The Annals of Statistics
1975, Vol. 3, No. 2, 334-349

A RESTRICTED SUBSET SELECTION APPROACH TO RANKING AND SELECTION PROBLEMS[1]

BY THOMAS J. SANTNER

Purdue University

Let $\pi_1, \cdots, \pi_k$ be k populations with π_i characterized by a scalar $\lambda_i \in \Lambda$, a specified interval on the real line. The object of the problem is to make some inference about $\pi_{(k)}$, the population with largest λ_i. The present work studies rules which select a random number of populations between one and m where the upper bound, m, is imposed by inherent setup restrictions on the experimenter. Formally the goal can be viewed as a generalization of the subset selection and indifference zone approaches. A selection procedure is defined in terms of a set of consistent sequences of estimators for the λ_i's. It is proved the infimum of the probability of a correct selection occurs at a point in the preference zone for which the parameters are as close together as possible. Conditions are given which allow evaluation of this last infimum. The number of non-best populations selected, the total number of populations selected, and their expectations are studied both asymptotically and for fixed n. Other desirable properties of the rule are also studied.

1. Introduction and summary. Let $(\mathscr{X}, \mathscr{B}, P_i)$, $i = 1, \cdots, k$ be k probability spaces hereafter referred to as populations and denoted as π_i, $i = 1, \cdots, k$. 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 $\lambda_i = \lambda(P_i) \in \Lambda$ a known interval on the real line. Let $\lambda_{[1]} \leqq \cdots \leqq \lambda_{[k]}$ be the ordered λ_i's, $\Omega = \{\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_k) \mid \lambda_i \in \Lambda \ \forall i\}$ 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 $(2 \leqq T \leqq k)$ 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 experiment goal to pursue. One approach

Received July 1973; revised January 1974.

[1] This paper forms a part of the author's doctoral dissertation submitted to Purdue University. This research was supported by the Office of Naval Research under contract N00014-A-0226-00014.

AMS 1970 *subject classifications.* Primary 62F07; Secondary 62G30.

Key words and phrases. Ranking and selection, restricted subset size, multiple decision.

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, 1965) 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 populations less obvious, this rule selects a larger subset but guarantees than no more than m populations are selected. Such procedures will be called *restricted subset selection procedures*.

In Section 2 the problem will be formalized and a class of procedures, $\{R(n)\}$ (one for each sample size n), proposed for its solution. The probability of a correct selection using $R(n)$ for arbitrary underlying $\boldsymbol{\lambda}$ is derived. In Section 3 a two-stage reduction is used to determine the infimum of the probability of a correct selection over the preference zone. Section 4 is devoted to a study of the properties of the sequence $\{R(n)\}$ and individual rules $R(n)$, while in Section 5 the number of populations and number of non-best populations selected by $R(n)$ are studied both asymptotically and for fixed n. In particular their supremum over Ω is derived. Some applications of the general theory are made in Section 6 to univariate normal populations for selection based on means and to multivariate normal populations for Maholanobis distance to a known control population. For the normal means problem, tables of the required sample sizes, expected number of selected populations, and comparisons with the fixed subset size procedures of Desu and Sobel (1968) can be found in Gupta and Santner (1973).

Finally it should be noted that optimality criteria for choosing the form of $R(n)$ or its defining statistic, T_n, are not considered here.

2. Formulation of the problem. Each π_i yields i.i.d. observations $\{X_{ij}\}$ which are also independent between populations. X_{ij} has cdf F_i corresponding to $P_i \in \mathscr{P}$ 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 $\mathscr{X}^n$ and

$$T_n(X_{i1}, \cdots, X_{in}) = T_{in} \to_p \lambda_i \qquad \text{as } n \to \infty .$$

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 $G_n(y \mid \lambda_i)$ with support $E_n^{\lambda_i}$ depends on F_i only through λ_i and is absolutely continuous with respect to Lebesgue measure with pdf $g_n(y \mid \lambda_i)$. Also for each n it is assumed $\{G_n(y \mid \lambda) \mid \lambda \in \Lambda\}$ forms a stochastically increasing family.

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

$$\text{(2.1)} \qquad \begin{aligned} &\text{(i)} \quad p(\cdot) \quad \text{is continuous and non-decreasing on} \quad \Lambda , \\ &\text{(ii)} \quad p(\lambda) < \lambda , \quad \lambda \in \Lambda , \\ &\text{(iii)} \quad p : \Lambda' \to_{\text{onto}} \Lambda \quad \text{where} \quad \Lambda' = \{\lambda \in \Lambda \mid p(\lambda) \in \Lambda\} . \end{aligned}$$

Define

$$\Omega(p) = \{\boldsymbol{\lambda} \in \Omega \mid \lambda_{[k-1]} \leqq p(\lambda_{[k]})\}$$
$$\Omega^0(p) = \{\boldsymbol{\lambda} \in \Omega \mid \lambda_{[1]} = \lambda_{[k-1]} = p(\lambda_{[k]})\} .$$

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 $\Omega^0(p)$ contains the so-called *least favorable configurations* in $\Omega(p)$.

EXAMPLE 2.1.

(1) $p_1(\lambda) = \lambda - \delta_1(\delta_1 > 0) \Rightarrow \Omega(p_1) = \{\lambda \mid \lambda_{[k]} - \lambda_{[k-1]} \geqq \delta_1\}$, a location type preference zone.

(2) $p_2(\lambda) = \delta_2^{-1}\lambda(\delta_2 > 1 \text{ and } \Lambda \subset (0, \infty)) \Rightarrow \Omega(p_2) = \{\boldsymbol{\lambda} \mid \lambda_{[k]} \geqq \delta_2\lambda_{[k-1]}\}$, a scale type preference zone.

(3)
$$\begin{aligned} p_3(\lambda) &= \lambda - \delta_1(\delta_1 > 0) , && 0 \leqq \lambda \leqq \delta_1\delta_2/(\delta_2 - 1) \\ &= \delta_2^{-1}\lambda(\delta_2 > 1) , && \lambda \geqq \delta_1\delta_2/(\delta_2 - 1) \end{aligned}$$

$(\Lambda = [0, \infty)) \Rightarrow \Omega(p_3) = \Omega(p_2) \cap \Omega(p_1)$, a mixed 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 choice $m = 1$ and $m = k$ respectively if the weaker $p(\lambda) \leqq \lambda$ 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 $h_n(\cdot): E_n \to R^1$ where $\bigcup_{\lambda \in \Lambda} E_n^\lambda \subset E_n$ and satisfies

$$\text{(2.2)} \qquad \begin{array}{ll} \text{(i)} & \text{For each } n \text{ and } x,\ h_n(x) > x . \\ \text{(ii)} & \text{For each } n,\ h_n(x) \text{ is continuous and strictly increasing in } x . \\ \text{(iii)} & \text{For each } x,\ h_n(x) \to x \text{ as } n \to \infty . \end{array}$$

Define the procedure:

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

EXAMPLE 2.2. For $h_n(x) = x + d/n^{\frac{1}{2}}$

$$R(n): \text{Select } \pi_i \Leftrightarrow T_{in} \geqq \max\{T_{[k-m+1]n}, T_{[k]n} - d/n^{\frac{1}{2}}\} .$$

Goal. Given P^*, $p(\cdot)$ and the sequence $R(n)$ find the common sample size n necessary to achieve

$$\text{(2.4)} \qquad P_{\boldsymbol{\lambda}}[CS \mid R(n)] \geqq P^* \quad \forall\, \boldsymbol{\lambda} \in \Omega(p) .$$

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

THEOREM 2.1. *For any* $\boldsymbol{\lambda} \in \Omega$

$$\text{(2.5)} \qquad P_{\boldsymbol{\lambda}}[CS \mid R(n)] = \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \int_{-\infty}^{\infty} \prod_{j \in \mathscr{S}_\nu^l(k)} G_n^{(j)}(y) \times \prod_{j \in \bar{\mathscr{S}}_\nu^l(k)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\}\, dG_n^{(k)}(y)$$

where

$\{\mathscr{S}_\nu^l(i) \mid \nu = 1, \cdots, \binom{k-1}{l}\}$ *is the collection of all subsets of size* l *from* $U(i) = \{1, \cdots, k\} - \{i\}$,

$\bar{\mathscr{S}}_\nu^l(i) = U(i) - \mathscr{S}_\nu^l(i)$,

$G_n^{(j)}(y) = G_n(y \mid \lambda_{[j]})$.

The proof of Theorem 2.1 is straightforward and omitted.

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.

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.

$$\inf_{\Omega(p)} P_{\boldsymbol{\lambda}}[CS \mid R(n)] = \inf_{\Omega^0(p)} P_{\boldsymbol{\lambda}}[CS \mid R(n)] = \inf_{\lambda \in \Lambda'} \psi(\lambda, n) \quad \textit{where}$$

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

PROOF. It suffices to show for all $\boldsymbol{\lambda} \in \Omega$, $P_{\boldsymbol{\lambda}}[CS \mid R(n)] \geqq \inf_{\lambda \in \Lambda'} \psi(\lambda, n)$. Define

$$\begin{aligned}\phi(\mathbf{T}) &= 1\,, \qquad T_{(k)} \geqq \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\} \\ &= 0\,, \qquad T_{(k)} < \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\}\end{aligned}$$

where the n is suppressed for ease of notation, and then $P_{\boldsymbol{\lambda}}[CS \mid R(n)] = E_{\boldsymbol{\lambda}}[\phi(\mathbf{T})]$. Now both $T_{[k-m+1]}$ and $T_{[k]}$ are non-decreasing functions of $T_{(l)}$ for $l < k$ when all other components of $(T_{(1)}, \cdots, T_{(k)})$ are fixed. This implies $\phi(\mathbf{T})$ is monotone in $T_{(l)}$ for $l < k$ if all other components of $(T_{(1)}, \cdots, T_{(k)})$ are fixed. The result now follows from an application of the lemma by Mahamunulu (1967) and Alam and Rizvi (1966).

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

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

The following lemma due to Gupta and Panchapakesan (1972) can be applied to give conditions for such monotone behavior.

LEMMA 3.1. *Let $F(\cdot \mid \lambda) \mid \lambda \in \Lambda\}$ be a family of absolutely continuous distributions on the real line with continuous densities $f(\cdot \mid \lambda)$ and $\phi(x, \lambda)$ 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 $\lambda \in \Lambda$*

$$(3.1) \qquad f(x \mid \lambda)\frac{\partial \phi(x, \lambda)}{\partial \lambda} - \frac{\partial F(x \mid \lambda)}{\partial \lambda}\frac{\partial \phi(x, \lambda)}{\partial x} \geqq 0 \qquad \textit{for} \quad \text{a.e.} \quad x\,,$$

and

$$(3.2) \qquad \text{(i)} \quad \textit{for all} \quad \lambda \in \Lambda\,, \quad \frac{\partial \phi(x, \lambda)}{\partial x} \quad \textit{is Lebesgue integrable on} \quad R\,;$$

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

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

and $h(x)$ *is Lebesgue integrable on* R.

REMARK 3.1. Even though not explicitly mentioned in Gupta and Panchapakesan's paper, regularity conditions are required in the proof of their result and (3.2) is one such set of conditions. The assumptions (3.3) on $G_n(y \mid \lambda)$ insure that (3.2) holds in our application of Lemma 3.1 to $\psi(\lambda, n)$. For any $[\lambda_1, \lambda_2] \subset \Lambda'$ and $\lambda_3 \in \Lambda'$ there exist $e_1(y)$ and $e_2(y)$ such that

$$(3.3) \qquad \text{(i)} \quad \left|\frac{\partial G_n(y \mid p(\lambda))}{\partial \lambda}\right| \leqq e_1(y) \quad \forall\, \lambda \in [\lambda_1, \lambda_2]$$

$$\text{where} \quad \left(\int e_1(y)\, dG_n(y \mid \lambda_3)\right)\left(\int e_1(h_n(y))\, dG_n(y \mid \lambda_3)\right) < \infty$$

$$\text{(ii)} \quad \left|\frac{\partial G_n(y \mid \lambda)}{\partial \lambda}\right| \leqq e_2(y) \quad \forall\, \lambda \in [\lambda_1, \lambda_2]$$

$$\text{where} \quad \left(\int e_2(y)\, dG_n(h_n(y) \mid \lambda_3)\right)\left(\int e_2(y)\, dG_n(y \mid \lambda_3)\right) < \infty\,.$$

THEOREM 3.2. *If $E_n^\lambda = E_n$ for all $\lambda \in \Lambda'$, if $G_n(y \mid \lambda)$ is continuously differentiable and satisfies* (3.3) *and all derivatives in* (3.4) *and* (3.5) *exist and if for all $\lambda \in \Lambda'$*

$$g_n(y \mid \lambda) \frac{\partial G_n(h_n(y) \mid p(\lambda))}{\partial \lambda} - h_n'(y) g_n(h_n(y) \mid p(\lambda)) \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} \geqq 0 \quad \text{a.e. } in \quad E_n\,, \tag{3.4}$$

$$g_n(y \mid \lambda) \frac{\partial G_n(y \mid p(\lambda))}{\partial \lambda} - g_n(y \mid p(\lambda)) \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} \geqq 0 \qquad \text{a.e. } in \quad E_n\,, \tag{3.5}$$

then $\psi(\lambda, n)$ is non-decreasing in λ.

PROOF. Note that $\psi(\lambda, n) = \int_{E_n} \phi(y, \lambda)\, dG_n(y \mid \lambda)$ for the choice

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

Hence

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

where

$$K_n(y, \lambda) = G_n(y \mid p(\lambda)) / G_n(h_n(y) \mid p(\lambda) \qquad \text{and}$$
$$b(y) = (k - m)\tbinom{k-1}{k-m} y^{k-m-1} (1 - y)^{m-1}\,.$$

So (3.1) becomes: for all $\lambda \in \Lambda'$

$$\begin{aligned} & g_n(y \mid \lambda)\Big[(k-1) \frac{G_n(h_n(y) \mid p(\lambda))}{\partial \lambda} \{G_n(h_n(y) \mid p(\lambda))\} I(K_n(y, \lambda); k - m, m) \\ & + b(K_n(y, \lambda)) \left\{G_n(h_n(y) \mid p(\lambda)) \frac{\partial G_n(y \mid p(\lambda))}{\partial \lambda}\right. \\ & \left.- G_n(y \mid p(\lambda)) \frac{\partial G_n(h_n(y) \mid p(\lambda))}{\partial \lambda}\right\}\Big] - \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} \\ & \times [(k-1) G_n(h_n(y) \mid p(\lambda)) g_n(h_n(y) \mid p(\lambda)) h_n'(y) I(K_n(y, \lambda); k - m, m) \\ & + b(K_n(y, \lambda))\{G_n(h_n(y) \mid p(\lambda)) g_n(y \mid p(\lambda)) \\ & - h_n'(y) G_n(y \mid p(\lambda)) g_n(h_n(y) \mid p(\lambda))\}] \geqq 0 \qquad \text{a.e. in} \quad E_n\,. \end{aligned} \tag{3.6}$$

By rearranging terms, (3.6) can be seen to hold if for all $\lambda \in \Lambda'$

$$\left\{g_n(y \mid \lambda) \frac{\partial G_n(y \mid p(\lambda))}{\partial \lambda} - \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} g_n(y \mid p(\lambda))\right\} \geqq 0 \qquad \text{a.e. in} \quad E_n \tag{3.7}$$

and

$$\left\{g_n(y\,|\,\lambda)\,\frac{\partial G_n(h_n(y)\,|\,p(\lambda))}{\partial\lambda} - h_n'(y)g_n(h_n(y)\,|\,p(\lambda))\,\frac{\partial G_n(y\,|\,\lambda)}{\partial\lambda}\right\}$$

$$\begin{aligned}(3.8)\qquad &\times \{(k-1)I(K_n(y,\lambda); k-m, m)G_n(h_n(y)\,|\,p(\lambda)) \\ &- b(K_n(y,\lambda))G_n(y\,|\,p(\lambda))\} \geqq 0 \qquad \text{a.e. in} \quad E_n\,.\end{aligned}$$

Now a simple computation shows that for any $0 \leqq a \leqq c \leqq 1$ and $1 < m < k$

$$(3.9)\qquad (k-1)cI(a/c; k-m, m) \geqq ab(a/c)$$

and so the second factor in (3.8) is nonnegative since for all $y \in E^n$ and $\lambda \in \Lambda'$ we have $0 \leqq G_n(y\,|\,p(\lambda)) \leqq G_n(h_n(y)\,|\,p(\lambda)) \leqq 1$. Hence (3.8) and (3.7) reduce to (3.4) and (3.5). Similar arguments show that (3.3) implies the regularity conditions required for Lemma 3.1 and hence completes the proof.

REMARK 3.2. The proofs of Theorem 3.2 and Lemma 3.1 also show that if (3.4) and (3.5) are identically zero, then $\psi(\lambda, n)$ 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 $\{R(n)\}$. The properties of both the sequence $\{R(n)\}$ and the individual rules $R(n)$ will be studied. It will first be shown that any $(P^*, p(\cdot))$ requirement can be met by choosing a sufficiently large common sample size n. For $\boldsymbol{\lambda} \in \Omega$ let

$$(4.1)\qquad p_{\boldsymbol{\lambda}}^{\,n}(i) = P_{\boldsymbol{\lambda}}[R(n) \text{ selects } \pi_{(i)}]\,.$$

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

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

$p_{\boldsymbol{\lambda}}^{\,n}(i)$ is $\uparrow$ in $\lambda_{[i]}$ when all other components of $\boldsymbol{\lambda}$ are fixed,
is $\downarrow$ in $\lambda_{[j]}$ $(j \neq i)$ when all other components of $\boldsymbol{\lambda}$ are fixed.

THEOREM 4.1. *If there exists $N \geqq 1$ and $\lambda_0 \in \Lambda'$ such that for all $n \geqq N$ $\inf_{\lambda\in\Lambda'} \psi(\lambda, n) = \psi(\lambda_0, n)$, then any sequence $\{R(n)\}$ defined by* (2.3) *is consistent* wrt $\Omega(p)$.

PROOF. From the hypothesis of the theorem and the result of Theorem 3.1 we have for all $n \geqq N$

$$(4.2)\qquad \inf_{\Omega(p)} P_{\boldsymbol{\lambda}}[CS\,|\,R(n)] = \int \nu(y,\lambda_0)\,dG_n(y\,|\,\lambda_0) \qquad \text{where}$$

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

Also $T_{in} \to_p \lambda_i$ as $n \to \infty \Leftrightarrow G_n(y\,|\,\lambda_i) \to \{^{0,\,y<\lambda_i}_{1,\,y>\lambda_i}$.
Since $\psi(\lambda_0, n) \leqq 1$, it suffices to show that for all $1 > \varepsilon' > 0\ \exists\, M \ni$ for all $n \geqq M$, $\int \nu(y,\lambda_0)\,dG_n(y\,|\,\lambda_0) > 1 - \varepsilon'$. Since $p(\lambda_0) < \lambda_0$, $\exists\, \alpha \ni p(\lambda_0) < \alpha < \lambda_0$. Given

$1 > \varepsilon' > 0$ let $\varepsilon = 1 - (1 - \varepsilon')^{\frac{1}{2}}$ and choose $M > N \ni$

(a) $G_n(\alpha \mid \lambda_0) < \varepsilon$ (since $\alpha < \lambda_0$)

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

So for all $y > \alpha$

(a) $$1 \geqq G_n(h_n(y) \mid p(\lambda_0)) \geqq G_n(h_n(\alpha) \mid p(\lambda_0))$$

(b) $$G_n(y \mid p(\lambda_0)) \geqq G_n(\alpha \mid p(\lambda_0))$$

which implies that for all $y > \alpha$

$$\nu(y, \lambda_0) \geqq \{G_n(h_n(\alpha) \mid p(\lambda_0))\}^{k-1} I\left(\frac{G_n(\alpha \mid p(\lambda_0))}{G_n(h_n(y) \mid p(\lambda_0))}; k - m, m\right) \geqq 1 - \varepsilon .$$

So finally for all $n \geqq M$,

$$\begin{aligned}\int \nu(y, \lambda_0)\, dG_n(y \mid \lambda_0) &\geqq \int_\alpha^\infty \nu(y, \lambda_0)\, dG_n(y \mid \lambda_0) \\ &\geqq (1 - \varepsilon) \int_\alpha^\infty dG_n(y \mid \lambda_0) \\ &\geqq 1 - \varepsilon' .\end{aligned}$$ □

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

PROOF. Since $p_\lambda^n(i) = E_\lambda[W_i(n)]$, where

$$W_i(n) = 1, \quad T_{(i)} \geqq \max\{T_{[k-m+1]}, h_n^{-1}(T_{[k]})\} \tag{4.3}$$
$$= 0, \quad \text{otherwise},$$

the result of Mahamunulu–Alam–Rizvi can again be used to show the desired monotonicity. Arguments similar to those in the proof of Theorem 2.1 show that

(A) $W_i(n)$ is non-increasing in $T_{(j)}$ $(j \neq i)$ when all other components of **T** are fixed

(B) $W_i(n)$ is non-decreasing in $T_{(i)}$ when all other components of **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.

DEFINITION 4.3. The rule R is *monotone* means for all $1 \leqq i < j \leqq k$ and $\boldsymbol{\lambda} \in \Omega$, $P_\lambda[R \text{ selects } \pi_{(j)}] \geqq P_\lambda[R \text{ selects } \pi_{(i)}]$.

DEFINITION 4.4. The rule R is *unbiased* means for all $1 \leqq i < k$ and $\boldsymbol{\lambda} \in \Omega$ $P_\lambda[R \text{ does not select } \pi_{(i)}] \geqq P_\lambda[R \text{ does not select } \pi_{(k)}]$.

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

PROOF. Since monotonicity implies unbiasedness it suffices to show that $P_\lambda^n(i) \leqq p_\lambda^n(i + 1)$ for any $i = 1, \cdots, k - 1$ and $\lambda \in \Omega$. Assuming for notational ease that $\lambda_i = \lambda_{[i]}$ it follows that

$$\begin{aligned} p_\lambda{}^n(i) &= p^n_{(\lambda_1,\cdots,\lambda_k)}(i) \\ &\leqq p^n_{(\lambda_1,\cdots,\lambda_{i-1},\lambda_{i+1},\lambda_{i+1},\lambda_{i+2},\cdots,\lambda_k)}(i) && \text{since } p_\lambda{}^n(i) \text{ is } \uparrow \text{ in } \lambda_{[i]} \\ &= p^n_{(\lambda_1,\cdots,\lambda_{i-1},\lambda_{i+1},\lambda_{i+1},\lambda_{i+2},\cdots,\lambda_k)}(i+1) \\ & && \text{since both } \pi_{(i)} \text{ and } \pi_{(i+1)} \text{ have the same cdf} \\ &\leqq p^n_{(\lambda_1,\cdots,\lambda_{i-1},\lambda_i,\lambda_{i+1},\cdots,\lambda_k)}(i+1) && \text{since } p_\lambda{}^n(i+1) \downarrow \text{ in } \lambda_{[i]} \\ &= p_\lambda{}^n(i+1)\,. && \square \end{aligned}$$

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

5. Number of selected populations. If $S(n)$ is the number of populations selected by $R(n)$, if $T(n)$ is the number of non-best populations selected by $R(n)$ and if $p_\lambda{}^n(i)$ and $W_i(n)$ are defined by (4.1) and (4.3) respectively, then the following representations hold

$$(5.1) \qquad S(n) = \textstyle\sum_{i=1}^k W_i(n)\,, \qquad T(n)\sum_{i=1}^{k-1} W_i(n)\,, \qquad p_\lambda{}^n(i) = E_\lambda[W_i(n)]\,.$$

THEOREM 5.1. *For any* $\boldsymbol{\lambda} \in \Omega$

$$E_\lambda[S(n)] = \textstyle\sum_{i=1}^k \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \int_{-\infty}^{\infty} \prod_{j\in\mathscr{S}_\nu{}^l(i)} G_n{}^{(j)}(y) \prod_{j\in\bar{\mathscr{S}}_\nu{}^l(i)} \{G_n{}^{(j)}(h_n(y)) - G_n{}^{(j)}(y)\}\, dG_n{}^{(i)}(y)\,.$$

PROOF. It follows from (5.1) that $E_\lambda[S(n)] = \sum_{i=1}^k p_\lambda[R(n) \text{ select } \pi_{(i)}]$. An argument similar to that in the proof of Theorem 2.1 serves to evaluate each term in the sum and completes the proof.

REMARK 5.1. The expected value of $T(n)$ can be derived in a similar manner.

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 $E_\lambda[S(n)]$ and $E_\lambda[T(n)]$ over Ω.

THEOREM 5.2. *For any* $\boldsymbol{\lambda} \ni \lambda_{[k]} > \lambda_{[k-1]}$, $p_\lambda{}^n(i) \to 1$ *for* $i = k$ *and* $\to 0$ *for* $i < k$ *as* $n \to \infty$.

PROOF. *Case* A: $i = k$. The strong monotonicity of $R(n)$ implies $p_\lambda{}^n(k) \geqq p^n_{\lambda'}(k)$ where $\boldsymbol{\lambda}' = (\lambda_{[k-1]}, \cdots, \lambda_{[k-1]}, \lambda_{[k]})$. Now taking $\lambda_0 = \lambda_{[k]}$ and $p(\lambda_0) = \lambda_{[k-1]}$ in the proof of Theorem 4.1 yields $p^n_{\lambda'}(k) \to 1$ as $n \to \infty$, and gives the result.

Case B: $1 \leqq i < k$. Let

$$(5.2) \qquad f_i{}^{l,\nu}(y) = \textstyle\prod_{j\in\mathscr{S}_\nu{}^l(i)} G_n{}^{(j)}(y) \prod_{j\in\bar{\mathscr{S}}_\nu{}^l(i)} \{G_n(h_n(y)) - G_n(y)\}\,;$$

then $p_\lambda{}^n(i) = \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \int f_i{}^{l,\nu}(y)\, dG_n{}^{(i)}(y)$ and it suffices to show $\int f_i{}^{l,\nu}(y)\, dG_n{}^{(i)}(y) \to 0$ as $n \to \infty$ for all l and ν.

Subcase (1): For l and ν such that $k \in \bar{\mathscr{S}}_\nu{}^l(i)$. Pick α and $\alpha' \ni \lambda_{[k-1]} < \alpha < \alpha' < \lambda_{[k]}$. Now since $h_n(\alpha) \to \alpha$ and $\alpha' < \lambda_{[k]}$, for all $\varepsilon > 0$ there exists $\exists\ N \ni$ for all

$n \geqq N$

$$G_n^{(k)}(\alpha') < \varepsilon/2\,, \qquad h_n(\alpha) < \alpha'\,, \qquad G_n^{(i)}(\alpha) > 1 - \varepsilon/2$$

which implies that for all $n \geqq N$ and $y < \alpha$

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

So finally

$$\begin{aligned} 0 &\leqq \int f_i^{l,\nu}(y)\, dG_n^{(i)}(y) = \int_{\infty}^{\alpha} f_i^{l,\nu}(y)\, dG_n^{(i)}(y) + \int_{\alpha}^{\infty} f_i^{l,\nu}(y)\, dG_n^{(i)}(y) \\ &\leqq \int_{\infty}^{\alpha} \varepsilon/2\, dG_n^{(i)}(y) + \int_{\alpha}^{\infty} 1\, dG_n^{(i)}(y) \\ &< \varepsilon \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall\, n \geqq N\,. \end{aligned}$$

Subcase (2): For l and ν such that $k \in \mathscr{S}_\nu^l(i)$. Again using a straightforward argument like the above, the desired result follows.

COROLLARY 5.1. *For any $\boldsymbol{\lambda} \in \Omega$ $\lambda_{[k]} > \lambda_{[k-1]}$, $W_i(n) \to_p 1$ for $i = k$ and $\to_p 0$ for $i < k$ as $n \to \infty$.*

PROOF. For any $\varepsilon > 0$, $P_\lambda[|W_k(n) - 1| > \varepsilon] = P_\lambda[W_k(n) = 0] = 1 - p_\lambda{}^n(k) \to 0$ as $n \to \infty$ and for $i < k$, $P_\lambda[|W_i(n)| > \varepsilon] = P_\lambda{}^n(i) \to 0$ as $n \to \infty$.

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) and $(S(n) - 1) \leqq (S(n) - 1)^2$ together with the convergence in probability of the $W_i(n)$ random variables we obtain

COROLLARY 5.2. *For $\boldsymbol{\lambda} \in \Omega$ such that $\lambda_{[k]} > \lambda_{[k-1]}$*

(1) *$S(n) \to_{p_1}$ and $T(n) \to_p 0$ as $n \to \infty$ and hence*

(2) *$E_\lambda[S(n)] \to 1$ and $E_\lambda[T(n)] \to 0$ 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 $E_\lambda[S(n)]$ in Ω occurs at some point $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_k)$ for which $\lambda_{[1]} = \lambda_{[k]}$. The condition (5.3) will be assumed in some of the theorems which follow.

(5.3) (i) $E_n{}^\lambda = E_n$ for all $\lambda \in \Lambda$.

(ii) For any $[\lambda_1, \lambda_2] \subset \Lambda$ there exists $e_3(y)$ depending only on λ_1 and $\lambda_2 \ni \left|\dfrac{\partial G_n(y \mid \lambda)}{\partial \lambda}\right| \leqq e_3(y)$ where

$$\left(\int e_3(y)\, dG_n(h_n(y) \mid \lambda')\right)\left(\int e_3(h_n(y))\, dG_n(y \mid \lambda')\right) < \infty \quad \text{for all } \lambda' \geqq \lambda_2\,.$$

THEOREM 5.3. *If (5.3) is satisfied and for all λ_1, λ_2 in Λ with $\lambda_1 \leqq \lambda_2$*

$$\frac{\partial G_n(h_n(y) \mid \lambda_1)}{\partial \lambda_1}\, g_n(y \mid \lambda_2) - \frac{\partial G_n(y \mid \lambda_1)}{\partial \lambda_1}\, g_n(h_n(y) \mid \lambda_2) h_n{}'(y) \geqq 0 \qquad \text{a.e. in } E_n \tag{5.4}$$

then $E_\lambda[S(n)]$ is non-decreasing in $\lambda_{[1]}$ on $\Lambda(\lambda_{[2]}) = \{\lambda \in \Lambda \mid \lambda \leqq \lambda_{[2]}\}$ for any fixed $(\lambda_{[2]}, \cdots, \lambda_{[k]})$.

PROOF. Fix $\lambda_{[2]} \leqq \cdots \leqq \lambda_{[k]}$ for the following argument and then $E_{\lambda}[S(n)] = T_1(\boldsymbol{\lambda}) + T_2(\boldsymbol{\lambda})$ where

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

$$T_2(\boldsymbol{\lambda}) = \sum_{i=2}^{k} \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \int_{E_n} f_i^{l,\nu}(y)\, dG_n^{(i)}(y)$$

where $f_i^{l,\nu}(y)$ is defined by (5.2).
Now $T_2(\boldsymbol{\lambda})$ can be rewritten as

$$T_2(\boldsymbol{\lambda}) = \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \sum_{i \ni 1 \in \mathscr{S}_\nu{}^l(i)} \int_{E_n} f_i^{l,\nu}(y)\, dG_n^{(i)}(y)$$
$$+ \sum_{l=k-m}^{k-1} \sum_{\nu=1}^{\binom{k-1}{l}} \sum_{i \ni 1 \in \mathscr{S}_\nu{}^l} \int_{E_n} f_i^{l,\nu}(y)\, dG_n^{(i)}(y) .$$

For any $A \subset \{1, \cdots, k\}$ of size s, let $\{\mathscr{S}_\nu{}^l(A) \mid \nu = 1, \cdots, \binom{k-s}{l}\}$ be the collection of all subsets of size l from $\{1, \cdots, k\} - A$. Note that for any fixed $l = k - m, \cdots, k - 1$ and $i = 2, \cdots, k$, $\{\mathscr{S}_\nu{}^l(i) \mid 1 \in \mathscr{S}_\nu(i)\} = \{\mathscr{S}_\nu{}^{l-1}(1, i) \cup \{1\} \mid \nu = 1, \cdots, \binom{k-2}{i-1}\}$, while for any $l = k - m, \cdots, k - 2$ and $i = 2, \cdots, k$, $\{\mathscr{S}_\nu{}^l(i) \mid 1 \notin \mathscr{S}(i)\} = \{\mathscr{S}_\nu{}^l(1, i) \mid \nu = 1, \cdots, \binom{k-2}{l}\}$.
So

$$T_2(\boldsymbol{\lambda}) = \sum_{l=k-m}^{k-1} \sum_{i=2}^{k} \sum_{\nu=1}^{\binom{k-2}{l-1}} \int_{E_n} W_i^{l,\nu}(y) G_n^{(1)}(y)\, dG_n^{(i)}(y)$$
$$+ \sum_{l=k-m}^{k-2} \sum_{i=2}^{k} \sum_{\nu=1}^{\binom{k-2}{l}} \int_{E_n} Z_i^{l,\nu}(y) \{G_n^{(1)}(h_n(y)) - G_n^{(1)}(y)\}\, dG_n^{(i)}(y)$$

where

(1) $W_i^{l,\nu}(y) = \prod_{j \in \mathscr{S}_\nu{}^{l-1}(1,i)} G_n^{(j)}(y) \prod_{j \in \bar{\mathscr{S}}_\nu{}^{l-1}(1,i)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\}$

(2) $Z_i^{l,\nu}(y) = \prod_{j \in \mathscr{S}_\nu{}^{l}(1,i)} G_n^{(j)}(y) \prod_{j \in \bar{\mathscr{S}}_\nu{}^{l}(1,i)} \{G_n^{(j)}(h_n(y)) - G_n^{(j)}(y)\}$.

Next integrating $T_1(\boldsymbol{\lambda})$ by parts and noting that for fixed $l = k - m, \cdots, k - 1$ and $i = 2, \cdots, k$, $\{\mathscr{S}_\nu{}^l(1) \mid i \in \mathscr{S}_\nu{}^l(1)\} = \{\mathscr{S}_\nu{}^{l-1}(1, i) \cup \{i\} \mid \nu = 1, \cdots, \binom{k-2}{l-1}\}$, while for any $l = k - m, \cdots, k - 2$ and $i = 2, \cdots, k$, $\{\mathscr{S}_\nu{}^l(1) \mid i \in \mathscr{S}_\nu{}^l(1)\} = \{\mathscr{S}_\nu{}^l(1, i) \mid \nu = 1, \cdots, \binom{k-2}{l}\}$, we obtain that

$$T_1(\boldsymbol{\lambda}) = 1 - \sum_{l=k-m}^{k-1} \sum_{i=2}^{k} \sum_{\nu=1}^{\binom{k-2}{l-1}} \int_{E_n} W_i^{l,\nu}(y) G_n^{(1)}(y)\, dG_n^{(i)}(y)$$
$$- \sum_{l=k-m}^{k-1} \sum_{i=2}^{k} \sum_{\nu=1}^{\binom{k-2}{l}} \int_{E_n} Z_i^{l,\nu}(y) G_n^{(1)}(y) \{g_n^{(i)}(h_n(y)) h_n'(y)$$
$$- g_n^{(i)}(y)\}\, dy .$$

Hence combining and cancelling terms it follows that

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

and finally

$$\frac{dE_\lambda[S(n)]}{d\lambda_{[1]}} = \sum_{l=k-m}^{k-2} \sum_{i=2}^{k} \sum_{\nu=1}^{\binom{k-2}{l}} \int_{E_n} Z_i^{l,\nu}(y) \tag{5.5}$$
$$\times \left\{ \frac{\partial G_n^{(1)}(h_n(y))}{\partial \lambda_{[1]}} g_n^{(i)}(y) - \frac{G_n^{(1)}(y)}{\partial \lambda_{[1]}} g_n^{(i)}(h_n(y)) h_n'(y) \right\} dy .$$

But (5.4) gives for every $i = 2, \cdots, k$

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

which implies the derivative in (5.5) is nonnegative and completes the proof.

REMARK 5.3. Condition (5.4) 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* $\lambda_{[2]} \leqq \cdots \leqq \lambda_{[k]}$, $\{dE_\lambda[S(n)]/d\lambda_{[1]}\} \geqq 0$ *for* $\lambda_{[1]}$ *in* $\Lambda(\lambda_{[2]})$, *then the* $\sup_\Omega E_\lambda[S(n)] = \sup_{\lambda \in \Lambda} \gamma(\lambda, n)$ *where*

$$(5.6) \qquad \gamma(\lambda, n) = k \int_{E_n^\lambda} \{G_n(h_n(y) \mid \lambda)\}^{k-1} I\left(\frac{G_n(y \mid \lambda)}{G_n(h_n(y) \mid \lambda)}; k-m, m\right) dG_n(y \mid \lambda) .$$

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

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

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

But from the previous proof

$$\frac{\partial E_{\boldsymbol{\lambda}'}[S(n)]}{\partial \lambda_{[1]}}\bigg|_{\boldsymbol{\lambda}(q)} \geqq 0 .$$

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

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

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

$$\begin{aligned} &\{G_n(h_n(y) \mid \lambda)\}^{k-3} \left\{g_n(y \mid \lambda) \frac{\partial G_n(h_n(y) \mid \lambda)}{\partial \lambda} - \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} g_n(h_n(y) \mid \lambda) h_n'(y)\right\} \\ &\quad \times \left\{(k-1G)_n(h_n(y) \mid \lambda) I\left(\frac{G_n(y \mid \lambda)}{G_n(h_n(y) \mid \lambda)}; k-m, m\right)\right. \\ &\quad \left. - G_n(y \mid \lambda) b\left(\frac{G_n(y \mid \lambda)}{G_n(h_n(y) \mid \lambda)}\right)\right\} \geqq 0 \qquad \forall \lambda \text{ and a.e. } y \\ &\Leftrightarrow \left\{g_n(y \mid \lambda) \frac{\partial G_n(h_n(y) \mid \lambda)}{\partial \lambda} - \frac{\partial G_n(y \mid \lambda)}{\partial \lambda} g_n(h_n(y) \mid \lambda) h_n'(y)\right\} \geqq 0 \\ &\qquad \forall \lambda \text{ and a.e. } y \end{aligned}$$

since the third factor is nonnegative by (3.9). 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

$$E_\lambda[T(n)] = E_\lambda[S(n)] - p_\lambda^n(k) . \tag{5.7}$$

COROLLARY 5.4. *If the hypotheses of Corollary* 5.3 *hold then* $\sup_\Omega E_\lambda[T(n)] = [(k-1)/k] \sup_{\lambda \in \Lambda} \gamma(\lambda, n)$ *where* $\gamma(\lambda, n)$ *is defined by* (5.6).

PROOF. For any $\boldsymbol{\lambda} = (\lambda_{[1]}, \cdots, \lambda_{[k]}) \in \Omega$ let $\boldsymbol{\lambda}([k]) = (\lambda_{[k]}, \cdots, \lambda_{[k]})$; then for all $\boldsymbol{\lambda} \in \Omega$ the hypotheses imply $E_\lambda[S(n)] \leqq E_{\lambda([k])}[S(n)]$. Also the strong monotonicity of $R(n)$ implies $p_\lambda^n(k) \geqq p^n_{\lambda([k])}(k)$. So by (5.7)

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

implies $\sup_\Omega E_\lambda[T(n)] = (k-1)/k \sup_{\lambda \in \Lambda} \gamma(\lambda, n)$.

REMARK 5.5. From Corollary 3.3 it follows that $\gamma(\lambda, n)$ is non-decreasing in λ if the hypotheses of Theorem 5.3 hold 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 $\pi_i \sim N(\mu_i, \sigma^2)$, $i = 1, \cdots, k$ where the common variance σ^2 is known and the experimenter is interested in selecting the population having largest μ_i. We take $T_{in} = (1/n) \sum_{j=1}^n X_{ij}$ and then $\lambda_i = \mu_i$ and $G_n(y \mid \lambda_i) = \Phi\{n^{\frac{1}{2}}(y - \mu_i)/\sigma\}$ where Φ is the cdf of an $N(0, 1)$ random variable.

Since this is a location parameter problem we take $p(\mu) = \mu - \delta (\delta > 0)$ $h_n(x) = x + d\sigma/n^{\frac{1}{2}}$ and obtain $\Omega(p) = \{\mu \mid \mu_{[k]} - \mu_{[k-1]} \geqq \delta\}$ and $R(n)$: Select $\pi_i \Leftrightarrow \bar{X}_i \geqq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/n^{\frac{1}{2}}\}$. Using Theorem 3.1 and Corollary 5.3 it can be seen that

$$\inf_{\Omega(p)} P[CS \mid R(n)] = \int_{-\infty}^{\infty} \left\{\Phi\left(y + d + \frac{(n\sigma)^{\frac{1}{2}}}{\sigma}\right)\right\}^{k-1} \times I\left(\frac{\Phi(y+d)}{\Phi(y + d + (n\sigma)^{\frac{1}{2}}/\sigma)} ; k-m, m\right) d\Phi(y) \tag{6.1}$$

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

One choice of $\{R(n)\}$ can be made by setting the right-hand side of (6.2) equal to $1 + \varepsilon$ 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 of Desu and Sobel (1968) and tables of constants required to implement the proposed procedure are given in Gupta and Santner (1973).

II. Now suppose π_i is p-variate normal with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}(N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}))$ for $i = 0, 1, \cdots, k$. The common $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}_0$ are both known and π_0 may be thought 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 $\lambda_i = (\boldsymbol{\mu}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_0)$. Gupta (1966), Alam and Rizvi (1966) and Gupta and Studden (1970) have considered this problem. We take $T_{in} = (X_{ij} - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(X_{ij} - \boldsymbol{\mu}_0)$, $p(\lambda) = p_3(\lambda)$ of Example 2.1, $h_n(x) = d^{1/n}x (d > 1)$ and

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

so that $T_{in} \to_p p + \lambda_i$ as $n \to \infty$ and $G_n(y \mid \lambda_i) = F_{np}(y \mid \lambda_i)$, $\Omega(p) = \Omega_1 \cap \Omega_2$ where $\Omega_1 = \{\boldsymbol{\lambda} \mid \lambda_{[k]} - \lambda_{[k-1]} \geqq \delta_1\}$ and $\Omega_2 = \{\boldsymbol{\lambda} \mid \lambda_{[k]} \geqq \delta_2 \lambda_{[k-1]}\}$. Also $R(n)$: Select $\pi_i \Leftrightarrow T_{in} \geqq \max\{T_{[k-m+1]n}, d^{-1/n}T_{[k]n}\}$.

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

$$\begin{aligned} \frac{\partial F_p(y \mid \lambda)}{\partial \lambda} &= \tfrac{1}{2}[F_{p+2}(y \mid \lambda) - F_p(y \mid \lambda)] \\ &= -f_{p+2}(y \mid \lambda) \qquad \text{where} \quad f_p(y \mid \lambda) = \frac{dF_p(y \mid \lambda)}{dy}\,, \end{aligned}$$

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

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

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

$$\begin{aligned} \psi(\lambda, n) &= \int_0^\infty \{F_{np}(yd^{1/n} \mid (\lambda - \delta_1))\}^{k-1} \\ &\quad \times I\left(\frac{F_{np}(y \mid (\lambda - \delta_1))}{F_{np}(y \mid (\lambda - \delta_1))}\,; k - m, m\right) dF_{np}(y \mid \lambda)\,, \qquad \lambda \in I_1 \\ &= \int_0^\infty \{F_{np}(yd^{1/n} \mid \delta_2^{-1}\lambda)\}^{k-1} \\ &\quad \times I\left(\frac{F_{np}(y \mid \delta_2^{-1}\lambda)}{F_{np}(yd^{1/n} \mid \delta_2^{-1}\lambda)}\,; k - m, m\right) dF_{np}(y \mid \lambda)\,, \qquad \lambda \in I_2 \end{aligned} \tag{6.3}$$

where $I_1 = [\delta_1, \delta_1\delta_2/(\delta_2 - 1))$ and $I_2 = [\delta_1\delta_2/(\delta_2 - 1), \infty)$.

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 $f_p(y \mid \lambda)$ density listed above, a piecewise application of Theorem 3.2 on I_1 and I_2 shows that $\psi(\lambda, n)$ is $\downarrow$ in λ on I_1 and $\uparrow$ in λ on I_2 and hence $\inf_{\Omega(p)} P[CS \mid R(n)] = \psi(\delta_1\delta_2/(\delta_2 - 1), n)$. Theorem 4.1 applies since $\delta_1\delta_2/(\delta_2 - 1) \in \Lambda = [0, \infty)$, and hence $\inf_{\Omega(p)} P_\lambda[CS \mid R(n)] \to 1$ as $n \to \infty$. All other usual properties hold for $R(n)$ and in particular (5.4) holds (as verified by Panchapakesan (1969)) and hence $\sup_\Omega E[S(n)] = \sup_{\lambda \geqq 0} \gamma(\lambda, n)$ and $\gamma(\lambda, n) \uparrow$ in

λ where

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

Using a probability argument this supremum can be evaluated as $\lim_{\lambda\to\infty} \gamma(\lambda, n) = m$. We obtain $\sup_\Omega E[S(n)] = m$. Details of the above problem as well as other applications to regular and nonregular problems can be found in Santner (1973).

Acknowledgment. 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. The author wishes also to thank the referee and Professor W. J. Hall whose constructive suggestions led to significant improvements in the presentation of the 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, **3**. Univ. of Chicago Press.

[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, Univ. of North Carolina, Chapel Hill.

[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* (P. R. Krishaiah, ed.) 457–475. Academic Press, New York.

[8] Gupta, S. S. and Nagel, K. (1971). On some contributions to multiple decision theory. In *Statistical Decision Theory and Related Topics* (S. S. Gupta and J. Yackel, eds.). 79–102. Academic Press, New York.

[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. (1973). On selection and ranking procedures—a restricted subset selection rule. In *Proceedings of the 39th Session of the International Statistical Institute*, Vienna, Austria. **1** 409–417.

[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* (R. C. Bose *et al.* eds.). 327–338. Univ. of North Carolina Press, Chapel Hill.

[12] Huang, W. T. (1972). Some contributions to sequential selection and ranking procedures. Mimeograph Series N. 295, Department of Statistics, Purdue Univ.

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

[14] Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *Ann. Math. Statist.* **38** 1079–1091.

[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 Univ.

[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* (S. S. Gupta and J. Yackel eds.). 275–298. Academic Press, New York.

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

[19] SOBEL, M. (1969). Selecting a subset containing at least one of the t best populations. In *Multivariate Analysis II* (P. R. Krishnaiah, ed.). 515–540. Academic Press, New York.

DEPARTMENT OF OPERATIONS RESEARCH
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850