Step 3.

.
.
.

Step k-1. Select π_i, $i \geq k-1$ and stop, if

$$\tilde{X}_{k-1} \geq \mu_0 - d_{k-1} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_{k-1} and go to Step k

Step k. Select π_k and stop, if

$$\tilde{X}_k \geq \mu_0 - d_k \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_k.

Here d_i's are the smallest values such that δ_3 satisfies the P*-condition.

Evaluation of d_i's

For any i, $1 \leq i \leq k$,

$$\inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}(CS|\delta_3) = \inf_{\underline{\mu}\in\Omega_i} P_{\underline{\mu}}\left(\bigcup_{j=1}^{k-i+1} \{\tilde{X}_j \geq \mu_0 - d_j \frac{\sigma}{\sqrt{n}}\}\right)$$

$$= P_{\underline{\mu}^*}\left(\bigcup_{j=1}^{k-i+1} \{\tilde{X}_j \geq \mu_0 - d_j \frac{\sigma}{\sqrt{n}}\}\right)$$

$$= P(Z_{k-i+1} \geq -d_{k-i+1})$$

whenever $\underline{\mu}^* = (\mu_0, -\infty, \ldots, -\infty, \underbrace{\mu_0, \ldots, \mu_0}_{i}) \in \bar{\Omega}_i$.

Since Z_i is N(0,1), it implies $d_{k-i+1} = d$ for all i, and

$$d = \Phi^{-1}(P^*).$$

Hence, we have the following theorem:

Theorem 3.5. Selection procedure δ_3 satisfies the P*-condition with $d_i = d$, $i = 1,\ldots,k$, which do not depend on i. Hence the procedure is not changed if the statistics $\tilde{X}_i$ are replaced by X_i, the sample mean of population π_i for $i = 1,\ldots,k$.

The following procedure δ_4 was given by Gupta and Sobel (1958), without assuming any ordering prior:

Definition 3.5. The selection procedure δ_4 is defined as follows:

$$\delta_4: \text{ Select } \pi_i \text{ if and only if } X_i \geq \mu_0 - d\frac{\sigma}{\sqrt{n}} \quad i = 1,\ldots,k$$

where d is the smallest constant such that δ_4 satisfies the P*-condition.

It was shown that the value d is determined by the equation

$$\Phi(-d) = 1 - P^{*\frac{1}{k}} \text{ i.e. } d = \Phi^{-1}(P^{*\frac{1}{k}}).$$

3.5. Some Proposed Selection Procedures $\delta_i^{(2)}$, $i = 1, 2, 3, 4$ When μ_0 is Unknown

Case II. μ_0 unknown, common σ^2 known, common sample size n.

Definition 3.6. We define a selection procedure $\delta_1^{(2)}$ by replacing the inequalities

$$\hat{X}_{i:k} \geq \mu_0 - d_{i:k}^{(1)}\frac{\sigma}{\sqrt{n}}, \quad i = 1,\ldots,k$$

in procedure δ_1 (Definition 3.1) with

$$\hat{X}_{i:k} \geq X_0 - d_{i:k}^{(2)}\frac{\sigma}{\sqrt{n}}, \quad i = 1,\ldots,k, \text{ respectively.}$$

Here $X_0 = \sum_{i=1}^{n} X_{0i}/n$, $d_{i:k}^{(2)}$, $i = 1,\ldots,k$ are the smallest constants such that the selection procedure $\delta_1^{(2)}$ satisfies the P*-condition.

Similar to the Case I, we have the following theorem:

<u>Theorem 3.6</u>. For any i, $1 \leq i \leq k$, $d^{(2)}_{k-i+1:k}$ is determined by the equation

$$\int_{-\infty}^{\infty} P(V_i \geq t - d^{(2)}_{k-i+1:k}) d\Phi(t) = P^*. \qquad (3.15)$$

It is easy to see that $d^{(2)}_{k-i+1:k} = d^{(2)}_{1:i}$ and it is increasing in i. The following theorem gives us an identical form of the selection procedure $\delta_1^{(2)}$.

<u>Theorem 3.7.</u> The selection procedure $\delta_1^{(2)}$ is not changed if the statistics $\hat{X}_{i:k}$, $i = 1,\ldots,k$, are replaced by $\hat{\hat{X}}_{i:k}$, $i = 1,\ldots,k$, respectively.

Proof. The proof is straightforward and hence it is omitted.

The values $d^{(2)}_{1:i}$, $i = 1,\ldots,k$ are tabulated in Table II for $k = 1\ (1)\ 6, 8, 10, \infty$ and $P^* = .99, .975, .95, .925, .90$.

Similar to the Case I, we propose a selection procedure $\delta_2^{(2)}$ as follows:

<u>Definition 3.7</u>. We define a selection procedure $\delta_2^{(2)}$ by

$$\delta_2^{(2)}: \text{ Select } \pi_i \text{ if and only if } \hat{X}_{i:k} \geq X_0 - d\frac{\sigma}{\sqrt{n}} \qquad i = 1,\ldots,k$$

where d is the smallest value such that $\delta_2^{(2)}$ satisfies the P*-condition. Then, similar to procedure δ_2 we have $d = d^{(2)}_{1:k}$.

Next, we define a selection procedure $\delta_3^{(2)}$ which is similar to δ_3.

Definition 3.8. The selection procedure $\delta_3^{(2)}$ is defined by replacing $\tilde{X}_i \geq \mu_0 - d_i \frac{\sigma}{\sqrt{n}}$ in δ_3 (Definition 3.4) by $\tilde{X}_i \geq X_0 - d_i' \frac{\sigma}{\sqrt{n}}$, $i = 1,\ldots,k$ where $d_1',\ldots,d_k'$ are the smallest values such that $\delta_3^{(2)}$ satisfies the P*-condition.

Similar to Theorem 3.5 we have:

Theorem 3.8. The selection procedure $\delta_3^{(2)}$ satisfies the P*-condition with $d_i' = d$, $i = 1,\ldots,k$ where d is determined by the equation

$$\int_{-\infty}^{\infty} \Phi(d-t)d\Phi(t) = P^*. \tag{3.16}$$

And $\delta_3^{(2)}$ is not changed if the statistics $\tilde{X}_i$ is replaced by X_i, the sample mean of population π_i for $i = 1,\ldots,k$.

The following selection procedure $\delta_4^{(2)}$ was proposed by Gupta and Sobel (1958):

Definition 3.9. The selection procedure $\delta_4^{(2)}$ is defined by

$\delta_4^{(2)}$: Select π_i if and only if $X_i \geq X_0 - d \frac{\sigma}{\sqrt{n_i}}$ $i = 1,\ldots,k$

where d is determined by the following equation.

$$\int_{-\infty}^{\infty} \prod_{i=1}^{k} [\Phi(u\sqrt{\frac{n_i}{n_0}} + d)]\phi(u)du = P^*. \tag{3.17}$$

For the special case $n_i = n$ $(i = 0, 1, \ldots, k)$

$$\int_{-\infty}^{\infty} \Phi^k(t+d)\,\phi(t)\,dt = P^*. \tag{3.18}$$

Under the normal distribution N(0,1), the tables of d-values satisfying the Equation (3.18) for several values of P^* are given in Bechhofer (1954) for k = 1 (1) 10 and in Gupta (1956) for k = 1 (1) 50.

3.6. Some Proposed Selection Procedures $\delta_i^{(3)}$, i = 1, 2, 3, 4 for the Normal Means Problem When Common Variance σ^2 is Unknown

Case III. μ_0 known, common variance σ^2 unknown, $n_i = n > 1$.

Definition 3.10. We define the selection procedure $\delta_1^{(3)}$ by replacing the inequalities

$$\hat{X}_{i:k} \geq \mu_0 - d_{i:k}^{(1)} \frac{\sigma}{\sqrt{n}} \qquad i = 1, \ldots, k$$

in procedure δ_1 (Definition 3.1) by

$$\hat{X}_{i:k} \geq \mu_0 - d_{i:k}^{(3)} \frac{S}{\sqrt{n}} \qquad i = 1, \ldots, k, \text{ respectively,}$$

where $d^{(3)}$'s are the smallest values such that $\delta_1^{(3)}$ satisfies the P^*-condition; S^2 denotes the pooled estimator of σ^2 based on $\nu = k(n-1)$, that is

$$S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - X_i)^2/\nu. \tag{3.19}$$

Note that $\frac{\nu S^2}{\sigma^2}$ has the chi-square distribution χ_ν^2 with ν degrees of freedom. The following theorem then follows:

<u>Theorem 3. 9</u>. The equation which determines the constant $d_{k-i+1:k}^{(3)}$ is

$$P(V_i \geq -d_{k-i+1:k}^{(3)} \frac{S}{\sigma}) = P^* \qquad (3.20)$$

or

$$\int_0^\infty P(V_i \geq -d_{k-i+1:k}^{(3)} y) q_\nu(y) dy = P^* \qquad (3.21)$$

where $q_\nu(y)$ is the density of $\frac{S}{\sigma}$.

We can rewrite Formula (3.21) as

$$\int_0^\infty P(V_i \geq -d_{k-i+1:k}^{(3)} \sqrt{\frac{t}{\nu}}) d\chi_\nu^2(t) = P^*$$

or

$$\int_0^\infty P(V_i \geq -d_{k-i+1:k}^{(3)} \sqrt{\frac{2t}{\nu}}) \frac{t^{\frac{\nu}{2}-1} e^{-t}}{\Gamma(\frac{\nu}{2})} dt = P^*. \qquad (3.22)$$

Remark 3.2. The values of $d^{(3)}_{k-i+1:k}$, $i = 1,\ldots,k$ depend on $\nu = k(n-1)$; also $d^{(3)}_{k-i+1:k} \neq d^{(3)}_{1:i}$.

By using Rabinowitz and Weiss table (1959) (with N=24 and n of their table equal to 0) we have evaluated and tabulated the values of $d^{(3)}_{k-i+1:k}$, $i=1,\ldots,k$, in Table III, for $k = 2\ (1)\ 6$, $P^* = .99, .975, .95, .925, .90$, with common sample size $n = 3, 5, 9$, and 21.

For $k \geq 6$ and $n > 21$, i.e. $\nu > 120$ we can reasonably well approximate $d^{(3)}_{k-i+1:k}$ by $d^{(1)}_{1:i}$.

Definition 3.11. We define the selection procedure $\delta_2^{(3)}$ by

$$\delta_2^{(3)}: \text{ Select } \pi_i \text{ if and only if } \hat{X}_{i:k} \geq \mu_0 - d^{(3)} \frac{S}{\sqrt{n}} \quad i = 1,\ldots,k$$

where S is defined as in procedure $\delta_1^{(3)}$, and $d^{(3)}$ is the smallest constant such that $\delta_2^{(3)}$ satisfies the P*-condition.

As before, it can be shown that $d^{(3)} = d^{(3)}_{1:k}$.

Remark 3.3. In Case III the selection procedure $\delta_1^{(3)}$ will not be changed if we replace the isotonic statistics $\hat{X}_{i:k}$ by $\hat{\hat{X}}_{i:k}$, respectively. But this is not necessarily true for selection procedure $\delta_2^{(3)}$.

Definition 3.12. The selection procedure $\delta_3^{(3)}$ is defined to have the same form as procedure $\delta_3^{(2)}$ except that the inequality defined in the ith step of procedure $\delta_3^{(2)}$ is replaced by

$$X_i \geq \mu_0 - d \frac{S}{\sqrt{n}} \quad \text{for } i = 1,\ldots,k.$$

The proof of the following theorem uses the same arguments as that in Case I, hence it is omitted.

<u>Theorem 3.10</u>. The equation which determines the constant d of selection procedure $\delta_3^{(3)}$ is

$$\int_0^\infty \Phi(yd)q_\nu(y)dy = P^*. \qquad (3.23)$$

Gupta and Sobel (1958) gave a selection procedure $\delta_4^{(3)}$ in this case. It is as follows:

$\delta_4^{(3)}$: Select π_i if and only if $X_i \geq \mu_0 - D \frac{S}{\sqrt{n_i}}$ $\quad i = 1,\ldots,k$

and the equation which determines D is

$$\int_0^\infty \Phi^k(yD)q_\nu(y)dy = P^*, \qquad (3.24)$$

$$\text{where } \nu = \sum_{i=1}^{k} (n_i - 1).$$

3.7. <u>Some Proposed Selection Procedures $\delta_i^{(4)}$, i = 1, 2, 3, 4 for the Normal Means Problem When Both Control μ_0 and Common Variance σ^2 are Unknown</u>

<u>Case IV</u>. μ_0 unknown, common variance σ^2 unknown and common sample size n.

Here we replace μ_0 in each selection procedure $\delta_j^{(3)}$ by X_0, $1 \leq j \leq 4$, and get procedures $\delta_j^{(4)}$, $1 \leq j \leq 4$, respectively. The constants $d_{k-i+1:k}^{(4)}$, $i = 1,\ldots,k$, of procedure $\delta_1^{(4)}$ are determined by

$$\int_0^\infty \int_{-\infty}^\infty P(V_i \geq u - d^{(4)}_{k-i+1:k} \sqrt{\tfrac{t}{\nu}}) d\Phi(u) d\chi^2_\nu(t) = P^*. \qquad (3.25)$$

The constant d of procedure $\delta_2^{(4)}$ is

$$d = d^{(4)}_{1:k}.$$

The constants d of procedures $\delta_3^{(4)}$ and $\delta_4^{(4)}$ are determined by

$$\int_0^\infty \int_{-\infty}^\infty \Phi^r(u+\sqrt{\tfrac{t}{\nu}}d) d\Phi(u) d\chi^2_\nu(t) = P^* \qquad (3.26)$$

with r = 1 and k, respectively, and their values for selected values of P*, k and ν are given in Gupta and Sobel (1957) and Dunnett (1955).

3.8. Properties of the Selection Procedures

Under simple ordering prior, it is natural to require that an ideal selection procedure is isotonic as defined below:

Definition 3.13. A selection procedure δ is isotonic if it selects π_i with parameter μ_i, and if $\mu_i < \mu_j$, then it also selects π_j. Procedure δ is weak isotonic or monotone if

$P(\pi_i$ is selected$|\delta) \leq P(\pi_j$ is selected$|\delta)$ whenever $\mu_i < \mu_j$.

It is easy to see that any isotonic selection procedure is weak isotonic, but the converse is not true.

Now, let $\delta_i^{(1)} = \delta_i$, i = 1, 2, 3, 4.

Theorem 3.11. The selection procedures $\delta_1^{(i)}$, $\delta_2^{(i)}$ and $\delta_3^{(i)}$ are isotonic and procedure $\delta_4^{(i)}$ is monotone, for i = 1, 2, 3, 4.

Proof. The proof follows immediately from the definitions of the procedures.

Given observations $\underline{X} = \underline{x} = (x_0,\ldots,x_k)$ where x_i is the sample mean of population π_i, $i = 1,\ldots,k$, and $x_0 = \mu_0$ if μ_0 is known, otherwise x_0 is the sample mean of population π_0. Let

$$\psi_i(\underline{x}, \delta) = P(\pi_i \text{ included in the selected subset} | \underline{X} = \underline{x}, \delta) \qquad \text{for } i = 1,\ldots,k.$$

Definition 3.14. A selection procedure δ is called translation-invariant if for any $\underline{x} \in \mathbb{R}^{k+1}$, $c \in \mathbb{R}$

$$\psi_i(x_0 + c, x_1 + c,\ldots,x_k + c; \delta) = \psi_i(x_0,\ldots,x_k; \delta), \quad i = 1,\ldots,k.$$

Theorem 3.12. The selection procedures $\delta_1^{(i)}$, $\delta_2^{(i)}$, $\delta_3^{(i)}$ and $\delta_4^{(i)}$ are translation-invariant for i = 1, 2, 3, 4.

Proof. Proof is straightforward and hence omitted.

Expected Number (Size) of Bad Populations in the Selected Subset

Suppose the control μ_0 is known and we have common sample size n and common known variance σ^2; without loss of generality, we assume that $\mu_0 = 0$ and $\sigma/\sqrt{n} = 1$. Let $E(S'|\delta)$ denote the expected number of bad populations in the selected subset in using the selection procedure

δ, then for any j, $0 \le j \le k$,

$$\sup_{\underline{\mu}\in\Omega_{k-j}} E_{\underline{\mu}}(S'|\delta_1)$$

$$= \sup_{\underline{\mu}\in\Omega_{k-j}} \sum_{r=1}^{j} P_{\underline{\mu}}\left(\bigcup_{\ell=1}^{r} \{\hat{\hat{X}}_{\ell:k} \ge - d^{(1)}_{\ell:k}\}\right)$$

$$= \sum_{r=1}^{j} P\left(\bigcup_{\ell=1}^{r} \{\hat{\hat{Z}}_{\ell:j} \ge - d^{(1)}_{\ell:k}\}\right) . \qquad (3.27)$$

On the other hand, for procedure δ_2

$$\sup_{\underline{\mu}\in\Omega_{k-j}} E(S'|\delta_2) = \sum_{r=1}^{j} P\left(\bigcup_{\ell=1}^{r} \{\hat{\hat{Z}}_{\ell:j} \ge - d^{(1)}_{1:k}\}\right) . \qquad (3.28)$$

From (3.28) we see that the supremum for δ_2 is increasing in j and is greater than or equal to the supremum for δ_1 given in (3.27), since

$$d^{(1)}_{\ell:k} = d^{(1)}_{1:k-\ell+1} \le d^{(1)}_{1:k}.$$

Therefore, we have the following theorem (see also the remark just before Def. 3.4).

<u>Theorem 3.13</u>. For any i, $0 \le i \le k$

$$\sup_{\underline{\mu}\in\Omega_i} E(S'|\delta_2) \ge \sup_{\underline{\mu}\in\Omega_i} E(S'|\delta_1),$$

$$\sup_{\underline{\mu}\in\Omega} E(S'|\delta_2) = \sup_{\underline{\mu}\in\Omega_0} E(S'|\delta_2).$$

<u>Theorem 3.14</u>. In Section 3.1, Case I, for any j, $0 \le j \le k$

$$\sup_{\underline{\mu}\in\Omega_{k-j}} E(S'|\delta_3) = j - q(1-q^j)/P^* \qquad (3.29)$$

where $q = 1 - P^*$.

Proof.

$$\sup_{\underline{\mu}\in\Omega_{k-j}} E(S'|\delta_3)$$

$$= \sup_{\underline{\mu}\in\Omega_{k-j}} \sum_{i=1}^{j} P_{\underline{\mu}}(\text{select } \pi_i|\delta_3)$$

$$= \sup_{\underline{\mu}\in\Omega_{k-j}} \sum_{i=1}^{j} P_{\underline{\mu}}(\max_{1\le r\le i} X_r \ge -d)$$

$$= \sum_{i=1}^{j} (1 - \prod_{r=1}^{i} F(-d))$$

$$= j - \sum_{i=1}^{j} q^i$$

$$= j - q(1-q^j)/P^*$$

where $q = (1-P^*)$.

Theorem 3.15. $\sup_{\underline{\mu}\in\Omega_{k-j}} E(S'|\delta_3)$ is increasing in j, hence

$$\sup_{\underline{\mu}\in\Omega_{k-j}} E(S'|\delta_3) = \sup_{\Omega_0} E(S'|\delta_3) = k - q(1-q^k)/P^*. \tag{3.30}$$

Proof. Since

$$(j+1) - \sum_{i=1}^{j+1} q^i - (j - \sum_{i=1}^{j} q^i) = 1 - q^{j+1} > 0.$$

In Case I, Gupta (1965) showed that

$$\sup_{\underline{\mu}\in\Omega} E(S'|\delta_4) = kP^{*\frac{1}{k}}. \tag{3.31}$$

Let us define the event $A_i = \{\hat{\hat{Z}}_{i:k} \geq -d^{(1)}_{i:k}\}$, $i = 1,\ldots,k$; then we have

<u>Lemma 3.2</u>. $P(\bigcup_{i=1}^{j} A_i \cap A_{j+1}) > P(\bigcup_{i=1}^{j} A_i)\, P^*$ for all j, $1 \leq j \leq k-1$, $k \geq 2$.

Proof: $P(\bigcup_{i=1}^{j} A_i \cap A_{j+1})$

$$= P(\bigcup_{i=1}^{j} \{\hat{\hat{Z}}_{i:j} \geq -d^{(1)}_{i:k}\} \cap A_{j+1})$$

$$= P(\bigcup_{i=1}^{j} \{\hat{\hat{Z}}_{i:j} \geq -d^{(1)}_{i:k}\})\, P(A_{j+1})$$

$$> P(\bigcup_{i=1}^{j} A_i) P(A_{j+1})$$

$$= P(\bigcup_{i=1}^{j} A_i)\, P^*.$$

The above inequality is a result of the fact

$A_i \subset \{\hat{\hat{Z}}_{i:j} \geq -d^{(1)}_{i:k}\}$ for all $i = 1,\ldots,j$; $j = 1,\ldots,k-1$.

<u>Theorem 3.16</u>. For all $k \geq 2$, $\sup_{\Omega_0} E(S'|\delta_1) < \sup_{\Omega_0} E(S'|\delta_3)$.

Proof: To prove the theorem it is sufficient to show that for all given $k \geq 2$, $P(\bigcup_{i=1}^{j} A_i) \leq 1 - (1-P^*)^j$ for all j and strictly inequality holds for some j, $1 \leq j \leq k$.

It holds for $j = 1$, since $P(A_1) = P^*$. Suppose $P(\bigcup_{i=1}^{j} A_i) \leq 1-(1-P^*)^j$ is true for some j, $1 \leq j \leq k-1$, then

$$P(\bigcup_{i=1}^{j+1} A_i) = P(\bigcup_{i=1}^{j} A_i) + P^* - P(\bigcup_{i=1}^{j} A_i \cap A_{j+1})$$

$$< P(\bigcup_{i=1}^{j} A_i) + P^* - P(\bigcup_{i=1}^{j} A_i)P^*$$

$$\leq P^* + (1-P^*)(1-(1-P^*)^j)$$

$$= 1 - (1-P^*)^{j+1}.$$

Hence by induction principle, the proof is finished.

This theorem tells us that procedure δ_1 is better than δ_3 in the sense that in Ω_0 it tends to select smaller number of bad populations, however, procedure δ_1 is not uniformly better than δ_3. In some cases (see Section 5), δ_3 is slightly better than δ_1.

When the ordering prior among the unknown parameters is unknown, we can use the selection procedure of Gupta and Sobel (1958) or use the ordering of the sample means as the ordering of unknown parameters and apply the selection procedure which is originally used under ordering prior. In the normal case with the latter approach, the substitution implies that the isotonic regression of the sample means turns to the usual ordered sample means, and that the selection procedures $\delta_2^{(i)}$, i = 1, 2, 3, 4, are of the same type as $\delta_4^{(i)}$ (i = 1,2,3,4), respectively, and the selection procedures $\delta_j^{(i)}$, j = 1,3, i = 1,2,3,4 are of the same form as $\delta_5^{(i)}$, i=1,2,3,4, respectively, which are equivalent to the procedures proposed by Naik (1975) and Broström (1977), independently (see also Holm (1979)).

4. Selection Rules for the Location Parameter Under Partial Ordering Prior Assumption

Assume that we have only a partial ordering prior of k unknown location parameters, that is the parameter space

$\Omega' = \{\underline{\theta} | \underline{\theta} \in \mathbb{R}^k$ and there is a partial order relation "$\lesssim$" among θ_i's$\}$

Our approach is to partition the set $\{\theta_1,\ldots,\theta_k\}$ into several subsets, say $B_0,\ldots,B_\ell$, so that $B_i \cap B_j = \emptyset$, if $i \neq j$, $\bigcup_{j=1}^{\ell} B_j = \{\theta_1,\ldots,\theta_k\}$ and for each B_j $(j = 1,\ldots,\ell)$ there is a simple order on it and there is no order relation among the elements of subset B_0.

Let $b_i = |B_i|$, the number of elements contained in B_i, $i = 0,\ldots,\ell$, so we have

$$\sum_{i=0}^{\ell} b_i = k.$$

If we denote the new induced partial order by "$\lesssim'$", then we have a parameter space $\Omega'' \supset \Omega'$. We use an example to illustrate how to find an induced partial order.

<u>Example</u>. Suppose $k = 8$, and we have a partial ordering prior $\theta_1 \lesssim \theta_5$, $\theta_1 \lesssim \theta_8$, $\theta_1 \lesssim \theta_2 \lesssim \theta_3 \lesssim \theta_4$, and $\theta_2 \lesssim \theta_6 \lesssim \theta_7$. We use a "tree" to represent this partial ordering as in Figure 1.

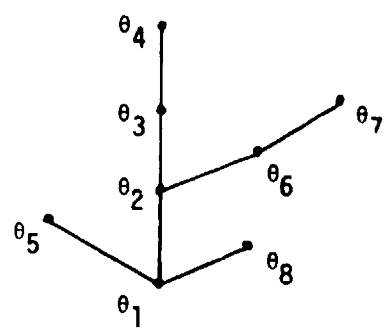

Figure 1. Original partial ordering

Then we have an induced partial ordering $\theta_1 \leq \theta_2 \leq \theta_3 \leq \theta_4$, $\theta_6 \leq \theta_7$ as in Figure 2.

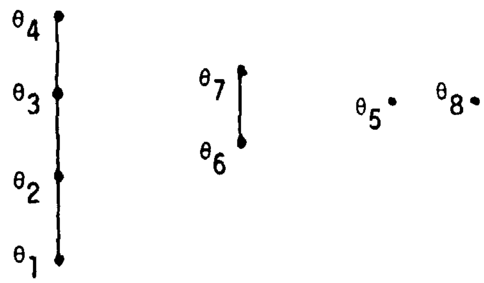

Figure 2. Induced partial ordering.

And

$$B_0 = \{\theta_5, \theta_8\}$$

$$B_1 = \{\theta_1, \theta_2, \theta_3, \theta_4\}$$

$$B_2 = \{\theta_6, \theta_7\}.$$

It is clear that the induced partial order is not unique, for example, we can partition $\{\theta_1,\ldots,\theta_8\}$ into three other subsets B_0', B_1', B_2' where

$$B_0' = \{\theta_5, \theta_8\}$$

$$B_1' = \{\theta_1, \theta_2, \theta_6, \theta_7\}$$

$$B_2' = \{\theta_3, \theta_4\}.$$

For the location parameter case, a selection procedure δ^P can be defined as follows:

Definition 4.1. We define a selection procedure δ^P as follows:

Suppose $B_0,\ldots,B_\ell$ are induced subsets and that for each subset B_j, $j = 1,\ldots,\ell$ there is a simple order on it. We choose a proper selection procedure δ for each subset B_j, such that the corresponding probability of a correct selection is not less than $P_j^* = P^{*\frac{b_j}{k}}$. For subset B_0 we may use selection procedure δ_4 or δ_5 with $P_0^* = P^{*\frac{b_0}{k}}$.

Theorem 4.1. The selection procedure δ^P satisfies the P*-condition.

Proof.

$$\begin{aligned}
&\inf_{\underline{\theta}\in\Omega'} P_{\underline{\theta}}(CS|\delta^P) \\
&\geq \inf_{\underline{\theta}\in\Omega''} P_{\underline{\theta}}(CS|\delta^P) \\
&\geq \prod_{i=1}^{\ell} \inf_{\Omega'_{B_i}} P(CS|\delta^P) \\
&\geq P^{*\left(\sum_{i=0}^{\ell} \frac{b_i}{k}\right)} = P^*
\end{aligned}$$

where Ω'_{B_i} is the parameter space associated with the subset B_i.

5. Comparisons of the Performance of Basic Rules for the Normal Means Problem

In this section we describe results of a Monte Carlo study to compare the performance of selection procedures δ_1, δ_2, δ_3, and δ_4. Suppose we have k independent populations, each population with distribution $N(\mu_i, \sigma^2)$, with common known variance σ^2 and common sample size n. Assume that the mean μ_0 of the control is known; without loss of generality we assume that $\mu_0 = 0$ and $\sigma/\sqrt{n} = 1$.

In the simulation study, we used Rubin and Hinkle's RVP-Random Variable Package, Purdue University Computing Center, to generate random numbers. For each k, we generated one random number (variable) for each population, then applied each selection procedure separately and repeated it ten thousand times; we used the relative frequencies as an approximation of the exact values of the associated performance characteristics for each procedure. In Table IV we use the following notations:

$\underline{\mu} = (\mu_1,\ldots,\mu_k)$, μ_i is the parameter of population π_i.

PS = P(CS)

PI = P(correctly rejecting all bad populations)

PC = P(correct classification of all population)

where the correct classification means that we select all good populations and reject all bad populations.

EI = Expected number (size) of bad populations contained in the selected subset.

$$EJ = \sum_{\mu_i < \mu_0} (\mu_i - \mu_0)^2 \, P\,(\pi_i \text{ is selected})$$

ES = Expected size of the selected subset.

Table IV.1 consists of four parts, namely, the four values of k = 2,3,4,5, for each value of k we assume that we have two bad populations. In this case based on the performance characteristics PI, PC, EI or EJ, we found the performance ordering as follows:

$$\delta_1 \succ \delta_2 \succ \delta_3 \succ \delta_4.$$

where $\delta_1 \succ \delta_2$ means that δ_1 is better than δ_2.

In Table IV.2 we assume that we have three bad populations for $k = 3$, and that both populations are bad for $k = 2$, this table indicates the same trend as Table IV.1, i.e. $\delta_1 \succ \delta_2 \succ \delta_3 \succ \delta_4$. If k is increased by adding strictly good (parameter strictly larger than control) populations, then $EI(\delta_i)$, $i = 1,2$ does not increase. This is because $\hat{\hat{X}}_{i:k} > \hat{\hat{X}}_{i:k+1}$ a.s. $1 \leq i \leq k$.

In Table IV.3 we assume that for each k, $k = 2,3,4,5$ that every population is bad. Based on the quantities PI, PC, EI and EJ, we find that the performance is as follows:

$$\delta_1 \succ \delta_2 \succ \delta_3 \succ \delta_4 .$$

This is the same result as before.

Table IV.4 has the same structure as before, but for each value of k, $k = 2,3,4,5$, we assume that the first population is the one and only one bad population with parameter -1 which is less than the control $\mu_0 = 0$. A glance at the table indicates that the performance, based on the characteristics PI, PC, EI and ES, can roughly be ordered as follows:

$$\delta_3 \gtrsim [\delta_2, \delta_1] \succ \delta_4 .$$

i.e. procedure δ_3 is the best and is slightly better than δ_2 and δ_1, δ_2 and δ_1 are very close and both are better than δ_4. As the number of populations k increases from two to five and the three additional populations are good populations with parameter 1, 2, and 3, respectively, we find that $EI(\delta_i, k = 5) - EI(\delta_i, k = 2)$, $i = 1,2,3,4$, is 0.0124, 0.0124, 0.0031, 0.121, respectively. This means that when k increases and the additional populations are good, then procedure δ_4 is the most sensitive procedure with k and thus not good in terms of EI. δ_3 seems to perform better in terms of EI while δ_1 and δ_2 are about the same.

In Table IV.5 we assume that the ordering prior of unknown parameter is incorrect; i.e. the true configuration (-2, -1,0,1,2) is replaced by (-1,-2,1,0,2). The simulation results indicate that, based on PI, PC, EI and EJ we have performance $\delta_1 > \delta_2 > [\delta_3,\delta_4]$. Thus here again δ_1 is the best. If we compare Table IV.5 with Table IV.1, we see that δ_4 does not change (the small differences are because of random fluctuations), $EI(\delta_3)$ and $EJ(\delta_3)$ increase quite appreciably.

From these five tables, it appears that, in general, the <u>overall performance</u> of these procedures is $\delta_1 > \delta_2 > \delta_3 > \delta_4$, if the ordering prior is correct. If there is no information regarding the prior ordering, then δ_4 or δ_5 seem to be an appropriate procedure to use.

TABLE I

Table of $d^{(1)}_{1:k}$ values (satisfying (3.5) and (3.14)) necessary to carry out the procedure δ_1 for the normal means problem under the simple ordering prior.

$d^{(1)}_{1:k}$	p^*				
k	.99	.975	.95	.925	.90
1	2.3264	1.9600	1.6449	1.4395	1.2816
2	2.3337	1.9775	1.6780	1.4872	1.3430
3	2.3339	1.9787	1.6817	1.4942	1.3538
4	2.3339	1.9787	1.6823	1.4956	1.3563
5	2.3339	1.9787	1.6824	1.4960	1.3571
6	2.3339	1.9787	1.6824	1.4960	1.3573
∞	2.3340	1.9787	1.6824	1.4960	1.3574

TABLE II

Table of $d^{(2)}_{1:k}$ values (satisfying (3.15)) necessary to carry out the procedure $\delta^{(2)}_1$ for the normal means problem under simple ordering prior.

$d^{(2)}_{1:k}$	p^*				
k	.99	.975	.95	.925	.90
1	3.2886	2.7711	2.3258	2.0355	1.8122
2	3.3449	2.8494	2.4267	2.1530	1.9434
3	3.3605	2.8730	2.4589	2.1917	1.9874
4	3.3673	2.8840	2.4723	2.2105	2.0091
5	3.3711	2.8901	2.4832	2.2215	2.0219
6	3.3734	2.8941	2.4890	2.2286	2.0303
8	3.3761	2.8988	2.4960	2.2375	2.0406
10	3.3776	2.9014	2.5000	2.2426	2.0440
∞	3.3787	2.9032	2.5021	2.2448	2.0487

TABLE III

Table of $d^{(3)}_{i:k} \equiv D(i:k)$ values (satisfying (3.22)) necessary to carry out the procedure $\delta^{(3)}_1$ for the normal means problem with common sample size n (common variance unknown) under simple ordering prior.

	n = 3					n = 5				
P* =	.990	.975	.950	.925	.900	.990	.975	.950	.925	.900
D(1:2) =	6.6993	4.9977	3.8822	3.2780	2.8643	5.0853	4.0805	3.3310	2.8889	2.5700
D(2:2) =	6.4966	4.8106	3.6930	3.0802	2.6558	5.0177	3.9939	3.2210	2.7578	2.4194
D(1:3) =	4.5435	3.5731	2.8815	2.4867	2.2081	3.8313	3.1412	2.6072	2.2858	2.0518
D(2:3) =	4.5272	3.5534	2.8567	2.4569	2.1735	3.8269	3.1334	2.5941	2.2674	2.0279
D(3:3) =	4.4443	3.4604	2.7482	2.3337	2.0361	3.7914	3.0813	2.5205	2.1750	1.9180
D(1:4) =	3.6058	2.8997	2.3763	2.0701	1.8512	3.1905	2.6405	2.2073	1.9440	1.7514
D(2:4) =	3.6037	2.8965	2.3713	2.0633	1.8424	3.1901	2.6397	2.2050	1.9401	1.7457
D(3:4) =	3.5958	2.8853	2.3553	2.0427	1.8173	3.1877	2.6347	2.1956	1.9261	1.7271
D(4:4) =	3.5480	2.8241	2.2776	1.9501	1.7108	3.1640	2.5965	2.1383	1.8520	1.6372
D(1:5) =	3.0674	2.4966	2.0637	1.8068	1.6216	2.7891	2.3203	1.9470	1.7186	1.5513
D(2:5) =	3.0670	2.4959	2.0625	1.8049	1.6189	2.7891	2.3201	1.9465	1.7178	1.5497
D(3:5) =	3.0659	2.4940	2.0590	1.7999	1.6123	2.7889	2.3194	1.9448	1.7147	1.5451
D(4:5) =	3.0613	2.4865	2.0473	1.7840	1.5922	2.7873	2.3158	1.9374	1.7033	1.5295
D(5:5) =	3.0281	2.4406	1.9854	1.7081	1.5032	2.7693	2.2851	1.8893	1.6399	1.4518
D(1:6) =	2.7100	2.2230	1.8473	1.6223	1.4593	2.5082	2.0935	1.7608	1.5566	1.4064
D(2:6) =	2.7099	2.2228	1.8470	1.6216	1.4583	2.5082	2.0935	1.7607	1.5563	1.4059
D(3:6) =	2.7098	2.2224	1.8461	1.6202	1.4563	2.5081	2.0934	1.7604	1.5555	1.4046
D(4:6) =	2.7091	2.2211	1.8436	1.6163	1.4509	2.5080	2.0929	1.7590	1.5530	1.4007
D(5:6) =	2.7060	2.2156	1.8343	1.6033	1.4340	2.5068	2.0900	1.7528	1.5432	1.3871
D(6:6) =	2.6810	2.1788	1.7823	1.5379	1.3562	2.4922	2.0640	1.7109	1.4871	1.3179

TABLE III (continued)

Table of $d_{i:k}^{(3)} = D(i:k)$ values (satisfying (3.22)) necessary to carry out the procedure $\delta_1^{(3)}$ for the normal means problem with common sample size n (common variance unknown) under simple ordering prior.

	n = 9					n = 21				
P^* =	.900	.975	.950	.925	.900	.990	.975	.950	.925	.900
D(1:2) =	4.5081	3.7260	3.1050	2.7240	2.4424	4.2169	3.5394	2.9828	2.6332	2.3714
D(2:2) =	4.4746	3.6720	3.0240	2.6191	2.3154	4.1973	3.5006	2.9165	2.5422	2.2570
D(1:3) =	3.5468	2.9598	2.4876	2.1963	1.9809	3.3948	2.8606	2.4208	2.1456	1.9405
D(2:3) =	3.5452	2.9557	2.4789	2.1824	1.9616	3.3940	2.8583	2.4142	2.1342	1.9238
D(3:3) =	3.5245	2.9189	2.4196	2.1031	1.8637	3.3801	2.8290	2.3627	2.0622	1.8326
D(1:4) =	3.0155	2.5272	2.1312	1.8862	1.7051	2.9192	2.4639	2.0881	1.8532	1.6785
D(2:4) =	3.0153	2.5267	2.1298	1.8835	1.7008	2.9192	2.4637	2.0872	1.8511	1.6749
D(3:4) =	3.0143	2.5238	2.1231	1.8724	1.6850	2.9187	2.4616	2.0818	1.8415	1.6607
D(4:4) =	2.9990	2.4947	2.0746	1.8064	1.6027	2.9074	2.4374	2.0381	1.7801	1.5827
D(1:5) =	2.6678	2.2412	1.8935	1.6777	1.5183	2.5999	2.1964	1.8628	1.6541	1.4991
D(2:5) =	2.6678	2.2412	1.8932	1.6771	1.5171	2.5999	2.1964	1.8627	1.6537	1.4982
D(3:5) =	2.6677	2.2409	1.8921	1.6749	1.5134	2.5999	2.1962	1.8619	1.6519	1.4950
D(4:5) =	2.6670	2.2385	1.8865	1.6654	1.4998	2.5995	2.1945	1.8572	1.6435	1.4825
D(5:5) =	2.6546	2.2140	1.8446	1.6078	1.4274	2.5899	2.1733	1.8187	1.5891	1.4132
D(1:6) =	2.4174	2.0344	1.7206	1.5256	1.3814	2.3667	2.0006	1.6975	1.5077	1.3668
D(2:6) =	2.4174	2.0344	1.7206	1.5255	1.3811	2.3667	2.0006	1.6974	1.5076	1.3665
D(3:6) =	2.4175	2.0344	1.7204	1.5250	1.3801	2.3667	2.0006	1.6973	1.5072	1.3657
D(4:6) =	2.4174	2.0341	1.7195	1.5231	1.3769	2.3667	2.0004	1.6966	1.5056	1.3629
D(5:6) =	2.4169	2.0322	1.7146	1.5147	1.3647	2.3664	1.9990	1.6924	1.4981	1.3516
D(6:6) =	2.4064	2.0106	1.6772	1.4630	1.2995	2.3578	1.9801	1.6577	1.4488	1.2887

TABLE IV.1

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 5) under simple ordering prior.

$P^* = .900$

	δ_1	δ_2	δ_3	δ_4
$k = 2$, $\underline{\mu} = (-2,-1)$				
PS	1.0000	1.0000	1.0000	1.0000
PI	.3420	.3252	.3001	.1719
PC	.3420	.3252	.3001	.1719
EI	.8673	.8841	.9389	1.0950
EJ	1.4952	1.5120	1.6559	2.1831
ES	.8673	.8841	.9389	1.0950
$k = 3$, $\underline{\mu} = (-2,-1,0)$				
PS	.9535	.9573	.9696	.9632
PI	.3437	.3407	.3037	.1233
PC	.2972	.2980	.2703	.1175
EI	.8585	.8615	.9350	1.2126
EJ	1.4651	1.4681	1.6421	2.4996
ES	1.8120	1.8188	1.9046	2.1758
$k = 4$, $\underline{\mu} = (-2,-1,0,1)$				
PS	.9596	.9606	.9715	.9747
PI	.3269	.3254	.2936	.0874
PC	.2865	.2860	.2651	.0851
EI	.8802	.8817	.9431	1.3062
EJ	1.5015	1.5030	1.6532	2.7378
ES	2.8387	2.8412	2.9142	3.2808
$k = 5$, $\underline{\mu} = (-2,-1,0,1,2)$				
PS	.9562	.9564	.9690	.9765
PI	.3333	.3331	.2984	.0746
PC	.2895	.2895	.2674	.0725
EI	.8835	.8837	.9480	1.3712
EJ	1.5339	1.5341	1.6872	2.9450
ES	3.8386	3.8390	3.9167	4.3477

TABLE IV.2

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 5) under simple ordering prior.

$P^* = .900$

k = 2, $\underline{\mu}$ = (-3,-2)				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.7551	.7380	.7342	.5912
PC	.7551	.7380	.7342	.5912
EI	.2632	.2803	.3035	.4395
EJ	1.1443	1.2127	1.4025	2.1590
ES	.2632	.2803	.3035	.4395
k = 3, $\underline{\mu}$ = (-3,-2,-1)				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.3362	.3156	.2837	.1090
PC	.3362	.3156	.2837	.1090
EI	.8937	.9166	1.0275	1.3290
EJ	1.6654	1.6952	2.1746	3.5318
ES	.8937	.9166	1.0275	1.3290
k = 4, $\underline{\mu}$ = (-3,-2,-1,0)				
	δ_1	δ_2	δ_3	δ_4
PS	.9579	.9616	.9737	.9731
PI	.3257	.3225	.2801	.0759
PC	.2836	.2841	.2538	.0736
EI	.9118	.9160	1.0419	1.4675
EJ	1.7093	1.7165	2.2324	4.1380
ES	1.8697	1.8776	2.0156	2.4406
k = 5, $\underline{\mu}$ = (-3,-2,-1,0,1)				
	δ_1	δ_2	δ_3	δ_4
PS	.9582	.9590	.9714	.9796
PI	.3292	.3281	.2877	.0602
PC	.2874	.2871	.2591	.0582
EI	.8962	.8976	1.0172	1.5283
EJ	1.6554	1.6577	2.1429	4.3912
ES	2.8536	2.8559	2.9884	3.5078

TABLE IV.3

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 5) under simple ordering prior.

$P^* = .900$

$k = 2,\ \underline{\mu} = (-4,-3)$				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.9613	.9560	.9585	.9130
PC	.9613	.9560	.9585	.9130
EI	.0392	.0445	.0448	.0876
EJ	.3563	.4040	.4263	.8493
ES	.0392	.0445	.0448	.0876
$k = 3,\ \underline{\mu} = (-4,-3,-2)$				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.7587	.7359	.7300	.4997
PC	.7587	.7359	.7300	.4997
EI	.2599	.2835	.3201	.5574
EJ	1.1340	1.2324	1.5547	2.9908
ES	.2599	.2835	.3201	.5574
$k = 4,\ \underline{\mu} = (-4,-3,-2,-1)$				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.3348	.3114	.2814	.0747
PC	.3348	.3114	.2814	.0747
EI	.9003	.9282	1.0440	1.4745
EJ	1.7013	1.7437	2.2947	4.3666
ES	.9003	.9282	1.0440	1.4745
$k = 5,\ \underline{\mu} = (-4,-3,-2,-1,-0.5)$				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.1117	.1045	.0615	.0036
PC	.1117	.1045	.0615	.0036
EI	1.7460	1.7600	1.9734	2.4985
EJ	1.8147	1.8275	2.4965	5.0978
ES	1.7460	1.7600	1.9734	2.4985

TABLE IV.4

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 5) under simple ordering prior.

$P^* = .900$

$k = 2$, $\underline{\mu} = (-1,0)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9453	.9490	.9579	.9470
PI	.3854	.3854	.3937	.2676
PC	.3307	.3344	.3516	.2530
EI	.6146	.6146	.6063	.7324
EJ	.6146	.6146	.6063	.7324
ES	1.5599	1.5636	1.5642	1.6794
$k = 3$, $\underline{\mu} = (-1,0,1)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9531	.9535	.9638	.9616
PI	.3741	.3741	.3826	.2044
PC	.3272	.3276	.3464	.1970
EI	.6259	.6259	.6174	.7956
EJ	.6259	.6259	.6174	.7956
ES	2.5771	2.5777	2.5803	2.7574
$k = 4$, $\underline{\mu} = (-1,0,1,2)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9580	.9582	.9640	.9765
PI	.3664	.3664	.3834	.1683
PC	.3244	.3246	.3474	.1640
EI	.6336	.6336	.6166	.8317
EJ	.6336	.6336	.6166	.8317
ES	3.5902	3.5904	3.5801	3.8081
$k = 5$, $\underline{\mu} = (-1,0,1,2,3)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9554	.9554	.9623	.9794
PI	.3730	.3730	.3906	.1465
PC	.3284	.3284	.3529	.1431
EI	.6270	.6270	.6094	.8535
EJ	.6270	.6270	.6094	.8535
ES	4.5812	4.5812	4.5714	4.8329

TABLE IV.5

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 5) under simple ordering prior.

$P^* = .900$

$k = 2,\ \underline{\mu} = (-1,-2)$				
	δ_1	δ_2	δ_3	δ_4
PS	1.0000	1.0000	1.0000	1.0000
PI	.5405	.5349	.2937	.1722
PC	.5405	.5349	.2937	.1722
EI	.8331	.8387	1.3151	1.0904
EJ	2.2116	2.2340	3.4232	2.1578
ES	.8331	.8387	1.3151	1.0904
$k = 3,\ \underline{\mu} = (-1,-2,1)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9932	.9943	.99[illegible]7	.9976
PI	.5365	.5349	.2987	.1190
PC	.5297	.5292	.2944	.1189
EI	.8347	.8363	1.3116	1.2154
EJ	2.2252	2.2316	3.4155	2.4919
ES	1.8279	1.8306	2.3073	2.2130
$k = 4,\ \underline{\mu} = (-1,-2,1,0)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9921	.9923	.9973	.9746
PI	.5271	.5269	.2894	.0873
PC	.5192	.5192	.2867	.0849
EI	.8498	.8500	1.3235	1.3077
EJ	2.2685	2.2693	3.4553	2.7474
ES	2.8390	2.8395	3.3207	3.2822
$k = 5,\ \underline{\mu} = (-1,-2,1,0,2)$				
	δ_1	δ_2	δ_3	δ_4
PS	.9906	.9906	.9958	.9795
PI	.5317	.5316	.2937	.0711
PC	.5223	.5222	.2895	.0693
EI	.8461	.8462	1.3217	1.3593
EJ	2.2510	2.2514	3.4406	2.8830
ES	3.8341	3.8342	4.3173	4.3388

BIBLIOGRAPHY

Andersen, E. S. (1953). On the fluctuation of sums of random variables. Math. Scand., 1, 263-285.

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. Ann. Math. Statist., 26, 641-647.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference under Order Restrictions. John Wiley & Sons, New York.

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., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures. Univ. of Chicago Press, Chicago.

Broström, G. (1977). An improved procedure for selecting all populations better than a standard. Tech. Report 1977-3, Inst. of Math. and Statist., Univ. of Umeå, Sweden.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. J. Amer. Statist. Assoc., 50, 1090-1121.

Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. II, 2nd Edition. John Wiley & Sons, New York.

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

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

Gupta, S. S. (1967). On selection and ranking procedures. Trans. Amer. Soc. Qual. Control., pp. 151-155.

Gupta, S. S. and Hsiao, P. (1980). On Γ-minimax, minimax and Bayes procedures for selecting populations close to a control. (submitted).

Gupta, S. S., and Kim, W. C. (1980). Γ-minimax and minimax decision rules for comparison of treatments with a control. Recent Developments in Statistical Inference and Data Analysis. North-Holland Publishing Company, 55-71.

Gupta, S. S. and Panchapakesan, S. (1972). On a class of subset selection procedures. Ann. Math. Statist., 43, 814-822.

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

Gupta, S. S. and Sobel, M. (1957). On a statistic arises in selection and ranking problems. Ann. Math. Statist. 28, 957-67.

Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist., 29, 235-244.

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

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. Psychometrika, 29, 115-129.

Lehmann, E. L. (1961). Some Model I problems of selection. Ann. Math. Statist., 32, 990-1012.

Marcus, R. (1976). The power of some tests of the equality of normal means against an ordered alternative. Biometrika, 63, 1, 177-183.

Naik, U. D. (1975). Some selection rules for comparing p processes with a standard. Comm. Statist., 4, 519-535.

Paulson, E. (1962). A sequential procedure for comparing several experimental categories with a standard or control. Ann. Math. Statist., 33, 438-443.

Rabinowitz, P. and Weiss, G. (1959). Tables of abscissas and weights for numerical evaluation of integrals of the form $\int_0^\infty e^{-x}x^n f(x)dx$. .th. Tables and Other Aids to Comput., Vol. XIII, 68, 285-294.

Williams, D. A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics, 27, 103-117.

Williams, D. A. (1977). Inference procedures for monotonically ordered normal means. Biometrika, 64, 9-14.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Mimeograph Series #81-24	2. GOVT ACCESSION NO. AD-A109288	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Isotonic Procedures for Selecting Populations Better Than a Control under Ordering Prior		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Mimeo. Series #81-24
7. AUTHOR(s) Shanti S. Gupta and Hwa-Ming Yang		8. CONTRACT OR GRANT NUMBER(s) ONR N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE July 1981
		13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION ST[ATEM]ENT (of [th]e abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY [NO]TES

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

Selection procedure, Subset selection, P*-condition, Stepwise procedure, Ordering prior, Isotonic regression, Isotonic estimators, Random walk, Isotonic procedure, Monte Carlo comparisons.

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

The problem of selecting a subset containing all populations better than a control under an ordering prior is considered. Three new selection procedures which satisfy a desirable basic requirement on the probability of a correct selection are proposed and studied. Two of the three procedures use the isotonic regression over the sample means of the k-treatments with respect to the given ordering prior. Tables of constants which are necessary to carry out the selection procedures with isotonic approach for the selection of unknown means of normal populations are given. The results including Monte Carlo studies indicate

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DEAIN SLC 1

that, in general, the stepwise procedure δ_1, using isotonic estimators, is the best.