A Restricted Subset Selection Approach to Ranking and Selection Problems*

by

Thomas J. Santner

Purdue University

Department of Statistics
Division of Mathematical Sciences

Mimeograph Series #322

June 1973

*This research was supported by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University.

A Restricted Subset Selection Approach to

Ranking and Selection Problems*

by

Thomas J. Santner
Purdue University

1. Introduction and Summary. Let $(\mathscr{X}, \mathscr{B}, P_i)$, $i = 1,\ldots,k$ be k probability spaces hereafter referred to as populations and denoted as π_i, $i = 1,\ldots,k$. Specifically it is assumed $\mathscr{X}$ is a finite dimensional Euclidean space, $\mathscr{B}$ is the associated Borel sigma field and P_i is an unknown probability measure belonging to a specified family of probability measures, $\mathscr{P}$. Each π_i is characterized by an unknown scalar $\lambda_i = \lambda_i(P_i) \varepsilon \Lambda$ a known interval on the real line. Let $\lambda_{[1]} \leq \ldots \leq \lambda_{[k]}$ be the ordered λ_i's, $\Omega = \{\underset{\sim}{\lambda} = (\lambda_1,\ldots,\lambda_k) | \lambda_i \varepsilon \Lambda \forall_i\}$ the space of all possible underlying configurations of λ_i's and $\pi_{(i)}$ the (unknown) population with parameter $\lambda_{[i]}$. It is assumed there is no a priori knowledge of the correct pairing of the elements in $\{\pi_i\}$ and $\{\pi_{(i)}\}$. The goal is to define a procedure R to select the "best" population where for sake of definiteness $\pi_{(k)}$ is taken to be the best population. In some cases $\pi_{(1)}$ might be the best population. Of course if $T(2 \leq T \leq k)$ populations all have $\lambda_i = \lambda_{[k]}$, the selection of any of these tied populations accomplishes the goal.

This ranking and selection problem was formulated as a multiple decision problem and specific cases solved by early research workers. The theory in this field has undergone a somewhat dichotomous development arising from the detailed formulation of a reasonable experimental goal to pursue.

*This research was supported by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

One approach pioneered by Bechhofer (1954) has been to allow the experimenter to select one population which is guaranteed to be $\pi_{(k)}$ with at least probability P* whenever the unknown parameters lie outside some subset, or zone of indifference, of the entire parameter space. This has been termed the indifference zone approach. A variety of authors have contributed papers employing this approach and the monograph by Bechhofer, Kiefer and Sobel (1968) contains an extensive bibliography. In particular the procedure of Mahamunulu (1967) for selecting a fixed size subset of size m which contains at least c of the t best populations employs this approach.

In contrast to the indifference zone approach, Gupta (1956, 65) proposed a formulation, called the subset selection approach, in which the experimenter obtains a subset of the k populations for which there is fixed minimum probability P* over the entire parameter space that the best population is included. The procedure selects a random number of populations between one and k, the actual number depending on the data. A few recent contributors in this area are Panchapakesan (1969, 1971), Gupta and Nagel (1971), McDonald (1972) and Huang (1972). A unified account of some of the general theory can be found in Gupta and Panchapakesan (1972).

The goal in this paper is to study single sample procedures which give more flexibility to the experimenter than does either the fixed subset size rule or the 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 that a particular population is best, this type of rule retains the advantage of the subset selection procedure over the fixed size subset rule in allowing selection of fewer than m populations. On the other hand if the data make the choice of the best population less obvious, this rule selects a larger subset for further study but guarantees that no more than m populations are selected. Such procedures will be called restricted subset selection procedures.

2. Formulation of the Problem. Each π_i yields iid observations $\{X_{ij}\}_{j=1}^{\infty}$ which are also independent between populations. X_{ij} has cdf F_i corresponding to $P_i \varepsilon \mathcal{P}$ which is now assumed to be a parametric family. Furthermore it is assumed there exists a sequence of Borel measurable functions $\{T_n\}$ so that T_n is defined on $\mathcal{X}^n$ and

$$T_n(X_{i1},\ldots,X_{in}) = T_{in} \overset{p}{\to} \lambda_i \text{ as } n\to\infty.$$

In practice it suffices to assume $\{T_{in}\}$ converges to a monotone function of λ_i so that the resulting selection problem is equivalent to the original one. The assumptions concerning T_{in} are that its cdf $G_n(y|\lambda_i)$ with support $E_n^{\lambda_i}$ depends on F_i only through λ_i and is absolutely continuous with respect to Lesbesgue measure with pdf $g_n(y|\lambda_i)$. Also for each n it is assumed $\{G_n(y|\lambda)|\lambda\varepsilon\Lambda\}$ forms a stochastically increasing family.

An indifference zone will be defined in Ω by means of a function $p:\Lambda \to R$ such that

$$(2.1)\begin{cases} \text{(i)} & p(\cdot) \text{ is continuous and non decreasing on } \Lambda \\ \text{(ii)} & p(\lambda) < \lambda \;\forall\; \lambda \,\varepsilon\, \Lambda \\ \text{(iii)} & p:\Lambda' \overset{\text{onto}}{\to} \Lambda \text{ where } \Lambda' = \{\lambda\varepsilon\Lambda|p(\lambda)\varepsilon\Lambda\}. \end{cases}$$

Define

$$\Omega(p) = \{\underset{\sim}{\lambda}\varepsilon\Omega|\lambda_{[k-1]} \le p(\lambda_{[k]})\}$$

$$\Omega^o(p) = \{\underset{\sim}{\lambda}\varepsilon\Omega|\lambda_{[1]} = \lambda_{[k-1]} = p(\lambda_{[k]})\}.$$

The subspace $\Omega(p)$ represents those vectors of λ_i's for which the best and second best populations are sufficiently far apart so that the experimenter desires to insure detection of the best one with high probability. $\Omega(p)$ is called the preference zone, its complement the indifference zone and $\Omega^o(p)$ contains the so called least favorable configurations in $\Omega(p)$.

Example 2.1

$p(\lambda) = \lambda-\delta(\delta > 0) \Rightarrow \Omega(p) = \{\underset{\sim}{\lambda} | \lambda_{[k]}-\lambda_{[k-1]} \geq \delta\}$, a location type preference zone.

Remark 2.1. Since the emphasis in this paper is on the case $1 < m < k$ the strict inequality $p(\lambda) < \lambda$ insures that the indifference zone does not vanish. However it should be noted that the general theory formally reduces to give the results of Bechhofer and Gupta for the choices $m = 1$ and $m = k$ respectively if the weaker $p(\lambda) \leq \lambda$ is allowed.

Finally, a general procedure for selecting a restricted subset of the k populations will be defined. Let $\{h_n(\cdot)\}$ be a sequence of functions such that each $h_n(\cdot)$: $E_n \rightarrow R'$ where $\bigcup_{\lambda\varepsilon\Lambda} E_n^\lambda \subset E_n$ and satisfies

(2.2)
- (i) For each n and x, $h_n(x) > x$
- (ii) For each n, $h_n(x)$ is continuous and strictly increasing in x
- (iii) For each x, $h_n(x) \rightarrow x$ as $n\rightarrow\infty$.

Define the procedure:

(2.3) R(n): Select $\pi_i \Leftrightarrow T_{in} \geq \max\{T_{[k-m+1]n}, h_n^{-1}(T_{[k]n})\}$ where $T_{[1]n} \leq T_{[2]n} \leq \cdots \leq T_{[k]n}$ are the ordered estimators.

Example 2.2. For $h_n(x) = x + d/\sqrt{n} \Rightarrow$

$$R(n): \text{ Select } \pi_i \Leftrightarrow T_{in} \geq \max\{T_{[k-m+1]n}, T_{[k]n}-d/\sqrt{n}\} .$$

Goal Given P*, p(·) and the sequence {R(n)} find the common sample size n necessary to achieve

(2.4) $$P_{\underset{\sim}{\lambda}}[CS|R(n)] \geq P^* \;\; \forall \, \underset{\sim}{\lambda} \varepsilon \Omega(p)$$

The event [CS|R(n)] occurs iff the selected subset contains $\pi_{(k)}$.

Theorem 2.1. For any $\underset{\sim}{\lambda} \varepsilon \Omega$

(2.5) $$P_{\underset{\sim}{\lambda}}[CS|R(n)] = \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \int \prod_{j \varepsilon \mathscr{P}_\nu^p(k)} G_n^{(j)}(y) \prod_{j \varepsilon \bar{\mathscr{P}}_\nu^p(k)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\} dG_n^{(k)}(y)$$

where

$$\begin{cases} \{\mathscr{P}_\nu^p(i) \,|\, \nu = 1, \ldots, \binom{k-1}{p}\} \text{ is the collection of all subsets of size p from } U(i) = \{1,\ldots,k\} \setminus \{i\} \\ \bar{\mathscr{P}}_\nu^p(i) = U(i) \setminus \mathscr{P}_\nu^p(i) \\ G_n^{(j)}(y) = G_n(y|\lambda_{[j]}) \end{cases}$$

Proof. Let $T_{(i)}$ be the random variable corresponding to $\pi_{(i)}$ and

$$A_\nu^p = [T_{(k)} > T_{(j)} \; \forall \, j \varepsilon \mathscr{P}_\nu^p(k), \; T_{(k)} < T_{(j)} \forall j \varepsilon \bar{\mathscr{P}}_\nu^p(k)].$$

$$\begin{aligned} P_{\underset{\sim}{\lambda}}[CS|R(n)] &= P_{\underset{\sim}{\lambda}}[h_n(T_{(k)}) \geq T_{[k]}, \; T_{(k)} \geq T_{[k-m+1]}] \\ &= P_{\underset{\sim}{\lambda}}[h_n(T_{(k)}) \geq T_{[k]} \text{ and } T_{(k)} > \text{at least } (k-m) \; T_{(j)}\text{'s w/ } j < k] \\ &= \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} P_{\underset{\sim}{\lambda}}[h_n(T_{(k)}) \geq T_{[k]}, \; A_\nu^p] \\ &= \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} P_{\underset{\sim}{\lambda}}[T_{(k)} > T_{(j)} \; \forall j \varepsilon \mathscr{P}_\nu^p(k), \; T_{(k)} < T_{(j)} < h_n(T_{(k)}) \; \forall j \varepsilon \bar{\mathscr{P}}_\nu^p(k)] \end{aligned}$$

from which (2.5) can be immediately derived.

3. Infimum of the Probability of Correct Selection. The calculation of the infimum of the probability of a correct selection will be accomplished in two stages. In the first stage the k dimensional infimum will be reduced to a one dimensional infimum and in the second stage conditions will be given which allow final evaluation. The following lemma due to Mahamunulu (1967) and Alam and Rizvi (1966) will be needed

Lemma 3.1. Let $\underset{\sim}{X} = (X_1,\ldots,X_k)$ have $k \geq 1$ independent components such that for every i, X_i has cdf $H(\cdot|\theta_i)$. Suppose $\{H(\cdot|\theta)\}$ forms a stochastically increasing family. If $\phi(\underset{\sim}{X})$ is a montone function of X_i when all other components of $\underset{\sim}{X}$ are held fixed then $E_{\underset{\sim}{\theta}}[\phi(\underset{\sim}{X})]$ is monotone in θ_i in the same direction.

Let

$$I(y;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^y w^{a-1} (1-w)^{b-1}\, dw$$

denote the incomplete beta function with parameters a and b.

Theorem 3.1.

$$\inf_{\Omega(p)} P_{\underset{\sim}{\lambda}}[CS|R(n)] = \inf_{\Omega^o(p)} P_{\underset{\sim}{\lambda}}[CS|R(n)] = \inf_{\lambda\varepsilon\Lambda'} \psi(\lambda,n) \text{ where}$$

$$\psi(\lambda,n) = \int \{G_n(h_n(y)|p(\lambda))\}^{k-1}\, I\left(\frac{G_n(y|p(\lambda))}{G_n(h_n(y)|p(\lambda))};k-m,m\right) dG_n(y|\lambda)$$

Proof. It suffices to show $\forall\, \underset{\sim}{\lambda}\varepsilon\Omega$, $P_{\underset{\sim}{\lambda}}[CS|R(n)] \geq \inf_{\lambda\varepsilon\Lambda} \psi(\lambda,n)$.

Define $$\phi(\underset{\sim}{T}) = \begin{cases} 1, & T_{(k)} \geq \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\} \\ 0, & \qquad < \end{cases}$$

where the n is suppressed **for** ease of notation and then $P_{\underset{\sim}{\lambda}}[CS|R(n)] = E_{\underset{\sim}{\lambda}}[\phi(\underset{\sim}{T})]$. Recall that $T_{(i)}$ is the statistic corresponding to $\pi_{(i)}$. By Lemma 3.1 it suffices to show $\phi(T) \downarrow$ in $T_{(\ell)}$ $\forall \ell < k$. Suppose

$\underset{\sim}{T}' \ni T'_{(\ell)} > T_{(\ell)}$ and $T'_{(j)} = T_{(j)}$ $\forall j \nmid \ell$ and $\phi(\underset{\sim}{T}) = 0$. Now $\phi(\underset{\sim}{T}) = 0$

$$\Leftrightarrow \quad T_{(k)} < \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\}$$

$$\Leftrightarrow \begin{cases} \text{(a)} & h_n(T_{(k)}) < T_{[k]} \\ & \text{or} \\ \text{(b)} & T_{(k)} < T_{[k-m+1]} \end{cases}$$

If (a) holds $\Rightarrow$ either

$$\begin{cases} (1) & T'_{(\ell)} \leq T_{[k]} \Rightarrow T'_{[k]} = T_{[k]} \Rightarrow h_n(T'_{(k)}) = h_n(T_{(k)}) < T_{[k]} = T'_{[k]} \\ & \text{or} \\ (2) & T'_{(\ell)} > T_{[k]} \Rightarrow T'_{[k]} = T'_{[\ell]} \Rightarrow h_n(T'_{(k)}) < T_{[k]} < T'_{(k)}. \end{cases}$$

In either case $h_n(T'_{(k)}) < T'_{[k]} \leq \max\{T'_{[k]}, h_n(T'_{[k-m+1]})\} \Rightarrow \phi(\underset{\sim}{T}') = 0$.

A similar argument shows (b) also implies $\phi(\underset{\sim}{T}') = 0$.

So we get

$$P_{\underset{\sim}{\lambda}}[CS|R(n)] \geq \sum_{p=k-m}^{k-1} \binom{k-1}{p} \int \{G_n(y|p(\lambda_{[k]}))\}^p \{G_n(h_n(y)|p(\lambda_{[k]}) - G_n(y|p(\lambda_{[k]}))\}^{k-1-p}$$

$$dG_n(y|\lambda_{[k]})$$

$$= \psi(\lambda_{[k]}, n) \geq \inf_{\lambda \varepsilon \Lambda'} \psi(\lambda, n) \text{ which completes the proof.}$$

If $\psi(\lambda,n)$ is monotone (increasing say) in λ and there exists a smallest $\lambda_0 \varepsilon \Lambda'$ then the k dimensional infimum will be completely evaluated as

$$\inf_{\Omega(p)} P_{\underset{\sim}{\lambda}}[CS|R(n)] = \psi(\lambda_0, n).$$

The following two lemmas give sufficient conditions for such behavior. The first is due to Gupta and Panchapakesan (1972).

Lemma 3.2. Let $F(\cdot|\lambda)|\lambda\varepsilon\Lambda\}$ be a family of absolutely continuous distributions on the real line with continuous densities $f(\cdot|\lambda)$ and $\psi(x,\lambda)$ a bounded real valued function possessing first partial derivatives ϕ_x and ϕ_λ wrt x and λ respectively and satisfying regularity conditions (3.2). Then $E_\lambda[\phi(x,\lambda)]$ is non decreasing in λ provided for all $\lambda\varepsilon\Lambda$

(3.1) $$f(x/\lambda)\ \frac{\partial\phi(x,\lambda)}{\partial\lambda} - \frac{\partial F(x/\lambda)}{\partial\lambda}\ \frac{\partial\phi(x,\lambda)}{\partial x} \geq 0 \quad \text{for a.e.x.}$$

(3.2)

(i) For all $\lambda\varepsilon\Lambda$, $\frac{\partial\phi(x,\lambda)}{\partial x}$ is Lesbesgue integrable on R

(ii) For every $[\lambda_1,\lambda_2] \subset \Lambda$ and $\lambda_3\varepsilon\Lambda$ there exists h(x) depending only on λ_i, i=1,2,3 such that

$$\left|\frac{\partial\phi(x,\lambda)}{\partial\lambda}\ f(x|\lambda_3) - \frac{\partial F(x|\lambda)}{\partial\lambda}\ \frac{\partial\phi(x,\lambda_3)}{\partial x}\right| \leq h(x)\ \forall\,\lambda\varepsilon[\lambda_1,\lambda_2]$$

and h(x) is Lesbesgue integrable on R .

Also a straightforward computation shows the following result.

Lemma 3.3. For any $1 \leq \ell < n$ and $0 \leq a < c \leq 1$

$$n\ c\ I(a/c;\ell,n-\ell+1) \geq ab(a/c) \text{ where } b(y) = I'(y;\ell,n-\ell+1)$$

Remark 3.1. The following assumptions are essentially needed to insure that (3.2) holds in the following theorem. For any $[\lambda_1,\lambda_2] \subset \Lambda'$ and $\lambda_3\varepsilon\Lambda'$ there exist $e_1(y)$ and $e_2(y)$ such that

$$
(3.3)\quad \begin{cases}
\text{(i)} & \left|\dfrac{\partial G_n(y|p(\lambda))}{\partial\lambda}\right| \le e_1(y)\ \forall\ \lambda\varepsilon[\lambda_1,\lambda_2] \text{ where} \\
& (\int e_1(y)\,dG_n(y|\lambda_3))(\int e_1(h_n(y))\,dG_n(y|\lambda_3)) < \infty \\
\text{(ii)} & \left|\dfrac{\partial G_n(y|\lambda)}{\partial\lambda}\right| \le e_2(y)\ \ \forall\lambda\varepsilon[\lambda_1,\lambda_2] \text{ where} \\
& (\int e_2(y)\,dG_n(h_n(y)|\lambda_3))(\int e_2(y)\,dG_n(y|\lambda_3)) < \infty
\end{cases}
$$

Theorem 3.2. If $E_n^\lambda = E_n\ \forall\lambda\varepsilon\Lambda'$, $G_n(y|\lambda)$ is continuously differentiable and satisfies (3.3) and all derivatives in (3.4) and (3.5) exist and $\forall\lambda\varepsilon\Lambda'$

$$
(3.4)\qquad g_n(y|\lambda)\ \frac{\partial G_n(h_n(y)|p(\lambda))}{\partial\lambda} - h_n'(y)\,g_n(h_n(y)|p(\lambda))\ \frac{\partial G_n(y|\lambda)}{\partial\lambda} \ge 0 \text{ ae in } E_n
$$

$$
(3.5)\qquad g_n(y|\lambda)\ \frac{\partial G_n(y|p(\lambda))}{\partial\lambda} - g_n(y|p(\lambda))\ \frac{\partial G_n(y|\lambda)}{\partial\lambda} \ge 0 \text{ ae in } E_n
$$

then $\psi(\lambda,n)$ is nondecreasing in λ.

Proof. As indicated, the proof is an application of Lemma 3.2. Note that

$\psi(\lambda,n) = \int_{E_n}\phi(y,\lambda)\,dG_n(y|\lambda)$ for the choice

$$
\phi(y,\lambda) = \{G_n(h_n(y)|p(\lambda))\}^{k-1}\ I\left(\frac{G_n(y|p(\lambda))}{G_n(h_n(y)|p(\lambda))};k-m,m\right). \quad \text{Hence}
$$

$$
\begin{aligned}
\frac{\partial\phi(y,\lambda)}{\partial y} &= (k-1)\{G_n(h_n(y)|p(\lambda))\}^{k-2}\ g_n(h_n(y)|p(\lambda))h_n'(y)I(K_n(y,\lambda);k-m,m) \\
&\quad + \{G_n(h_n(y)|p(\lambda))\}^{k-3}\ b(K_n(y,\lambda))\{G_n(h_n(y)|p(\lambda))g_n(y|p(\lambda)) \\
&\qquad - G_n(y|p(\lambda))h_n'(y)\cdot g_n(h_n(y)|p(\lambda))\}
\end{aligned}
$$

$$\frac{\partial\phi(y,\lambda)}{\partial\lambda} = (k-1)\{G_n(h_n(y)|p(\lambda))\}^{k-2}\frac{\partial G_n(h_n(y)|p(\lambda))}{\partial\lambda} I(K_n(y,\lambda);k-m,m) + \{G_n(h_n(y)|p(\lambda))\}^{k-3} b(K_n(y,\lambda))\cdot\{G_n(h_n(y)|p(\lambda))\frac{\partial G_n(y|p(\lambda))}{\partial\lambda} - G_n(y|p(\lambda))\frac{\partial G_n(h_n(y)|p(\lambda))}{\partial\lambda}\}$$

where $K_n(y,\lambda) = \frac{G_n(y|p(\lambda))}{G_n(h_n(y)|p(\lambda))}$.

So (3.1) becomes; $\forall\lambda\varepsilon\Lambda'$

(3.6) $$g_n(y|\lambda)[(k-1)\frac{G_n(h_n(y)|p(\lambda))}{\partial\lambda}\{G_n(h_n(y)|p(\lambda))\}I(K_n(y,\lambda);k-m,m)+b(K_n(y,\lambda))\{G_n(h_n(y)|p(\lambda))\frac{\partial G_n(y|p(\lambda))}{\partial\lambda} - G_n(y|p(\lambda))\frac{\partial G_n(h_n(y)|p(\lambda))}{\partial\lambda}\}] - \frac{\partial G_n(y|\lambda)}{\partial\lambda}[(k-1)G_n(h_n(y)|p(\lambda))g_n(h_n(y)|p(\lambda))h_n'(y)I(K_n(y,\lambda);k-m,m) + b(K_n(y,\lambda))\{G_n(h_n(y)|p(\lambda))g_n(y|p(\lambda))-h_n'(y)G_n(y|p(\lambda))g_n(h_n(y)|p(\lambda))\}] \geq 0 \text{ ae in } E_n$$

By rearranging terms (3.6) can be seen to hold if $\forall\lambda\varepsilon\Lambda'$

(3.7) $$\{g_n(y|\lambda)\frac{\partial G_n(y|p(\lambda))}{\partial\lambda} - \frac{\partial G_n(y|\lambda)}{\partial\lambda}g_n(y|p(\lambda))\} \geq 0 \text{ ae in } E_n$$

and

(3.8) $$\{g_n(y|\lambda)\frac{\partial G_n(h_n(y)|p(\lambda))}{\partial\lambda} - h_n'(y)g_n(h_n(y)|p(\lambda))\frac{\partial G_n(y|\lambda)}{\partial\lambda}\} \times \{(k-1)I(K_n(y,\lambda);k-m,m)G_n(h_n(y)|p(\lambda))-b(K_n(y,\lambda))G_n(y|p(\lambda))\} \geq 0 \text{ ae in } E_n$$

But by Lemma 3.3 the second factor in (3.8) is non negative since $\forall y\varepsilon E^n$ and $\lambda\varepsilon\Lambda' \Rightarrow 0 \leq G_n(y|p(\lambda)) \leq G_n(h_n(y)|p(\lambda)) \leq 1$. Hence (3.8) and (3.7) reduce to (3.4) and (3.5). Similar arguments show that (3.3) imply the regularity conditions required for Lemma 3.2 and hence completes the proof.

Remark 3.2. The proofs of Theorem 3.2 and Lemma 3.2 also show that if (3.4) and (3.5) are identically zero than $\psi(\lambda,n)$ is independent of λ and if (3.4) and (3.5) are non positive then $\psi(\lambda,n)$ is non increasing in λ.

4. Properties of $\{R(n)\}$. Both the properties of the sequence $\{R(n)\}$ and the individual rules $R(n)$ will be studied. For $\underset{\sim}{\lambda}\varepsilon\Omega$ let

$$(4.1) \qquad p^n_{\underset{\sim}{\lambda}}(i) = P_{\underset{\sim}{\lambda}}[R(n) \text{ selects } \pi_{(i)}].$$

Def. 4.1. The sequence of rules $\{R(n)\}$ is consistent wrt Ω' means $\inf_{\Omega'} P[CS|R(n)] \to 1$ as $n\to\infty$.

Def. 4.2. The rule $R(n)$ is strongly monotone in $\pi_{(i)}$ means

$$p^n_{\underset{\sim}{\lambda}}(i) \text{ is } \begin{cases} \uparrow \text{ in } \lambda_{[i]} \text{ when all other components of } \underset{\sim}{\lambda} \text{ are fixed} \\ \downarrow \text{ in } \lambda_{[j]}\ (j \neq i) \text{ when all other components of } \underset{\sim}{\lambda} \text{ are fixed.} \end{cases}$$

Theorem 4.1. If there exists $N \geq 1$ and $\lambda_0\varepsilon\Lambda'$ such that $\forall\, n \geq N$ $\inf_{\lambda\varepsilon\Lambda'} \psi(\lambda,n) = \psi(\lambda_0,n)$, then any sequence $\{R(n)\}$ defined by (2.3) is consistent wrt any $\Omega(p)$.

Proof. From the hypothesis of the theorem and the result of Theorem 3.1 we have $\forall n \geq N$

$$\inf_{\Omega(p)} P_{\underset{\sim}{\lambda}}[CS|R(n)] = \int \nu(y,\lambda_0)\, dG_n(y|\lambda_0) \text{ where}$$

$$(4.2) \qquad \nu(y,\lambda) = \{G_n(h_n(y)|p(\lambda))\}^{k-1}\, I\left(\frac{G_n(y|p(\lambda))}{G_n(h_n(y)|p(\lambda))};\ k\text{-}m,m\right).$$

Also $T_{in} \overset{p}{\to} \lambda_i$ as $n\to\infty \Leftrightarrow G_n(y|\lambda_i) \to \begin{cases} 0\ , & y < \lambda_i \\ 1\ , & y > \lambda_i . \end{cases}$

Since $\psi(\lambda_0,n) \leq 1$, it suffices to show $\forall\ 1 > \epsilon' > 0\ \exists M \ni \forall n \geq M$

$$\int \nu(y,\lambda_0)\, dG_n(y|\lambda_0) > 1 - \epsilon' .$$

Since $p(\lambda_0) < \lambda_0 \exists\, \alpha \ni p(\lambda_0) < \alpha < \lambda_0$. Given $1 > \epsilon' > 0$ let $\epsilon = 1 - \sqrt{1-\epsilon'}$ and choose $M > N \ni$

(a) $G_n(\alpha|\lambda_0) < \epsilon$ (since $\alpha < \lambda_0$)

(b) $\{G_n(h_n(\alpha)|p(\lambda_0))\}^{k-1} I(G_n(\alpha|p(\lambda_0));k-m,m) > 1-\epsilon$ (since $h_n(\alpha) > \alpha > p(\lambda_0)$)

So $\forall y > \alpha$

$$\begin{cases} \text{(a)} & 1 \geq G_n(h_n(y)|p(\lambda_0)) \geq G_n(h_n(\alpha)|p(\lambda_0)) \\ \text{(b)} & G_n(y|p(\lambda_0)) \geq G_n(\alpha|p(\lambda_0)) \end{cases}$$

which implies that $\forall y > \alpha$

$$V(y,\lambda_0) \geq \{G_n(h_n(\alpha)|p(\lambda_0))\}^{k-1} I\left(\frac{G_n(\alpha|p(\lambda_0))}{G_n(h_n(y)|p(\lambda_0))};\ k-m,m\right) \geq 1-\epsilon.$$

So finally $\forall n \geq M$

$$\begin{aligned} \int V(y,\lambda_0)dG_n(y|\lambda_0) &\geq \int_\alpha^\infty V(y,\lambda_0)dG_n(y|\lambda_0) \\ &\geq (1-\epsilon)\int_\alpha^\infty dG_n(y|\lambda_0) \\ &\geq 1-\epsilon' \text{ and the proof is completed.} \end{aligned}$$

Remark 4.1. Theorem 4.1 shows that any (P^*,p) requirement can be met by choosing a sufficiently large common sample size n.

Theorem 4.2. Any rule R(n) of form (2.3) is strongly monotone in $\pi_{(i)}$ for any $i = 1,\ldots,k$.

Proof. Since $p_{\underset{\sim}{\lambda}}(i) = E_{\underset{\sim}{\lambda}}[\eta_i(\underset{\sim}{T})]$ where

$$\eta_i(\underset{\sim}{T}) = \begin{cases} 1 & , \quad T_{(i)} \geq \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\} \\ 0 & , \quad \text{otherwise} \end{cases}$$

the result of Lemma 3.1 can again be used to show the desired monotonicity. Arguments similar to those in the proof of Theorem 2.1 show that

(A) $\eta_i(\underset{\sim}{T})$ is non increasing in $T_{(j)}$ $(j \neq i)$ when all other components of $\underset{\sim}{T}$ are fixed

(B) $\eta_i(\underset{\sim}{T})$ is non decreasing in $T_{(i)}$ when all other components of $\underset{\sim}{T}$ are fixed

and hence complete the proof.

Gupta (1965) has proved that the subset selection rule which he studied possessed the properties of monotonicity and unbiasedness. Recall these definitions.

Def. 4.3. The rule R is monotone means $\forall 1 \leq i < j \leq k$ and $\underset{\sim}{\lambda} \varepsilon \Omega$

$P_{\underset{\sim}{\lambda}}[R \text{ selects } \pi_{(j)}] \geq P_{\underset{\sim}{\lambda}}[R \text{ selects } \pi_{(i)}]$

Def. 4.4. The rule unbiased means $\forall 1 \leq i < k$ and $\underset{\sim}{\lambda} \varepsilon \Omega$

$P_{\underset{\sim}{\lambda}}[R \text{ does not select } \pi_{(i)}] \geq P[R \text{ does not select } \pi_{(k)}]$

Corollary 4.1. All rules R(n) in the class defined by (2.3) are monotone and unbiased.

Proof. Since monotonicity implies unbiasedness it suffers to show that $p^n_{\underset{\sim}{\lambda}}(i) \leq p^n_{\underset{\sim}{\lambda}}(i+1)$ for any $i = 1,\ldots,k-1$ and $\underset{\sim}{\lambda} \varepsilon \Omega$. Assuming wlog that $\lambda_i = \lambda_{[i]}$ for notational ease it follows

$$p^n_{\underset{\sim}{\lambda}}(i) = p^n_{(\lambda_1,\ldots,\lambda_k)}(i)$$

$\leq p^n_{(\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1},\lambda_{i+1},\lambda_{i+2},\ldots,\lambda_k)}(i)$ since $p^n_{\underset{\sim}{\lambda}}(i)$ is $\uparrow$ in $\lambda_{[i]}$

$= p^n_{(\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1},\lambda_{i+1},\lambda_{i+2},\ldots,\lambda_k)}(i+1)$ since both $\pi_{(i)}$ and $\pi_{(i+1)}$ have the same cdf.

$\leq p^n_{(\lambda_1,\ldots,\lambda_{i-1},\lambda_i,\lambda_{i+1},\ldots,\lambda_k)}(i+1)$ since $p^n_{\underset{\sim}{\lambda}}(i+1)$ $\downarrow$ in $\lambda_{[i]}$

$= p^n_{\underset{\sim}{\lambda}}(i+1)$. This completes the proof of the corollary.

Remark 4.2. The above proof shows that any rule which is strongly monotone in $\pi_{(i)}$ for $i = 1,\ldots,k$ is monotone.

5. Number of Selected Populations. If $S(n)$ is the number of populations selected by $R(n)$, $T(n)$ is the number of non best populations selected by $R(n)$ and $p^n_{\underset{\sim}{\lambda}}(i)$ is defined as before by (4.1) and if

$$W_i(n) = \begin{cases} 0 , & T_{(i)} \geq \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\} \\ 1 , & \text{otherwise} \end{cases}$$

then the following representations hold

(5.1) $$S(n) = \sum_{i=1}^{k} W_i(n)$$

(5.2) $$T(n) = \sum_{i=1}^{k-1} W_i(n)$$

(5.3) $$p^n_{\underset{\sim}{\lambda}}(i) = E_{\underset{\sim}{\lambda}}[W_i(n)].$$

Theorem 5.1. For any $\underset{\sim}{\lambda} \varepsilon \Omega$

$$E_{\underset{\sim}{\lambda}}[S(n)] = \sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{j \varepsilon \mathcal{P}^p_\nu(i)} G_n^{(j)}(y) \prod_{j \varepsilon \bar{\mathcal{P}}^p_\nu(i)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\} dG_n^{(i)}(y)$$

Proof: Using (5.1) and an argument similar to that in the proof of Theorem (2.1) it can be seen that

$$E_{\underset{\sim}{\lambda}}[S(n)] = \sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} P_{\underset{\sim}{\lambda}}[h_n(T_{(i)}) \geq T_{(j)}\ \forall j \neq i,\ T_{(\ell)} > T_{(j)}\ \forall j \varepsilon \mathcal{P}^p_\nu(i),$$
$$T_{(i)} < T_{(j)}\ \forall j \varepsilon \bar{\mathcal{P}}^p_\nu(i)$$
$$= \sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} P_{\underset{\sim}{\lambda}}[T_{(i)} > T_{(j)} \forall j \varepsilon \mathcal{P}^p_\nu(i),\ T_{(i)} < T_{(j)} < h_n(T_{(j)} \forall j \varepsilon \bar{\mathcal{P}}^p_\nu(i)]$$

which gives the result.

Remark 5.1. The expected value of T(n) can be derived in a manner similar to the above using (5.2).

In the remainder of the section two topics will be studied:

(a) Asymptotic properties of the sequences {S(n)} and {T(n)}.

(b) The supremum of $E_{\underset{\sim}{\lambda}}[S(n)]$ and $E_{\underset{\sim}{\lambda}}[T(n)]$ over Ω.

Theorem 5.2. For any $\underset{\sim}{\lambda} \ni \lambda_{[k]} > \lambda_{[k-1]}$

$$(5.4) \qquad p^n_{\underset{\sim}{\lambda}}(i) \to \left\{\begin{matrix} 1 & , & i = k \\ 0 & , & 1 \le i < k \end{matrix}\right\} \quad \text{as } n\to\infty$$

Proof: Recall that $G^{(i)}_n(y) \to \left\{\begin{matrix} 1 & , & y > \lambda_{[i]} \\ 0 & , & y < \lambda_{[i]} \end{matrix}\right\}$ as $n\to\infty$ and pick $\alpha \ni$ $\lambda_{[k-1]} < \alpha < \lambda_{[k]}$. Let

$$f^{p,\nu}_i(y) = \prod_{j\in\mathscr{P}^p_\nu(i)} G^{(j)}_n(y) \prod_{j\in\bar{\mathscr{P}}^p_\nu(i)} \{G^{(j)}_n(h_n(y)) - G^{(j)}_n(y)\}$$

Case A: $i = k$

$$p^n_{\underset{\sim}{\lambda}}(k) = \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \int f^{p,\nu}_k(y)\, dG^{(k)}_n(y)$$

Subcase (1): For $k - m \le p \le k - 2$ and $1 \le \nu \le \binom{k-1}{p}$

$\Rightarrow \bar{\mathscr{P}}^p_\nu(k) \neq \phi$ and so $\forall y \in R$,

$$f^{p,\nu}_k(y) \le \prod_{j\in\bar{\mathscr{P}}^p_\nu(k)} \{G^{(j)}_n(h_n(y)) - G^{(j)}_n(y)\} \le \prod_{j\in\bar{\mathscr{P}}^p_\nu(k)} \{1-G^{(j)}_n(y)\}$$

Given $\epsilon > 0$ pick $N \ni \forall n \geq N$

$$\begin{cases} G_n^{(k)}(\alpha) < \epsilon/2 \\ \text{and} \\ \prod_{j \in \bar{\mathscr{S}}_\nu^p(k)} (1-G_n^{(j)}(\alpha)) < \epsilon/2 \end{cases}$$

So $\forall n \geq N$, $0 \leq \int f_k^{p,\nu}(y)\, dG_n^{(k)}(y)$

$$= \int_{-\infty}^{\alpha} f_k^{p,\nu}(y)\, dG_n^{(k)}(y) + \int_{\alpha}^{\infty} f_k^{p,\nu}(y)\, dG_n^{(k)}(y)$$

$$\leq \int_{-\infty}^{\alpha} 1\; dG_n^{(k)}(y) + \int_{\alpha}^{\infty} \epsilon/2\; dG_n^{(k)}(y)$$

$$\leq \epsilon$$

<u>Subcase (2)</u>: $p=k-1$ and $\nu=1$

Using the fact that $\mathscr{S}_\nu^{k-1}(k) = \{1,\ldots,k-1\}$ and $\bar{\mathscr{S}}_1^{k-1}(k) = \phi$ and an argument similar to that in the proof of Theorem 4.1 it can be proved that $\int f_k^{k-1,1}(y)\, dG_n^{(k)}(y) \to 1$ as $n\to\infty$.

<u>Case B: $1 \leq i \leq k-1$</u>

It suffices to show that $\int f_i^{p,\nu}(y)\, dG_n^{(i)}(y) \to 0$ as $n\to\infty$ $\forall$ p and ν.

<u>Subcase (1)</u>: For p and ν such that $k \in \mathscr{S}_\nu^p(i)$

Again using a straightforward argument of the above type the desired result follows.

<u>Subcase (2)</u>: For p and ν such that $k \in \bar{\mathscr{S}}_\nu^p(i)$.

Pick α' such that $\alpha < \alpha' < \lambda_{[k]}$. Now since $h_n(\alpha) \to \alpha$ and $\alpha' < \lambda_{[k]}$ $\exists$ N such that $\forall n \geq N$

$$\begin{cases} G_n^{(k)}(\alpha') < \epsilon/2 \\ h_n(\alpha) < \alpha' \\ G_n^{(i)}(\alpha) > 1 - \epsilon/2 \end{cases}$$

$\Rightarrow \forall n \geq N$ and $y < \alpha$

$$\begin{aligned} f_i^{p,\nu}(y) &\leq \{G_n^{(k)}(h_n(y)) - G_n^{(k)}(y)\} \\ &\leq G_n^{(k)}(h_n(y)) \\ &\leq G_n^{(k)}(h_n(\alpha)) \leq G_n^{(k)}(\alpha') < \epsilon/2. \end{aligned}$$

Finally we again obtain that $\int f_i^{p,\nu}(y)dG_n^{(i)}(y) < \epsilon \,\forall n \geq N$. This completes the proof of the theorem.

Corollary 5.1. For any $\underset{\sim}{\lambda}\varepsilon\Omega \ni \lambda_{[k]} > \lambda_{[k-1]}$

(5.5) $$W_i(n) \overset{p}{\to} \begin{cases} 1 \;, & i = k \\ 0 \;, & 1 \leq i \leq k-1 \end{cases} \quad \text{as } n\to\infty$$

Proof: For any $\epsilon > 0$, $P_{\underset{\sim}{\lambda}}[|W_k(n)-1| > \epsilon] \leq P_{\underset{\sim}{\lambda}}[W_k(n)=0] = 1-p_{\underset{\sim}{\lambda}}^n(k) \to 0$ as $n\to\infty$ and for $i < k$, $P_{\underset{\sim}{\lambda}}[|W_i(n)| > \epsilon] \leq p_{\underset{\sim}{\lambda}}^n(i) \to 0$ as $n\to\infty$ by (5.4).

Remark 5.2. Since all random variables studied in this section are uniformly bounded it follows that convergence in L^2 and probability are equivalent.

Using (5.1), (5.2) and $(S(n)-1) \leq (S(n)-1)^2$ together with the convergence in probability of the $W_i(n)$ random variables we obtain

Corollary 5.2. For $\underset{\sim}{\lambda}\varepsilon\Omega$ such that $\lambda_{[k]} > \lambda_{[k-1]}$

(1) $S(n) \overset{p}{\to} 1$ and $T(n) \overset{p}{\to} 0$ as $n\to\infty$ and hence

(2) $E_{\underset{\sim}{\lambda}}[S(n)] \to 1$ and $E_{\underset{\sim}{\lambda}}[T(n)] \to 0$ as $n\to\infty$.

The next results will study some properties of $S(n)$ when n is fixed. In particular, conditions will be given which guarantee that the supremum of $E_{\underset{\sim}{\lambda}}[S(n)]$ in Ω occurs at some point $\underset{\sim}{\lambda} = (\lambda_1,\ldots,\lambda_k)$ for which $\lambda_{[1]} = \lambda_{[k]}$. The condition (5.6) will be assumed in some of the theorems which follow.

(5.6)
$$\begin{cases} \text{(i)} & E_n^{\lambda} = E_n \quad \forall \lambda \varepsilon \Lambda \\ \text{(ii)} & \text{For any } [\lambda_1,\lambda_2] \subset \Lambda \text{ there exists } e_3(y) \text{ depending only} \\ & \lambda_1 \text{ and } \lambda_2 \ni \left| \dfrac{\partial G_n(y|\lambda)}{\partial \lambda} \right| \le e_3(y) \text{ where} \end{cases}$$

$$(\int e_3(y)\,dG_n(h_n(y)|\lambda'))(\int e_3(h_n(y))\,dG_n(y|\lambda')) < \infty \quad \forall \lambda' \ge \lambda_2$$

Theorem 5.3. If (5.6) is satisfied and $\forall \lambda_1$, λ_2 in Λ with $\lambda_1 \le \lambda_2$

(5.7)
$$\frac{\partial G_n(h_n(y)|\lambda_1)}{\partial \lambda_1} g_n(y|\lambda_2) - \frac{\partial G_n(y|\lambda_1)}{\partial \lambda_1} g_n(h_n(y)|\lambda_2)h_n'(y) \ge 0 \text{ ae in } E_n$$

then $E_{\underset{\sim}{\lambda}}[S(n)]$ is non decreasing in $\lambda_{[1]}$ on $\Lambda(\lambda_{[2]}) = \{\lambda\varepsilon\Lambda | \lambda \le \lambda_{[2]}\}$ for any fixed $(\lambda_{[2]},\ldots,\lambda_{[k]})$.

Proof. Fix $\lambda_{[2]} \le\cdots\le \lambda_{[k]}$ for the following argument and then

$E_{\underset{\sim}{\lambda}}[S(n)] = T_1(\underset{\sim}{\lambda}) + T_2(\underset{\sim}{\lambda})$ where

$$T_1(\underset{\sim}{\lambda}) = \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \int_{E_n} f_1^{p,\nu}(y)\,dG_n^{(1)}(y)$$

$$T_2(\underset{\sim}{\lambda}) = \sum_{r=2}^{k} \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \int_{E_n} f_r^{p,\nu}(y)\,dG_n^{(r)}(y) \text{ where}$$

$f_i^{p,\nu}(y)$ is defined as the proof of Theorem 5.3.

Now $T_2(\underset{\sim}{\lambda})$ can be rewritten as

$$T_2(\underset{\sim}{\lambda}) = \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \sum_{r \ni 1 \varepsilon \mathcal{P}_\nu^p(r)} \int_{E_n} f_r^{p,\nu}(y)dG_n^{(r)}(y)$$

$$+ \sum_{p=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{p}} \sum_{r \ni 1 \varepsilon \bar{\mathcal{P}}_\nu^p(r)} \int_{E_n} f_r^{p,\nu}(y)dG_n^{(r)}(y).$$

For any $A \subset \{1,\ldots,k\}$ of size s, let $\{\mathcal{P}_\nu^p(A) \mid \nu=1,\ldots,\binom{k-s}{p}\}$ be the collection of all subsets of size p from $\{1,\ldots,k\} - A$. Note that for any fixed $p=k-m,\ldots,k-1$ and $r=2,\ldots,k$

$$\{\mathcal{P}_\nu^p(r) \mid 1 \varepsilon \mathcal{P}_\nu^p(r)\} = \{\mathcal{P}_\nu^{p-1}(1,r) \cup \{1\} \mid \nu=1,\ldots,\binom{k-2}{p-1}\} \tag{5.8}$$

while for any $p=k-m,\ldots,k-2$ and $r=2,\ldots,k$

$$\{\mathcal{P}_\nu^p(r) \mid 1 \not\varepsilon \mathcal{P}_\nu^p(r)\} = \{\mathcal{P}_\nu^p(1,r) \mid \nu=1,\ldots,\binom{k-2}{p}\}. \tag{5.9}$$

So

$$T_2(\underset{\sim}{\lambda}) = \sum_{p=k-m}^{k-1} \sum_{r=2}^{k} \sum_{\nu=1}^{\binom{k-2}{p-1}} \int_{E_n} W_r^{p,\nu}(y)G_n^{(1)}(y)dG_n^{(r)}(y)$$

$$+ \sum_{p=k-m}^{k-2} \sum_{r=2}^{k} \sum_{\nu=1}^{\binom{k-2}{p}} \int_{E_n} z_r^{p,\nu}(y)\{G_n^{(1)}(h_n(y))-G_n^{(1)}(y)\}dG_n^{(r)}(y)$$

where

(1) $$W_r^{p,\nu}(y) = \prod_{j \varepsilon \mathcal{P}_\nu^{p-1}(1,r)} G_n^{(j)}(y) \prod_{j \varepsilon \bar{\mathcal{P}}_\nu^{p-1}(1,r)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\}$$

(2) $$z_r^{p,\nu}(y) = \prod_{j \varepsilon \mathcal{P}_\nu^{p}(1,r)} G_n^{(j)}(y) \prod_{j \varepsilon \bar{\mathcal{P}}_\nu^{p}(1,r)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\}.$$

Next integrating $T_1(\lambda)$ by parts and noting that for fixed p=k-m,...,k-1 and r=2,...,k

$$(5.10)\qquad \{\mathscr{S}_\nu^p(1)\,|\,r\varepsilon\mathscr{S}_\nu^p(1)\} = \{\mathscr{S}_\nu^{p-1}(1,r)U\{r\}\,|\,\nu=1,\dots,\binom{k-2}{p-1}\}$$

while for any p=k-m,...,k-2 and r=2,...,k

$$(5.11)\qquad \{\mathscr{S}_\nu^p(1)\,|\,r\varepsilon\mathscr{T}_\nu^p(1)\} = \{\mathscr{S}_\nu^p(1,r)\,|\,\nu=1,\dots,\binom{k-2}{p}\}$$

we obtain that

$$T_1(\lambda) = 1 - \sum_{p=k-m}^{k-1}\sum_{r=2}^{k}\sum_{\nu=1}^{\binom{k-2}{p-1}} \int_{E_n} W_r^{p,\nu}(y)G_n^{(1)}(y)dG_n^{(r)}(y)$$

$$- \sum_{p=k-m}^{k-1}\sum_{r=2}^{k}\sum_{\nu=1}^{\binom{k-2}{p}} \int_{E_n} z_r^{p,\nu}(y)G_n^{(1)}(y)\{g_n^{(r)}(h_n(y))h_n'(y)-g_n^{(r)}(y)\}dy.$$

Hence combining and cancelling terms it follows that

$$E_{\lambda}[S(n)] = 1 + \sum_{p=k-m}^{k-2}\sum_{r=2}^{k}\sum_{\nu=1}^{\binom{k-2}{p}} \int_{E_n} z_r^{p,\nu}(y)\{G_n^{(1)}(h_n(h))g_n^{(r)}(y)-G_n^{(1)}(y)g_n^{(r)}(h_n(y))h_n'(y)\}dy$$

and finally

$$(5.12)\qquad \frac{dE_{\lambda}[S(n)]}{d\lambda_{[1]}} = \sum_{p=k-m}^{k-2}\sum_{r=2}^{k}\sum_{\nu=1}^{\binom{k-2}{p}} \int_{E_n} z_r^{p,\nu}(y)\cdot$$

$$\{\frac{\partial G_n^{(1)}(h_n(y))}{\partial\lambda_{[1]}} g_n^{(r)}(y) - \frac{\partial G_n^{(1)}(y)}{\partial\lambda_{[1]}} g_n^{(r)}(h_n(y))h_n'(y)\}dy.$$

But (5.7) gives for every r=2,...,k

$$\frac{\partial G_n^{(1)}(h_n(y))}{\partial\lambda_{[1]}} g_n^{(r)}(y) - \frac{\partial G_n^{(1)}(y)}{\partial\lambda_{[1]}} g_n^{(r)}(h_n(y))h_n'(y) \geq 0 \text{ a.e. in } E_n$$

$\Rightarrow$ the derivative in (5.12) is non negative and completes the proof.

Remark 5.3. Condition (5.7) is essentially the same requirement as that made by Sobel (1969) and Gupta and Panchapakesan (1972) in order to show that $\sup_{\Omega} E[S]$ be attained for their rules when the distributions are identical. In location or scale parameter problems it reduces to the requirement of MLR.

Corollary 5.3. If for every fixed $\lambda_{[2]} \leq \cdots \leq \lambda_{[k]}$, $\dfrac{dE_{\lambda}[S(n)]}{d\lambda_{[1]}} \geq 0$ for $\lambda_{[1]}$ in $\Lambda(\lambda_{[2]})$, then the $\sup_{\Omega} E_{\underset{\sim}{\lambda}}[S(n)] = \sup_{\lambda\in\Lambda} \gamma(\lambda,n)$ where

$$(5.13) \qquad \gamma(\lambda,n) = k\int_{E_n^{\lambda}} \{G_n(h_n(y)|\lambda)\}^{k-1} I\left(\frac{G_n(y|\lambda)}{G_n(h_n(y)|\lambda)};k-m,m\right) dG_n(y|\lambda)$$

Furthermore if the hypotheses of Theorem 5.3 hold for $\lambda_1=\lambda_2$ then $\gamma(\lambda,n)$ is non decreasing in λ and hence if there is a greatest element $\lambda_0 \varepsilon \Lambda \Rightarrow \sup_{\Omega} E_{\underset{\sim}{\lambda}}[S(n)] = \gamma(\lambda_0,n)$.

Proof. It suffices to prove $\forall q < k$ and fixed $\lambda_{[q+1]} \leq \cdots \leq \lambda_{[k]}$ that $E_{\underset{\sim}{\lambda}(q)}[S(n)] \uparrow$ in λ on $\Lambda(\lambda_{[q+1]})$ where the underlying $\underset{\sim}{\lambda}(q) = (\lambda,\ldots,\lambda,\lambda_{[q+1]},\ldots,\lambda_{[k]})$. Let $\underset{\sim}{\lambda}' = (\lambda_{[1]},\ldots,\lambda_{[k]})$ and note from Theorem 5.1 that $E_{\underset{\sim}{\lambda}'}[S(n)]$ is invariant under permutations of the elements in $\underset{\sim}{\lambda}'$. So

$$\frac{dE_{\underset{\sim}{\lambda}(q)}[S(n)]}{d\lambda} = \sum_{i=1}^{q} \left.\frac{\partial E_{\underset{\sim}{\lambda}'}[S(n)]}{\partial\lambda_{[i]}}\right|_{\underset{\sim}{\lambda}(q)}$$

$$= \left.\frac{q\partial E_{\underset{\sim}{\lambda}'}[S(n)]}{\partial\lambda_{[1]}}\right|_{\underset{\sim}{\lambda}(q)}$$

But from the previous proof $\left.\dfrac{\partial E_{\underset{\sim}{\lambda}'}[S(n)]}{\partial\lambda_{[1]}}\right|_{\underset{\sim}{\lambda}(q)} \geq 0$.

Hence the supremum over Ω of $E[S(n)]$ occurs at some point where all the $\lambda_{[i]}$'s are equal.

Since $\gamma(\lambda,n) = E[\phi(Y,\lambda)]$ for

$$\phi(y,\lambda) = \{G_n(h_n(y)|\lambda)\}^{k-1}\ I\left(\frac{G_n(y|\lambda)}{G_n(h_n(y)|\lambda)};k-m,m\right),$$

Lemma 3.2 can be applied and the sufficient condition (3.1) that $\gamma(\lambda,n)$ be non decreasing reduces to

$$\{G_n(h_n(y)|\lambda)\}^{k-3}\ \{g_n(y|\lambda)\ \frac{\partial G_n(h_n(y)|\lambda)}{\partial\lambda} - \frac{\partial G_n(y|\lambda)}{\partial\lambda}\ g_n(h_n(y)|\lambda)h_n'(y)\}.$$

$$\{(k-1)G_n(h_n(y)|\lambda)\ I\left(\frac{G_n(y|\lambda)}{G_n(h_n(y)|\lambda)};k-m,m\right) - G_n(y|\lambda)b\left(\frac{G_n(y|\lambda)}{G_n(h_n(y)|\lambda)}\right)\} \geq 0 \quad \forall\lambda \text{ and ae } y$$

$$\Leftrightarrow\ \{g_n(y|\lambda)\ \frac{\partial G_n(h_n(y)|\lambda)}{\partial\lambda} - \frac{\partial G_n(y|\lambda)}{\partial\lambda}\ g_n(h_n(y)|\lambda)h_n'(y)\} \geq 0 \quad \forall\lambda \text{ and ae } y$$

since the third factor is non negative by Lemma 3.3. The final part of the result is obvious.

Remark 5.4. While the hypotheses of Theorem 5.3 imply those of Corollary 5.3 in the regular case, these hypotheses are also satisfied in some non regular problems, for example, in selection from uniform populations.

Note that the expected number of non best populations selected can be written in the form

$$(5.14) \qquad E_{\underset{\sim}{\lambda}}[T(n)] = E_{\underset{\sim}{\lambda}}[S(n)] - p^n_{\underset{\sim}{\lambda}}(k).$$

Corollary 5.4. If the hypotheses of Corollary 5.3 hold then

$\sup\limits_{\Omega} E_{\underset{\sim}{\lambda}}[T(n)] = \frac{(k-1)}{k} \sup\limits_{\lambda\in\Lambda} \gamma(\lambda,n)$ where $\gamma(\lambda,n)$ is defined by (5.13).

Proof. For any $\underset{\sim}{\lambda} = (\lambda_{[1]},\ldots,\lambda_{[k]})\in\Omega$ let $\underset{\sim}{\lambda}([k]) = (\lambda_{[k]},\ldots,\lambda_{[k]})$ then $\forall\,\underset{\sim}{\lambda}\in\Omega$ the hypotheses imply $E_{\underset{\sim}{\lambda}}[S(n)] \leq E_{\underset{\sim}{\lambda}([k])}[S(n)]$. Also the strong monotonicity of R(n) implies $p^n_{\underset{\sim}{\lambda}}(k) \geq p^n_{\underset{\sim}{\lambda}([k])}(k)$. So by (5.14)

$$E_{\underset{\sim}{\lambda}}[T(n)] \leq E_{\underset{\sim}{\lambda}([k])}[T(n)] = \frac{(k-1)}{k}\gamma(\lambda_{[k]},n)$$

$$\Rightarrow \sup_{\Omega} E_{\underset{\sim}{\lambda}}[T(n)] = \frac{(k-1)}{k}\sup_{\lambda\varepsilon\Lambda}\gamma(\lambda,n).$$

Remark 5.5. From Corollary 3.3 it follows that $\gamma(\lambda,n)$ is nondecreasing in λ if the hypotheses of Theorem 5.3 holds for $\lambda_1 = \lambda_2$.

6. Applications. In this section we apply the results of this paper to some problems of selecting from univariate and multivariate normal populations.

I. Suppose $\pi_i \sim N(\mu_i,\sigma^2)$, i=1,...,k where the common variance σ^2 is known and the experimenter is interested in selecting the population having largest μ_i. We take $T_{in} = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$ and then $\lambda_i = \mu_i$ and

$G_n(y|\lambda_i) = \Phi(\{n^{1/2}(y-\mu_i)\}/\sigma)$ where Φ is the cdf of a N(0,1) random variable.

Since this is a location parameter problem we take $p(\mu) = \mu-\delta(\delta > 0)$ and $h_n(x) = x+d\sigma/\sqrt{n}$ and obtain

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

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

Using Theorem 3.1 and Corollary 5.3 it can be seen that

$$(6.1)\quad \inf_{\Omega(p)} P[CS|R(n)] = \int_{-\infty}^{\infty}\{\Phi(y+d+\frac{\sqrt{n}\delta}{\sigma})\}^{k-1}\, I\left(\frac{\Phi(y+d)}{\Phi(y+d+\frac{\sqrt{n}\delta}{\sigma})};k-m,m\right)d\Phi(y)$$

$$(6.2)\quad \sup_{\Omega} E[S(n)] = k\int_{-\infty}^{\infty}\{\Phi(y+d)\}^{k-1}\, I\left(\frac{\Phi(y)}{\Phi(y+d)};k-m,m\right)d\Phi(y).$$

One choice of $\{R(n)\}$ can be made by setting the right hand side of (6.2) equal to $1 + \epsilon$ and solving for d. Having chosen the sequence $\{R(n)\}$, the proper sample size can be found by equating the right hand side of (6.1) to P* and solving for n. Additional details including comparison with the fixed size procedure Desu and Sobel (1968) and tables of constants required to implement the proposed procedure are given in Gupta and Santner (1972).

II. Now suppose π_i is p variate normal with mean vector $\underset{\sim}{\mu}_i$ and covariance matrix $\underset{\sim}{\Sigma}(N_p(\underset{\sim}{\mu}_i, \underset{\sim}{\Sigma}))$ for i=0,1,...,k. The common $\underset{\sim}{\Sigma}$ and μ_0 are both known and π_0 may be though of as a standard or control population. It is desired to select that population which is furthest away from π_0 in the sense of Mahalanobis distance so that $\lambda_i = (\underset{\sim}{\mu}_i - \underset{\sim}{\mu}_0)' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\mu}_i - \underset{\sim}{\mu}_0)$. Gupta (1966), Alam and Rizvi (1966) and Gupta and Studden (1970) have considered this problem. We take

$$T_{in} = (X_{ij} - \mu_0)' \, \Sigma^{-1}(X_{ij} - \mu_0)$$

$$p(\lambda) = \begin{cases} \lambda - \delta_1 \ (\delta_1 > 0) & , \quad 0 \le \lambda \le \delta_1\delta_2/(\delta_2 - 1) \\ \delta_2^{-1}\lambda \ (\delta_2 > 1) & , \quad \delta_1\delta_2/(\delta_2 - 1) \le \lambda \end{cases}$$

$$h_n(x) = d^{1/n}\, x, \quad d > 1$$

$$F_p(x|\lambda) = e^{-\lambda/2} \sum_{j=0}^{\infty} \frac{\lambda^j}{2^j j!} E_{p+2j}(x) \text{ where } E_q(x) = \int_0^x \frac{y^{q/2-1} e^{-y/2}}{\Gamma(q/2)\, 2^{q/2}}\, dy$$

so that $T_{in} \overset{P}{\to} p + \lambda_i$ as $n \to \infty$

$$G_n(y|\lambda_i) = F_{np}(y|\lambda_i)$$

$$\Omega(p) = \Omega_1 \cap \Omega_2 \text{ where } \begin{cases} \Omega_1 = \{\underset{\sim}{\lambda} | \lambda_{[k]} - \lambda_{[k-1]} \ge \delta_1\} \\ \Omega_2 = \{\underset{\sim}{\lambda} | \lambda_{[k]} \ge \delta_2\, \lambda_{[k-1]}\}. \end{cases}$$

R(n): Select $\pi_i \Leftrightarrow T_{in} \geq \max\{T_{[k-m+1]n}, d^{-1/n}\, T_{[k]n}\}$

The following are known properties of $F_p(y|\lambda)$

$$\frac{\partial F_p(y|\lambda)}{\partial\lambda} = \frac{1}{2}\,[F_{p+2}(y|\lambda) - F_p(y|\lambda)] = -\, f_{p+2}(y|\lambda) \text{ where } f_p(y|\lambda) = \frac{dF_p(y|\lambda)}{dy},$$

$f_{p+2}(y|\lambda)/f_p(y|\lambda)$ ↓ in λ and $\{\lambda f_{p+2}(y|\lambda)\}/f_p(y|\lambda)$ ↑ in λ. In addition a result from Chapter 7 of Lehmann (1959) can be applied to show that $f_{p+2}(y|\lambda)/\{y f_p(y|\lambda)\}$ is non increasing in y and $y\lambda$.

Since $f_p(y|\lambda)$ has MLR, Theorem 3.1 can be applied to show that

$$\text{(6.3)} \qquad \inf_{\Omega(p)} P[CS|R(n)] = \inf_{\lambda \geq \delta_1} \psi(\lambda,n) \text{ where}$$

$$\psi(\lambda,n) = \begin{cases} \displaystyle\int_0^\infty \{F_{np}(yd^{1/n}|(\lambda-\delta_1))\}^{k-1}\, I\left(\frac{F_{np}(y|(\lambda-\delta_1))}{F_{np}(y|(\lambda-\delta_1))}; k-m, m\right) dF_{np}(y|\lambda), & \lambda\varepsilon I_1 \\ \displaystyle\int_0^\infty \{F_{np}(yd^{1/n}|\delta_2^{-1}\lambda)\}^{k-1}\, I\left(\frac{F_{np}(y|\delta_2^{-1}\lambda)}{F_{np}(yd^{1/n}|\delta_2^{-1}\lambda)}; k-m, m\right) dF_{np}(y|\lambda), & \lambda\varepsilon I_2 \end{cases}$$

where $I_1 = [\delta_1, \delta_1\delta_2/(\delta_2-1))$

$I_2 = [\delta_1\delta_2/(\delta_2-1), \infty)$

In this problem the one dimensional infimum $\psi(\lambda,n)$ is not independent of λ as was the case in the normal means problem. However for $1 < d < \delta_2$ and using the properties of the $f_p(y|\lambda)$ density listed above, a piecewise application of Theorem 3.2 on I_1 and I_2 shows that $\psi(\lambda,n)$ is ↓ in λ on I_1 and ↑ in λ on I_2 and hence

$$\text{(6.4)} \qquad \inf_{\Omega(p)} P[CS|R(n)] = \psi(\delta_1\delta_2/(\delta_2-1), n).$$

Theorem 4.1 applies since $\delta_1\delta_2/(\delta_2-1)\varepsilon\Lambda = [0,\infty)$ and hence

$\inf_{\Omega(p)} P_{\underset{\sim}{\lambda}}[CS|R(n)] \to 1$ as $n\to\infty$. All other usual properties hold for R(n) and in particular (5.7) holds (as verified by Panchapakesan (1969)) and hence

$$\sup_{\Omega} E[S(n)] = \sup_{\lambda\geq 0} \gamma(\lambda,n) \text{ and } \gamma(\lambda,n) \uparrow \text{ in } \lambda$$

where

$$\gamma(\lambda,n) = k\int_0^\infty \{F_{np}(yd^{1/n}|\lambda)\}^{k-1}\, I\left(\frac{F_{np}(y|\lambda)}{F_{np}(yd^{1/n}|\lambda)};k-m,m\right)dF_{np}(y|\lambda).$$

Using a probability argument this supremum can be evaluated as $\lim_{\lambda\to\infty}\gamma(\lambda,n)=m$. We obtain

$$(6.5) \qquad \sup_{\Omega} E[S(n)] = m.$$

Details of the above problem as well as other applications to regular and non regular problems can be found in Santner (1973).

Acknowledgement. My sincere thanks go to my dissertation advisor, Professor Shanti S. Gupta. His suggestions, criticisms and encouragement were sources of stimulation throughout the preparation of this paper.

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] Bechhofer, R. E., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures. Statistical Research Monographs, Vol. III, The University of Chicago Press, Chicago.

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

[5] 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.

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

[7] Gupta, S. S. (1966). On some selection and ranking procedures for multivariate normal populations using distance functions. In Multivariate Analysis (ed. P. R. Krishnaiah), pp. 457-475, Academic Press, N.Y.

[8] Gupta, S. S. and Nagel, K. (1971). On some contributions to multiple decision theory. In Statistical Decision Theory and Related Topics (ed. S. S. Gupta and J. Yackel), pp. 79-102, Academic Press, N.Y.

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

[10] Gupta, S. S. and Santner, T. J. (1972). Selection of a restricted subset of normal populations containing the one with largest mean. Mimeograph Series No. 299, Department of Statistics, Purdue University, Lafayette, Indiana.

[11] Gupta, S. S. and Studden, W. J. (1970). On some selection and ranking procedures with applications to multivariate populations. In Essays in Probability and Statistics (ed. R. C. Bose et al), pp. 327-338, University of North Carolina Press, Chapel Hill, N.C.

[12] Huang, W. T. (1972). Some contributions to sequential selection and ranking procedures. Mimeograph Series No. 295, Department of Statistics, Purdue University, Lafayette, Indiana.

[13] Lehmann, E. (1959). Testing Statistical Hypotheses. J. Wiley, New York.

[14] Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. Ann. Math. Statist. 38, 1079-1081.

[15] McDonald, G. C. (1972). Some multiple comparison selection procedures based on ranks. Sankyā Ser A. 34, 53-64.

[16] Panchapakesan, S. (1969). Some contributions to multiple decision (selection and ranking) procedures. Mimeograph Series No. 189, Department of Statistics, Purdue University, Lafayette, Indiana.

[17] Panchapakesan, S. (1971). On a subset selection procedure for the most probable event in a multinomial distribution. In Statistical Decision Theory and Related Topics (ed. S. S. Gupta and J. Yackel), pp. 275-298, Academic Press, N.Y.

[18] Santner, T. J. (1973). A restricted subset selection approach to ranking and selection problems. Mimeograph Series No. 318, Department of Statistics, Purdue University, Lafayette, Indiana.

[19] Sobel, M. (1969). Selecting a subset containing at least one of the t best populations. In Multivariate Analysis II (ed. P. R. Krishnaiah), pp. 515-540, Academic Press, N.Y.