Selection of a Restricted Subset of Normal Populations containing the One with Largest Mean*

by

Shanti S. Gupta and Thomas J. Santner

Purdue University

Division of Mathematical Sciences
Department of Statistics

Mimeograph Series #299

August 1972

*This research was supported in part by the Office of Naval Research Contract N00014-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Selection of a Restricted Subset of Normal Populations Containing the One with Largest Mean*

by

Shanti S. Gupta and Thomas J. Santner

Purdue University

1. Introduction and Terminology

One of the prime motivations for the use of subset selection procedures is to enable the experimenter to screen a group of populations selecting a subset of the best ones which will be further studied in a more intensive fashion. However, in practice, the experimenter has only limited resources to use for secondary exploration. Hence the goal in this paper is to give more flexibility to the experimenter than does the usual subset selection procedure by allowing him to specify an upper bound, m, on the number of populations included in the selected subset. Should the data clearly indicate a single population is best, this procedure still retains that advantage of the subset selection approach which would allow selection of fewer than the maximum number of populations, m. On the other hand, if the data make the choice of the best population less obvious this procedure still selects a subset for further study but guarantees that no more than m populations are selected.

---

*This research was supported in part by the Office of Naval Research Contract N00014-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Two special cases of this goal are the following. When $m = 1$ we select exactly one population and claim it is best. Such rules have been widely studied in the literature and in particular Bechhofer (1954) solves the normal means problem when the common variance is known under this formulation. When $m = k$ we select a subset whose size is a random variable $(1 \leq S \leq k)$ and claim the best population is a member of the subset. Such rules have also been widely studied in the literature and in particular Gupta (1956,65) solves the above normal means problem using such a procedure.

Formally our method of viewing the selection problem relates the subset selection formulation and the indifference zone formulation by showing both are special cases of a general theory. In practice our method allows us to blend some of the advantages of each method in the solution of the selection problem.

To fix ideas we introduce the following terminology which will distinguish the various types of rules used.

Let S be the number of populations selected by the procedure R. The goal is to select the "best" population. Ω is the set of all possible parameter configurations.

Definition 1.1: R is a fixed size subset rule means $\exists$ s $(1 \leq s < k)$ such that

$$P_\theta[S=s] = 1 \; \forall \theta \epsilon \Omega$$

Rules for which s=1 are also known as indifference zone rules and were introduced by Bechhofer (1954). In the more general case these rules were introduced by Mahamunulu (1966,67).

Definition 1.2: R is a restricted subset selection procedure means $\exists\ 1 < s < k$ such that $P_\theta[1 \leq S \leq s] = 1\ \forall\theta\epsilon\Omega$ and R is not a fixed size subset rule.

Definition 1.3: R is a subset selection procedure means $P_\theta[1 \leq S \leq k] = 1$ $\forall\theta\epsilon\Omega$ and R is neither a restricted subset selection procedure nor a fixed size subset selection procedure.

2. Statement of the Problem

Let $\pi_i \sim N(\mu_i, \sigma^2)$ for $i = 1, \ldots, k$ where the common σ^2 is known. Also let $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ be the ordered means and $\pi_{(i)}$ the population with mean $\mu_{[i]}$, the best population being $\pi_{(k)}$. We assume there is no a priori knowledge concerning the pairing of the $\{\pi_{(i)}\}$ and $\{\pi_i\}$. Let $\delta \geq 0$ and

$$\Omega = \{\mu = (\mu_1, \ldots, \mu_k) | \mu_i \in (-\infty, \infty) \forall i\}$$

$$\Omega(\delta) = \{\mu \in \Omega | \mu_{[k]} - \mu_{[k-1]} \geq \delta\}$$

$$\Omega^o(\delta) = \{\mu \in \Omega(\delta) | \mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta\}.$$

Goal G: Given P*, m and also possibly n and $\delta \geq 0$ define a procedure R based on a common sample size n from each population which selects a subset of the populations not exceeding m in size such that the subset contains the population $\pi_{(k)}$ and satisfies the basic probability requirement

$$P_\mu[CS|R] \geq P^*, \quad \forall\mu \in \Omega(\delta) \tag{2.1}$$

As we shall see later, by fixing δ, n, and $m < k$, the admissable range of P* values becomes

$$(1/k < P^* < (k-m)\binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1}\, d\Phi(t) \ .$$

The event $[CS|R]$ is the selection of any subset containing $\pi_{(k)}$.

We propose the following rule based on a sample of common size n from each of the k populations. As usual let $\bar{X}_i$ be the sample mean from π_i and $\bar{X}_{[1]} \leq \cdots \leq \bar{X}_{[k]}$.

<u>Rule R</u>: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\}$ (2.2)

The following are special cases of the goal G and rule R.

A) $m = k$, $\delta = 0$

$\Omega(0) = \Omega$

<u>G</u>: Choose a subset of $\{\pi_1, \ldots, \pi_k\}$ containing the best population such that $P_\mu[CS|R] \geq P^* \ \forall \mu \in \Omega$

<u>R</u>: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \bar{X}_{[k]} - d\sigma/\sqrt{n}$

These are the goal and procedure studied by Gupta (1956,65).

B) $m = 1$, $\delta > 0$

<u>G</u>: Choose a single population such that $P_\mu[CS|R] \geq P^* \ \forall \mu \in \Omega(\delta)$

<u>R</u>: Select that population π_i corresponding to $\bar{X}_{[k]}$.

Bechhofer (1954) studied this goal and procedure.

C) $m = s\,(1 < s < k)$, $d = +\infty$

<u>R</u>: Select the populations corresponding to $\bar{X}_{[k-s+1]}, \ldots, \bar{X}_{[k]}$

This procedure was studied by Mahamunulu (1966,67) and Desu and Sobel (1968). The procedure is a fixed size subset type and must satisfy (2.1).

3. Probability of a Correct Selection

We introduce the following notation. For every $\ell = 1,\ldots,k$ and for every $i = k-m,\ldots, k-1$ let $\{S_j^i(\ell): j = 1,\ldots,\binom{k-1}{i}\}$ be the collection of all subsets of size i from $\{1,\ldots,k\}-\{\ell\}$. Also let $\bar{S}_j^i(\ell) = \{1,\ldots,k\} - \{\ell\} - S_j^i(\ell)$.

<u>Theorem 3.1</u>. For any $\mu \in \Omega$, $P_\mu[CS|R] =$

$$\sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right) \prod_{\ell \in \bar{S}_j^i(k)} \left\{\Phi\left(t+d+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right)\right.$$

$$\left. - \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right)\right\} d\Phi(t)$$

Proof:

Let $\bar{X}_{(i)}$ denote the mean from population $\pi_{(i)}$, then

$$P[CS|R] = P[\bar{X}_{(k)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\}]$$

$$= P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n} \text{ for } \ell < k \text{ and } \bar{X}_{(k)} \geq \text{at least } (k-m)\ \bar{X}_{(\ell)}'s \text{ with } \ell \neq k].$$

Now for every $i = k-m,\ldots, k-1$ and $j = 1,\ldots,\binom{k-1}{i}$ let

$$A_j^i = [\bar{X}_{(k)} \geq \bar{X}_{(\ell)}\ \forall \ell \in S_j^i(k) \text{ and } \bar{X}_{(k)} < \bar{X}_{(\ell)}\ \forall \ell \in \bar{S}_j^i(k)]$$

$$\Rightarrow P[CS|R] = P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n}\ \forall \ell<k \text{ and } \bigcup_{i=k-m}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i]$$

$$= \sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n}\ \forall \ell < k \text{ and } A_j^i]$$

Now fix i and j

$$P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n} \quad \forall \ell < k \text{ and } A_j^i]$$

$$= P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} \ \forall \ell \in S_j^i(k) \text{ and } \bar{X}_{(k)} < \bar{X}_{(\ell)} < \bar{X}_{(k)} + d\sigma/\sqrt{n} \quad \forall \ell \in \bar{S}_j^i(k)]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})) \prod_{\ell \in \bar{S}_j^i(k)} \{\Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}) + d)$$

$$- \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}))\} \, d\Phi(t)$$

QED

Remark 3.1: As special cases we immediately obtain the results of Bechhofer (1954) and Gupta (1965)

A) Bechhofer ($m=1$, $0 < d < \infty$)

$$P[CS|R] = P[\bar{X}_{(k)} \geq \bar{X}_{[k]} - d\sigma/\sqrt{n} \text{ and } A_1^{k-1}]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell=1}^{k-1} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})) \, d\Phi(t)$$

since $A_1^{k-1} = [\bar{X}_{(k)} \geq \bar{X}_{(\ell)} \ \forall \ell < k] \subset [\bar{X}_{(k)} \geq \bar{X}_{(S)} - d\sigma/\sqrt{n} \ \forall \ell < k]$

B) Gupta ($m=k$, $0 < d < \infty$)

$$P[CS|R] = P[\bar{X}_{(k)} \geq \bar{X}_{(k)} - d\sigma/\sqrt{n} \text{ and } \bigcup_{i=0}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell=1}^{k-1} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}) + d) \, d\Phi(t)$$

since $\bigcup_{i=0}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i \supset [\bar{X}_{(k)} \geq \bar{X}_{[k]} - d\sigma/\sqrt{n}]$

Remark 3.2: An application of the dominated convergence theorem shows

$$P[CS|R] \to 1 \text{ as } \mu_{[k]} - \mu_{[k-1]} \to \infty .$$

Next we determine the infimum over $\Omega(\delta)$ of the probability of a correct selection.

<u>Theorem 3.2</u>. $\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^0(\delta)} P[CS|R]$

$$\sum_{i=k-m}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} \Phi^i\left(t+\frac{\sqrt{n}}{\sigma}\delta\right)\left\{\Phi\left(t+d+\frac{\sqrt{n}}{\sigma}\delta\right)-\Phi\left(t+\frac{\sqrt{n}}{\sigma}\delta\right)\right\}^{k-1-i} d\Phi(t)$$

Proof:

We use the following lemma due to Alam and Rizvi (1965) and also to D. Mahamunulu (1966).

Lemma

Let $X = (X_1, \ldots, X_k)$ have $k \geq 1$ independent components such that for every i, X_i has cdf $H(x_i|\theta_i)$. Suppose $\{H(x|\theta)\}$ form a stochastically increasing family. If $\phi(X)$ is a monotone function of x_i when all other components of X are held fixed, then $E_\theta[\phi(X)]$ is monotone in θ_i in the same direction.

$$\text{Let } \phi(X) = \begin{cases} 1, & \bar{X}_{(k)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise} \end{cases}$$

We claim $\phi(X)$ is non increasing in $\bar{X}_{(i)}$ for $i=1,\ldots,k-1$. Let

$\bar{X}_{(i)} < \bar{X}'_{(i)}$, $X = (\bar{X}_{(1)}, \ldots, \bar{X}_{(k)})$, $X' = (\bar{X}_{(1)}, \ldots, \bar{X}_{(i-1)}, \bar{X}'_{(i)}, \bar{X}_{(i+1)}, \ldots, \bar{X}_{(k)})$

$$\Rightarrow \max\left\{\bar{X}_{[k]} - d\sigma/\sqrt{n},\ \bar{X}_{[k-m+1]}\right\} \leq \max\left\{\bar{X}'_{[k]} - d\sigma/\sqrt{n},\ \bar{X}'_{[k-m+1]}\right\}$$

where the primes denote the order statistics from X'. So if $\phi(X) = 0 \Rightarrow \phi(X') = 0$. Hence

$P_\mu[CS|R] = E_\mu[\phi(X)]$ is nonincreasing in each of $\mu_{[1]},\ldots,\mu_{[k-1]}$ when all other means are fixed. So

$\inf\limits_{\Omega(\delta)} P[CS|R] = \inf\limits_{\Omega^0(\delta)} P[CS|R]$ and substituting the vector of means $(\mu_{[1]},\ldots,\mu_{[1]},\mu_{[1]}+\delta)$ gives the result.

QED

Remark 3.3: As special cases we get the results obtained by Gupta (1965), Bechhofer (1954) and Desu and Sobel (1968).

A) Bechhofer $(m=1,\ \delta>0)$

$$\inf_{\Omega(\delta)} P[CS|R] = \int_{-\infty}^{\infty} \Phi^{k-1}\left(t+\frac{\sqrt{n}}{\sigma}\delta\right)\, d\Phi(t)$$

B) Gupta $(m=k,\ 0<d<\infty,\ \delta=0)$

$$\inf_{\Omega} P[CS|R] = \sum_{i=0}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} \Phi^i(t)\{\Phi(t+d) - \Phi(t)\}^{k-i-1} d\Phi(t)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(t+d) \sum_{i=0}^{k-1} \binom{k-1}{i} \left[\frac{\Phi(t)}{\Phi(t+d)}\right]^i \left[1-\frac{\Phi(t)}{\Phi(t+d)}\right]^{k-1-i} d\Phi(t)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(t+d)\, d\Phi(t)$$

C) Desu and Sobel $(1\le m<k,\ d=+\infty,\ \delta>0)$

$$\inf P[CS|R] = \int_{-\infty}^{\infty} \sum_{i=k-m}^{k-1} \binom{k-1}{i} \Phi^i\left(t+\frac{\sqrt{n}}{\sigma}\delta\right)\left\{1-\Phi\left(t+\frac{\sqrt{n}}{\sigma}\delta\right)\right\}^{k-1-i} d\Phi(t)$$

$$= (k-m) \binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)][1-\Phi(t)]^{m-1}[\Phi(t)]^{k-m-1} \, d\Phi(t)$$

as we will later show.

4. Properties of R

Next we study the properties of the procedure R. To facilitate this study we let $p_\mu(i) = P_\mu[R \text{ selects } \pi_{(i)}]$ and recall the following two definitions.

<u>Definition 4.1</u>: R is a monotone procedure means $\forall \mu \in \Omega$ and $i < j$ $p_\mu(i) \leq p_\mu(j)$.

<u>Definition 4.2</u>: R is an unbiased procedure means $\forall \mu \in \Omega$ and $j < k$

$$P_\mu[R \text{ does not select } \pi_{(j)}] \geq P_\mu[R \text{ does not select } \pi_{(k)}].$$

Of course R monotone $\Rightarrow$ R unbiased. Other optimal properties are

<u>Definition 4.3</u>: R is consistent wrt Ω' means $\lim_{n\to\infty} \inf_{\Omega'} P[CS|R] = 1$

<u>Definition 4.4</u>: R is strongly monotone in $\pi_{(i)}$ means

$$p_\mu(i)\begin{cases} \uparrow \text{ in } \mu_{[i]} \text{ when all other components of } \mu \text{ are fixed} \\ \downarrow \text{ in } \mu_{[j]} \text{ when all other components of } \mu \text{ are fixed } (j \neq i) \end{cases}$$

Remard 4.1: If R is non decreasing for $\pi_{(k)}$ then

$$\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^0(\delta)} P[CS|R]$$

<u>Theorem 4.1</u>. For every $i = 1,\dots,k$ and for all procedures R of the form (2.2), R is strongly monotone in $\pi_{(i)}$

Proof:

1) We have already shown this result for $i = k$. Since for $i < k$ we have $p_\mu(i) = E_\mu[\eta(x)]$ where $\eta(x) = \begin{cases} 1, & \bar{X}_{(i)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise,} \end{cases}$

the same argument applies to give the desired conclusion.

QED

Corollary 4.1. All rules of the form (2.2) are monotone and unbiased.

The proof follows from the definition of monotonicity and the property of being strongly monotone in $\pi_{(i)} \forall i$.

Theorem 4.2. For every rule R of form (2.2) and every $\delta > 0$, R is consistent wrt $\Omega(\delta)$.

Proof:

We must show

$$\sum_{i=k-m}^{k-1} \binom{k-1}{i} \lim_{n\to\infty} \int_{-\infty}^{\infty} \Phi^i(t+\tfrac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\tfrac{\sqrt{n}}{\sigma}\delta) - \Phi(t+\tfrac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\, d\Phi(t) = 1$$

We note each integrand is bounded wrt Φ measure and so dominated convergence applies.

For every $i < k-1$ we have

$$\lim_{n\to\infty} \Phi^i(t+\tfrac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+\tfrac{\sqrt{n}}{\sigma}\delta+d) - \Phi(t+\tfrac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i} = 0$$

and for $i = k-1$

$\lim_{n\to\infty} \Phi^{k-1}(t+\tfrac{\sqrt{n}}{\sigma}\delta) = 1$. Hence the result follows.

QED

This theorem says that no matter what probability requirement ($\delta > 0$, P^*) is made and which rule is used, (2.1) can be made to hold by choosing a sufficiently large sample.

<u>Theorem 4.3</u>. For every n and rule R of form (2.2), $\lim_{n\to\infty} \inf_{\Omega(\delta)} P[CS|R] = 1$.

For every n, $m < k$, and $\delta > 0$, $\lim_{n\to\infty} \inf_{\Omega(\delta)} P[CS|R] =$

$$(k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}[1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)]\ \Phi^{k-m-1}(t)\ [1-\Phi(t)]^{m-1}\ d\Phi(t)$$

Proof:

The first result follows from dominated convergence. The second result follows from the same theorem and

$$\lim_{n\to\infty}\ \inf_{\Omega(\delta)} P[CS|R] = \sum_{i=k-m}^{k-1}\binom{k-1}{i}\int_{-\infty}^{\infty}\Phi^{i}(t+\frac{\sqrt{n}}{\sigma}\delta)\ \{1-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\ d\Phi(t)$$

$$= \int_{-\infty}^{\infty}(k-m)\binom{k-1}{k-m}\int_{1-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)}^{1} y^{m-1}(1-y)^{k-m-1}dy\ d\Phi(t)$$

Letting $w = \Phi^{-1}(1-y)$ and changing the order of integration yields

$$(k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}\int_{w-\frac{\sqrt{n}}{\sigma}\delta}^{\infty} d\Phi(t)[1-\Phi(w)]^{m-1}[\Phi(w)]^{k-m-1}\ d\Phi(w)$$

QED

<u>Remark 4.2</u>: The first part states that by taking δ sufficiently large we can attain any P^* probability requirement for any rule d based on any number of observations. The second result says that given an indifferent zone $\delta \geq 0$ and common sample size n we can not achieve all P^* values. We can only attain

$$P^* < (k-m)\binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1} d\Phi(t) < 1$$

This interpretation follows since $\inf_{\Omega(\delta)} P[CS|R]$ is a monotone non decreasing function of d.

Remark 4.3: Using the monotonicity of $\inf_{\Omega(\delta)} P[CS|R]$ we can obtain the following bounds: For $m < k$ and $d \geq 0$

$$\int_{-\infty}^{\infty} \Phi^{k-1}(t+\frac{\sqrt{n}}{\sigma}\delta)\, d\Phi(t) \leq \inf_{\Omega(\delta)} P[CS|R]$$

$$\leq (k-m)\binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1}\, d\Phi(t).$$

For the purpose of implementing the procedure R we have prepared Table I found at the end of the paper. The body of the table contains the values of $\frac{\sqrt{n}}{\sigma}\delta$ necessary to obtain $P^* = .75, .90, .975$ using rules $d = .4, .7, 1.3$ and 1.6 for $k = 3(1)5$ with $m = 2(1)k-1$ and also for $k = 6(1)10, 15, 20$ with $m = 2(1)5$. In general given P^*, d, k and m the corresponding $\frac{\sqrt{n}}{\sigma}\delta$ is the solution of the following equation:

$$P^* = \sum_{i=k-m}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} \Phi^i(t+\frac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\frac{\sqrt{n}}{\sigma}\delta) - \Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\, d\Phi(t).$$

To compare this rule to the fixed size subset rule, we have calculated

$$e(P^*,k,m,d) = \frac{n(d)}{n(\infty)}$$

where $n(d')$ is the sample size necessary to achieve probability requirement (2.1) using rule R with k,m and d'. The ratio shows the relative samples sizes of the restricted subset selection rule to the fixed size subset rule when both attain the same probability requirements. For larger d values this ratio is close to one indicating that in many cases a slight additional cost will allow use of a restricted subset selection procedure and still meet the same probability requirement. The exact savings in terms of $(m - E_\mu[S|R])$ depends of course on the underlying μ. Some exact comparisons for the equispaced means and slippage configurations will be described in the next section.

5. Expected Number of Selected Populations

As usual define

$$Y_i = \begin{cases} 1, & \bar{X}_{(i)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise} \end{cases}$$

which gives $S = \sum_{i=1}^{k} Y_i$ = number of populations selected.

Then $E_\mu[S] = \sum_{1}^{k} p_\mu(i)$

Theorem 5.1. For every $\mu \in \Omega$, $E_\mu[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^P(i)} \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})\right) \prod_{\ell \in \bar{S}_j^P(i)} \left\{\Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})+d\right) - \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})\right)\right\} d\Phi(t)$$

Proof:

From the above discussion we see that it suffices to calculate $p_\mu(i)$ for $i = 1,\ldots,k$. Using arguments similar to those above we get

$$P_\mu(i) = \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} P[\bar{X}_{(i)} > \bar{X}_{(\ell)} \ \forall \ell \in S_j^p(i) \text{ and } \bar{X}_{(i)} + d\sigma/\sqrt{n} > \bar{X}_{(\ell)} > \bar{X}_{(i)} \forall \ell \in \bar{S}_j^p(i)]$$

$$= \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^p(i)} \Phi(t + \frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})) \prod_{\ell \in \bar{S}_j^p(i)} \{\Phi(t + \frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]}) + d)$$

$$- \Phi(t + \frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]}))\} \, d\Phi(t)$$

QED

Remark 5.1: $E_\mu[S|R] \leq m \quad \forall \mu \in \Omega$

Remark 5.2: If $m = k \Rightarrow \sup_{\Omega(\delta)} E_\mu[S|R] = \sup_{\Omega^o(\delta)} E_\mu[S|R]$

This was proved by Gupta (1965).

Since $E_\mu[S|R]$ is increasing in d the experimenter should seek to use rules with small d. On the other hand for fixed δ and P* the smaller d is the larger n must be to achieve (2.1). Hence, the experimenter must decide what trade off between n, d, and δ he is willing to accept.

To investigate his interdependence in more detail we have tabulated $E[S|R]$ under the following configurations.

A) Equispaced Means $\mu = (\alpha, \alpha+\delta, \alpha+2\delta, \ldots, \alpha+(k-1)\delta)$

Given P*, d, $\frac{\sqrt{n}}{\sigma}\delta$, k and m, Table III displays $E_\mu[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^p(i)} \Phi(t + \frac{\sqrt{n}}{\sigma}(i-\ell)\delta) \prod_{\ell \in \bar{S}_j^p(i)} \{\Phi(t + d + \frac{\sqrt{n}}{\sigma}(i-\ell)\delta) - \Phi(t + \frac{\sqrt{n}}{\sigma}(i-\ell)\delta)\} \, d\Phi(t)$$

B) Slippage $\mu = (\alpha,\ldots,\alpha,\alpha+\delta)$

Again given P*, d, $\frac{\sqrt{n}}{\sigma}\delta$, k and m, Table IV gives

$$E_\mu[S|R] = \sum_{p=k-m}^{k-1} \binom{k-1}{p} \int_{-\infty}^{\infty} \Phi^p(t+\tfrac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\tfrac{\sqrt{n}}{\sigma}\delta)-\Phi(t+\tfrac{\sqrt{n}}{\sigma}\delta)\}^{k-1-p} d\Phi(t)$$

$$+ (k-1)\sum_{p=k-m}^{k-1} [\binom{k-2}{p-1} \int_{-\infty}^{\infty} \Phi(t-\tfrac{\sqrt{n}}{\sigma}\delta)\Phi^{p-1}(t)\{\Phi(t+d)-\Phi(t)\}^{k-1-p} d\Phi(t)$$

$$+ \binom{k-2}{p} \int_{-\infty}^{\infty} \Phi^p(t)\{\Phi(t+d-\tfrac{\sqrt{n}}{\sigma}\delta)-\Phi(t-\tfrac{\sqrt{n}}{\sigma}\delta)\}\{\Phi(t+d)-\Phi^p(t)\}^{k-2-p} d\Phi(t)].$$

The same two tables also list

A) $\sum_{i=1}^{k}(k-i+1)p_\mu(i)$, the expected sum of ranks of the selected populations ($\pi_{(k)}$ is assigned rank 1 etc.) and

B) $E_\mu[S|R]/m$, the expected proportion of selected populations.

As an application of the theory we give the following example. An experimenter is sampling from nine normal populations with $\pi_i \sim N(\mu_i,1)$. He wishes to select a subset of size at most four which contains the population with largest mean. For his screening process he wishes to have a probability of correct selection at least .975 whenever $\mu_{[9]} - \mu_{[8]} \geq .8$. As $E_\mu[S|R]$ is increasing in d and n, he wishes both to be small.

Examining the four rules specified in Table I he finds his choices

are

d	.4	.7	1.3	1.6
n	18	15	11	10

He decides to base his preliminary research on a sample of size 10 from each population and uses the rule:

R: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[6]}, \bar{X}_{[9]} - 1.6/\sqrt{10}\}$.

6. Extension to Location and Scale Parameter Family

We assume we are given independent random variables $X_1, \ldots, X_k$ from k populations $\pi_1, \ldots, \pi_k$ with cdf's $F_{\theta_i}(x)$ where

A) $F_\theta(x) = F(x-\theta)$ $(\theta\varepsilon(-\infty,\infty))$ in the location parameter case and

B) $F_\theta(x) = F(x/\theta)$ $(F(0)=0$ and $\theta > 0)$ in the scale parameter case.

Here the θ_i's are unknown but F is known. Our goal is to select a subset of the populations not exceeding m in size such that

$$P_\theta[CS] \geq P^* \quad \forall\theta\varepsilon\Omega(\delta) \tag{6.1}$$

The event [CS] is the selection of any subset containing population $\pi_{(k)}$ and

A) $\Omega(\delta) = \{\theta | \theta_{[k]} - \theta_{[k-1]} \geq \delta\}$ in the location parameter case and

B) $\Omega(\delta) = \{\theta | \theta_{[k]}/\theta_{[k-1]} \geq \delta\}$ in the scale parameter case.

As usual there is no knowledge of the correct pairing of the $\{\pi_j\}$ and $\{\pi_{(j)}\}$. Our rules are the following:

A) R: Select $\pi_i \Leftrightarrow X_i \geq \max\{X_{[k]} - d\ ,\ X_{[k-m+1]}\}$

where $d > 0$ in the location parameter case and (6.2)

B) R': Select $\pi_i \Leftrightarrow X_i \geq \max\{c\, X_{[k]}, X_{[k-m+1]}\}$ where $0 < c < 1$ in the scale parameter case (6.3)

As the results for the present cases are completely analogous to those in the normal case we present them only for the location parameter family and then without proofs.

<u>Theorem 6.1</u>. $P_\theta[CS|R] =$

$$\sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} F(t+\theta_{[k]} - \theta_{[\ell]}) \prod_{\ell \in \bar{S}_j^i(k)} \{F(t+d+\theta_{[k]} - \theta_{[\ell]}) - F(t+\theta_{[k]}-\theta_{[\ell]})\}\, dF(t)$$

and further

$$\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^0(\delta)} P[CS|R]$$

$$= \sum_{i=k-m}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} F^i(t+\delta)\{F(t+d+\delta) - F(t+\delta)\}^{k-1-i}\, dF(t)$$

The optimality properties of R and R' parallel those in the normal case.

<u>Theorem 6.2</u>. For every $i = 1,\ldots,k$ and every R of the form (6.2), R is non decreasing in $\pi_{(i)}$. Hence R is monotone and unbiased.

<u>Theorem 6.3</u>. For every indifference zone $\delta \geq 0$ and $m < k$.

$$\lim_{d\to\infty} \inf_{\Omega(\delta)} P[CS|R] = \sup_{d\geq 0} \inf_{\Omega(\delta)} P[CS|R]$$

$$= (k-m)\binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-F(w-\delta)][F(w)]^{k-m-1}[1-F(w)]^{m-1}\, dF(w) \qquad (6.4)$$

As before we are not able to attain all P* values for a given δ merely by choosing sufficiently large d. The right hand side of (6.4) is the upper bound on the attainable P* values.

<u>Theorem 6.4</u>. For every rule of form (6.2),

$$\lim_{\delta\to\infty} \inf_{\Omega(\delta)} P[CS|R] = 1.$$

Again analogous to the previous results, for any rule R and P* we can always choose an indifference zone large enough so that (6.2) holds. Finally we obtain the general expression for the expected number of populations selected.

<u>Theorem 6.4</u>. $E_\theta[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell\in S_j^P(i)} F(t+\theta_{[i]}-\theta_{[\ell]}) \prod_{\ell\in \bar{S}_j^P(i)} \{F(t+d+\theta_{[i]}-\theta_{[\ell]})-F(t+\theta_{[i]}-\theta_{[\ell]})\}dF(t).$$

Here as before the expected savings using a restricted subset procedure over the fixed size subset procedure depends upon the underlying F(x) and parameter point in question.

Table I

Tables $\frac{\sqrt{n}}{\sigma}\delta$ needed to attain P* levels .75, .90, .975 for the rules given by d = .4, .7, 1.3 and 1.6

k	m \ d	.4			.7		
		P*			P*		
		.75	.90	.975	.75	.90	.975
3	2	1.078	1.891	2.766	0.863	1.645	2.520
4	2	1.342	2.123	2.998	1.156	1.906	2.781
	3	1.286	2.067	2.942	1.017	1.767	2.642
5	2	1.516	2.266	3.141	1.348	2.098	2.973
	3	1.456	2.225	3.125	1.192	1.942	2.817
	4	1.448	2.216	3.104	1.151	1.932	2.807
6	2	1.647	2.397	3.272	1.490	2.209	3.084
	3	1.591	2.341	3.216	1.335	2.054	2.929
	4	1.570	2.321	3.196	1.283	2.033	2.908
	5	1.569	2.320	3.194	1.281	2.030	2.907
7	2	1.747	2.465	3.340	1.601	2.319	3.194
	3	1.690	2.440	3.315	1.438	2.157	3.032
	4	1.680	2.430	3.305	1.389	2.139	3.014
	5	1.667	2.417	3.292	1.370	2.120	2.995
8	2	1.830	2.549	3.424	1.684	2.403	3.278
	3	1.772	2.491	3.366	1.529	2.248	3.123
	4	1.758	2.475	3.350	1.468	2.212	3.087
	5	1.755	2.474	3.347	1.462	2.181	3.056
9	2	1.906	2.625	3.500	1.766	2.453	3.328
	3	1.841	2.560	3.435	1.609	2.296	3.171
	4	1.829	2.549	3.423	1.541	2.260	3.135
	5	1.822	2.541	3.416	1.526	2.245	3.120
10	2	1.956	2.675	3.550	1.837	2.525	3.337
	3	1.884	2.603	3.478	1.666	2.385	3.198
	4	1.871	2.590	3.470	1.601	2.319	3.194
	5	1.869	2.587	3.464	1.585	2.304	3.179
15	2	2.175	2.862	3.737	2.056	2.744	3.494
	3	2.101	2.820	3.695	1.894	2.582	3.394
	4	2.086	2.798	3.673	1.825	2.542	3.391
	5	2.080	2.791	3.660	1.794	2.513	3.388
20	2	2.321	3.008	3.821	2.207	2.895	3.645
	3	2.245	2.933	3.808	2.045	2.732	3.482
	4	2.218	2.905	3.780	1.968	2.656	3.468
	5	2.213	2.900	3.775	1.935	2.622	3.449

Table I (cont.)

k	m \ d	1.3			1.6		
		P*			P*		
		.75	.90	.975	.75	.90	.975
3	2	0.559	1.340	2.215	0.464	1.246	2.121
4	2	0.943	1.662	2.537	0.884	1.634	2.509
	3	0.547	1.297	2.172	0.365	1.115	1.990
5	2	1.178	1.897	2.772	1.130	1.849	2.724
	3	0.805	1.524	2.399	0.678	1.397	2.272
	4	0.609	1.359	2.234	0.389	1.107	1.982
6	2	1.326	2.044	2.906	1.308	1.995	2.870
	3	0.992	1.679	2.554	0.889	1.576	2.451
	4	0.783	1.502	2.377	0.610	1.329	2.204
	5	0.697	1.447	2.322	0.431	1.181	2.056
7	2	1.457	2.145	3.020	1.442	2.130	2.942
	3	1.127	1.814	2.627	1.047	1.735	2.485
	4	0.930	1.649	2.524	0.772	1.459	2.334
	5	0.828	1.547	2.422	0.600	1.318	2.193
8	2	1.556	2.244	3.056	1.544	2.231	2.981
	3	1.234	1.922	2.734	1.168	1.855	2.605
	4	1.048	1.736	2.611	0.918	1.606	2.356
	5	0.925	1.644	2.519	0.724	1.442	2.255
9	2	1.645	2.332	3.082	1.619	2.307	3.059
	3	1.327	2.015	2.765	1.269	1.957	2.707
	4	1.141	1.829	2.641	1.020	1.707	2.457
	5	1.023	1.711	2.586	0.850	1.537	2.350
10	2	1.725	2.412	3.162	1.706	2.394	3.144
	3	1.407	2.049	2.844	1.367	2.024	2.774
	4	1.219	1.907	2.657	1.117	1.805	2.555
	5	1.101	1.789	2.603	0.947	1.635	2.385
15	2	1.967	2.655	3.405	1.952	2.640	3.390
	3	1.692	2.349	3.099	1.659	2.315	3.065
	4	1.516	2.173	2.923	1.449	2.106	2.856
	5	1.388	2.075	2.825	1.284	1.941	2.691
20	2	2.138	2.794	3.544	2.116	2.772	3.522
	3	1.871	2.527	3.277	1.842	2.498	3.248
	4	1.695	2.351	3.101	1.644	2.301	3.051
	5	1.578	2.234	2.984	1.492	2.149	2.898

Table II

This table gives $n(d)/n(+\infty)$ where $n(a)$ is the sample size necessary for the rule R: select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[k]} - a\sigma/\sqrt{n},\ \bar{X}_{[k-m+1]}\}$ to satisfy $P[CS] \geq P^* \ \forall \mu \in \Omega(\delta)$.

k	m (d →)	.4		.7		1.3		1.6	
		P*		P*		P*		P*	
		.90	.975	.90	.975	.90	.975	.90	.975
3	2	2.999	1.999	2.270	1.659	1.506	1.282	1.302	1.175
4	2	1.895	1.615	1.527	1.389	1.161	1.156	1.123	1.131
	3	7.660	3.488	5.598	2.813	3.016	1.901	2.229	1.596
5	2	1.598	1.473	1.370	1.319	1.120	1.147	1.064	1.108
	3	3.459	2.480	2.634	2.015	1.623	1.402	1.363	1.311
	4	17.615	5.392	13.389	4.091	6.625	2.793	4.396	2.198
6	2	1.490	1.419	1.265	1.261	1.083	1.119	1.039	1.092
	3	2.595	2.093	1.997	1.735	1.334	1.319	1.176	1.215
	4	5.706	3.358	4.384	2.783	2.394	1.860	1.873	1.598
7	2	1.388	1.358	1.228	1.242	1.051	1.110	1.036	1.054
	3	2.234	1.921	1.746	1.607	1.235	1.207	1.129	1.080
	4	3.912	2.773	3.030	2.306	1.802	1.618	1.411	1.383
8	2	1.348	1.336	1.198	1.224	1.046	1.064	1.033	1.013
	3	1.983	1.784	1.615	1.535	1.181	1.177	1.101	1.069
	4	3.081	2.413	2.462	2.050	1.515	1.466	1.297	1.193
9	2	1.325	1.323	1.157	1.196	1.045	1.026	1.023	1.011
	3	1.859	1.714	1.495	1.461	1.151	1.110	1.086	1.064
	4	2.691	2.227	2.129	1.876	1.394	1.331	1.214	1.152

Table III

Using the rule R and under the configuration $(\alpha,\alpha+\delta,\ldots,\alpha+(k-1)\delta)$ this table gives in order the triple a) the expected number of selected populations, b) the expected sum of ranks of the selected populations and c) the expected proportion of selected populations ((a) divided by m)

Number of Populations Studied

k = 3

m	d \ $\sqrt{\frac{n}{\sigma}}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3111	1.2800	1.2237	1.1649	1.1156
		2.5300	2.1262	1.7606	1.4906	3.3121
		0.6555	0.6400	0.6118	0.5825	0.5578
	.7	1.5039	1.4588	1.3751	1.2839	1.2038
		2.9134	2.4731	2.0451	1.7073	1.4698
		0.7520	0.7294	0.6875	0.6420	0.6019

k = 4

m	d \ $\sqrt{\frac{n}{\sigma}}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3619	1.3090	1.2316	1.1660	1.1157
		3.2056	2.3924	1.8184	1.4971	1.3124
		0.6810	0.6545	0.6158	0.5830	0.5578
	.7	1.5691	1.4972	1.3862	1.2855	1.2039
		3.7113	2.8056	2.1237	1.7172	1.4704
		0.7845	0.7486	0.6931	0.6427	0.6020
3	.4	1.4391	1.3629	1.2568	1.1750	1.1183
		3.3970	2.5213	1.8765	1.5173	1.3183
		0.4797	0.4543	0.4189	0.3917	0.3728
	.7	1.7789	1.6483	1.4611	1.3139	1.2126
		4.2343	3.1766	2.3037	1.7845	1.4910
		0.5930	0.5494	0.4870	0.4380	0.4042

k = 5

m	d \ $\sqrt{\frac{n}{\sigma}}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3956	1.3208	1.2326	1.1660	1.1157
		3.8362	2.5299	1.8277	1.4973	1.3124
		0.6978	0.6604	0.6163	0.5830	0.5578
	.7	1.6097	1.5119	1.3875	1.2855	1.2039
		4.4502	2.9794	2.3170	1.7175	1.4704
		0.8048	0.7560	0.6938	0.6428	0.6020
3	.4	1.4995	1.3845	1.2588	1.1751	1.1183
		4.1402	2.6964	1.8893	1.5176	1.3183
		0.4998	0.4615	0.4196	0.3917	0.3728
	.7	1.8785	1.6862	1.4650	1.3140	1.2125
		5.2408	3.4475	2.3276	1.7851	1.4910
		0.6262	0.5621	0.4884	0.4380	0.4042
4	.4	1.5165	1.3920	1.2601	1.1752	1.1183
		4.1910	2.7184	1.8932	1.5180	1.3183
		0.3791	0.3480	0.3150	0.2938	0.2796
	.7	1.9571	1.7230	1.4724	1.3148	1.2126
		5.4774	3.5593	2.3499	1.7875	1.4912
		0.4893	0.4308	0.3681	0.3287	0.3031

Table IV

Using the rule R and under the configuration $(\alpha,\alpha,\ldots,\alpha+\delta)$ the table gives in order the triple a) the expected number of selected populations, b) the expected sum of ranks of the selected populations and c) the expected proportion of selected populations ((a) divided by m)

Number of Populations Studied

k = 3

m	d \ $\frac{\sqrt{n}}{\sigma}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3120	1.2996	1.2702	1.2270	1.1766
		2.5773	2.3611	2.1156	1.8627	1.6259
		0.6560	0.6498	0.6351	0.6135	0.5883
	.7	1.5052	1.4872	1.4437	1.3783	1.3003
		2.9629	2.7352	2.4657	2.1740	1.8861
		0.7526	0.7436	0.7219	0.6892	0.6502

k = 4

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3641	1.3529	1.3243	1.2792	1.2233
		3.3491	3.0598	2.7192	2.3554	2.0028
		0.6821	0.6765	0.6622	0.6396	0.6116
	.7	1.5720	1.5568	1.5169	1.4523	1.3696
		3.8654	3.5571	3.1877	2.7804	2.3685
		0.7860	0.7784	0.7585	0.7261	0.6848
3	.4	1.4423	1.4266	1.3877	1.3288	1.2583
		3.5441	3.2426	2.8768	2.4792	2.0908
		0.4808	0.4755	0.4626	0.4429	0.4194
	.7	1.7844	1.7578	1.6920	1.5915	1.4701
		4.3959	4.0606	4.6299	3.1363	2.6292
		0.5948	0.5859	0.5640	0.5305	0.4900

k = 5

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3993	1.3893	1.3622	1.3172	1.2587
		4.1254	4.7752	3.3491	2.8800	2.4125
		0.6997	0.6947	0.6811	0.6586	0.6294
	.7	1.6145	1.6015	1.5653	1.5033	1.4198
		4.7653	4.3894	3.9297	3.4130	2.8799
		0.8072	0.8007	0.7827	0.7516	0.7099
3	.4	1.5055	1.4904	1.4512	1.3887	1.3108
		4.4422	4.0725	3.6089	3.0886	2.5649
		.5018	.4968	.4837	.4629	.4369
	.7	1.8882	1.8635	1.7988	1.6946	1.5627
		5.5835	5.1660	4.6218	3.9837	3.3107
		0.6294	0.6212	0.5996	0.5649	0.5209
4	.4	1.5230	1.5067	1.4649	1.3990	1.3177
		4.4949	4.1216	3.6502	3.1198	2.5859
		0.3808	0.3767	0.3662	0.3498	0.3294
	.7	1.9692	1.9392	1.8631	1.7437	1.5964
		5.8267	5.3950	4.8181	4.1356	3.4161
		0.4923	0.4848	0.4658	0.4359	0.3991

REFERENCES

[1] Alam, K. and Rizvi, M. H. (1966). Selection from Multivariate Normal Populations. Ann. Inst. Statist. Math. 18, 307-318.

[2] 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.

[3] Desu, M. M. and Sobel, M. (1968). A fixed subset-size approach to the selection problem. Biometrika 55, 401-410.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill, N.C.

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

[6] Lehmann, E. L. (1959). Testing Statistical Hypotheses, John Wiley, New York.

[7] Mahamunulu, D. M. (1966). On a generalized goal in fixed sample ranking and selection problems. Technical Report No. 72, Department of Statistics, University of Minnesota.

[8] Mahamunulu, D. M. (1967). Some fixed sample ranking and selection problems. Ann. Math. Stat. 38, 1079-91.

[9] Nagel, K. (1970). On subset selection rules with certain optimality properties. Mimeograph Series No. 222, Department of Statistics, Purdue University.