ON THE PROBABILITY OF CORRECT SELECTION IN THE SUBSET SELECTION PROBLEM

by

Jan F. Bjørnstad
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-32

November 1979

ON THE PROBABILITY OF CORRECT SELECTION IN THE SUBSET SELECTION PROBLEM

by

Jan F. Bjørnstad
Purdue University

1. Introduction.

Subset selection deals with the problem of selecting a random non-empty subset of populations out of say, k populations, with the aim that the selected populations are "close" in some sense to the best population. In particular, a subset including the best population is called a correct selection (CS). The classical condition on subset selection procedures is to require $P(CS) \geq P^*$. Usually P^* is chosen to be greater than $1/k$.

There seems to be some confusion as to why and whether $1/k$ is an appropriate lower bound for P^*. Gupta (1965) states that one should choose $P^* \geq 1/k$ because for $P^* < 1/k$ there always exist a no-data decision rule. Gibbons, Olkin and Sobel (1977) states that $1/2^k$ is the appropriate lower bound, but the justification given for this bound is incorrect. The bound $1/2^k$ is claimed to be obtained by the rule that selects each population with probability 1/2. However, this is not a subset selection rule since it may select an empty set. Also, $P(CS)=1/2$ for this procedure. Bechhofer and Santner (1979) support the lower bound $1/k$ on the basis of certain minimax arguments for no-data decision-rules.

The aim of this note is to clarify this issue. This author thinks that a lower bound on P^* should depend only on the decision-space and the class of procedures under consideration. From this point of view it is shown in Section 1 that for some reasonable classes, 1/k is the correct lower bound, in the sense that no procedure in those classes can achieve a P* less than 1/k. However, it turns out that for several classes 1/k is not the appropriate bound. It is also shown that a procedure with P(CS)<1/k has certain undesirable properties, which gives an argument in favor of considering only classes of procedures that has 1/k as the lower bound.

Section 2 deals with no-data rules. It is shown that Gupta's statement is incorrect if one only considers permutation-invariant no-data procedures. Bechhofer and Santner's approach is also briefly discussed.

2. P(CS) for Monotone, Ordered and Permutation-Invariant Procedures.

We shall consider the following situation. The k populations are denoted by $\pi_1,\ldots,\pi_k$. π_i is characterized by a real-valued parameter θ_i. X_i is the observation from π_i. $X_1,\ldots,X_k$ are assumed to be independent. $F_{\theta_i}(x)$ is the distribution function of X_i. It is assumed that $F_\theta(x)$ is stochastically increasing in θ, and continuous in x for each θ in the parameter-space $\Theta \subset \mathbb{R}$. This class of distribution functions is denoted by $\mathcal{F}_c$. Let $\Omega=\Theta^k$. Let $\theta_{[1]}\le\cdots\le\theta_{[k]}, X_{[1]}\le\cdots\le X_{[k]}$ denote the ordered θ_i's and X_i's. $X_{(i)}, \pi_{(i)}$ correspond to $\theta_{[i]}$. $\pi_{(k)}$ is defined to be the best population. The decision-space $\mathscr{A}$ is the set of all non-empty subsets of $(\pi_1,\ldots,\pi_k)$. A subset selection-rule δ is for each observed $\underline{x}=(x_1,\ldots x_k)$ a probability-measure $\delta(a|\underline{x})$ over $a\in\mathscr{A}$. For a procedure δ, the individual selection probabilities are given by:

$$\psi_i^\delta(\underline{x}) = P(\text{selecting } \pi_i \text{ with rule } \delta|\underline{x}) = \sum_{a \ni i} \delta(a|\underline{x})$$

Let $\psi^{\delta}_{(i)}$ correspond to $\theta_{[i]}$. Then the classical P* condition is

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|\delta) = \inf_{\underline{\theta}\in\Omega} E_{\underline{\theta}}\psi^{\delta}_{(k)}(\underline{X}) = P^* \qquad (1)$$

Let us for convenience denote $\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|\delta)$ by $P^*(\delta)$. The range of possible values of P* will depend upon the class $\mathcal{L}$ under consideration. Suppose $\inf_{\delta\in\mathcal{L}} P^*(\delta) = \alpha$, then the principle is that P* should be at least α, because no rules in $\mathcal{L}$ can achieve a P* less than α. Consider for example the class $\mathcal{L}_G$ of Gupta's rules (see Gupta (1965)).

$$\psi_i = 1 \text{ iff } X_i \geq X_{[k]} - d \quad , \quad d \geq 0.$$

Here $\inf_{\delta\in\mathcal{L}_G} P^*(\delta)$ is 1/k, achieved by the rule corresponding to d=0. Hence for $\mathcal{L}_G$, P* should be at least 1/k.

This principle will be applied to different classes of procedures to find out if 1/k is the natural lower bound. In order to define the class $\mathcal{L}_I$ of permutation-invariant procedures, let g be a permutation of (1,...,k) such that gi is the new position of element i under permutation g. Then $g\underline{x}$ is defined by $(g\underline{x})_i = x_{g^{-1}i}$, and $ga = \{gi: i \in a\}$ for $a \in \mathcal{A}$.

Definition 1. $\delta \in \mathcal{L}_I$ if for each permutation g

$$\delta(ga|g\underline{x}) = \delta(a|\underline{x}) \qquad \forall a \in \mathcal{A}, \ \forall \underline{x}.$$

Definition 2. δ is said to be ordered if

$$\psi^{\delta}_i(\underline{x}) \leq \psi^{\delta}_j(\underline{x}) \qquad \text{when } x_i < x_j$$

$\mathcal{L}_0$ denotes the class of ordered procedures.

Definition 3. δ is called monotone if for each i,

$$\psi_i^\delta(\underline{x}) \leq \psi_i^\delta(\underline{y}) \quad \text{if} \quad x_i \leq y_i \text{ and } x_j \geq y_j \quad \forall j \neq i .$$

Let $\mathcal{L}_M$ be the class of monotone procedures.

$\mathcal{L}_I$, $\mathcal{L}_O$, $\mathcal{L}_M$ are the three basic classes of procedures we consider. If $\mathcal{L}$ is one of the three or a combination of these, the basic question to answer is whether or not the following statement is true:

$$\inf_{\delta \in \mathcal{L}} P^*(\delta) = 1/k \tag{2}$$

Let us first discuss the relationship between the three classes. Clearly a rule can be monotone and not ordered or vice versa. The following results also hold.

Lemma 1.

(i) $\delta \in \mathcal{L}_I \not\Rightarrow \delta \in \mathcal{L}_O$ (i.e. $\mathcal{L}_I - \mathcal{L}_O$ is non-empty.)

(ii) $\delta \in \mathcal{L}_O \not\Rightarrow \delta \in \mathcal{L}_I$ (i.e. $\mathcal{L}_O - \mathcal{L}_I$ is non-empty.)

(iii) $\mathcal{L}_{I,M} \subset \mathcal{L}_{I,O}$ and $\mathcal{L}_{I,O} - \mathcal{L}_{I,M}$ is non-empty.

Here $\mathcal{L}_{I,M} = \mathcal{L}_I \cap \mathcal{L}_M$ and $\mathcal{L}_{I,O} = \mathcal{L}_I \cap \mathcal{L}_O$.

Proof.

(i) is obvious. E.g. the rule that selects π_i if and only if $X_i = X_{[k-1]}$ is permutation-invariant but not ordered.

(ii). Consider the following rule δ:

If $X_1 = X_{[k]}$: select π_1

If $X_1 < X_{[k]}$: select $\pi_{[k]}$, $\pi_{[k-1]}$, where $\pi_{[i]}$ corresponds to $X_{[i]}$.

$\delta \in \mathcal{L}_O$, but $\delta \notin \mathcal{L}_I$.

(iii). First we note that $\delta\in\mathcal{L}_I$ implies that $\psi_i^\delta(\underline{x})=\psi_{gi}^\delta(g\underline{x})$ $\forall\,(g,i,\underline{x})$. Assume $\delta\in\mathcal{D}_{I,M}$. Let $\underline{x}$ be such that $x_i < x_j$.
We shall show that $\psi_i^\delta(\underline{x}) \le \psi_j^\delta(\underline{x})$.
Let g be the permutation with $gi=j$, $gj=i$, $g\ell=\ell$ $\forall\,\ell\ne i,j$, and let $y=gx$. Then $y_i=x_j>x_i$ and $y_\ell\le x_\ell$ $\forall\,\ell\ne i$. Hence $\psi_i^\delta(\underline{x})\le\psi_i^\delta(\underline{y})$ from Definition 3, and $\psi_j^\delta(\underline{x})=\psi_{gj}^\delta(g\underline{x})=\psi_i^\delta(\underline{y})\ge\psi_i^\delta(\underline{x})$, which proves the first statement. Let now $k\ge 3$, and consider the following rule.

$$\delta:\ \text{select } \pi_i \Leftrightarrow X_i \ge \min\left(X_{[k]},\frac{X_{[k]}}{\overline{X}}\right)\ ;\ \overline{X}=\frac{1}{k}\sum_{i=1}^{k}X_i \quad .$$

δ is clearly in $\mathcal{L}_{I,0}$. We shall show that δ is not monotone.
Let $x_i=3/2$ for $i\le k-1$ and $x_k=2$, and let $y_i=0$ for $i\le k-2$, $y_{k-1}=3/2$, $y_k=2$. Here $x_{k-1}=y_{k-1}$ and $y_j\le x_j$ $\forall\,j\ne k-1$. It is readily seen that $\psi_{k-1}(\underline{x})=1$ and $\psi_{k-1}(\underline{y})=0$. Q.E.D.

The results about $P^*(\delta)$ for the classes $\mathcal{L}_I$, $\mathcal{L}_0$, $\mathcal{L}_M$ are given in the following

Theorem 1.

(a) $\inf_{\delta\in\mathcal{D}_I} P^*(\delta) = 0$, provided $F_\theta(x)\to 1$ as $\theta\to\inf\Theta$.

(b) $\inf_{\delta\in\mathcal{L}_0} P^*(\delta) = (1/k)^2$; $\forall\,F_\theta\in\mathfrak{F}_c$.

(c) $\inf_{\delta\in\mathcal{L}_M} P^*(\delta) = 0$; $\forall\,F_\theta\in\mathfrak{F}_c$.

Proof.

(a) Consider the rule

$$\delta:\ \text{select } \pi_i \Leftrightarrow X_i = X_{[1]}\ .$$

$$\delta\in\mathcal{L}_1 \text{ so we may assume } \theta_i=\theta_{[i]}\quad .$$

Then

$$P_{\underline{\theta}}(CS|\delta) = P(X_k \le X_i \; ; \; \forall\, i \le k-1) = \int \prod_{i=1}^{k-1} P(X_i \ge x)\, dF_{\theta_k}(x)$$

$$= \int \prod_{j=1}^{k-1} (1-F_{\theta_j}(x))\, dF_{\theta_k}(x) \; .$$

Let $\theta_j \to \inf \Theta \;\; \forall\, j \le k-1$ and keep θ_k fixed. Then $F_{\theta_j}(x) \to 1, \forall\, x$ and from Lebesgues convergence theorem $P_{\underline{\theta}}(CS|\delta) \to 0$. Hence $P^*(\delta) = 0$.

(b) We observe that $a \ne \phi$ implies $\sum_{a \in \mathscr{A}} \delta(a|\underline{x}) = 1 \;\; \forall\, \underline{x}$ and hence

$$\sum_{i=1}^{k} \psi_i^{\delta}(\underline{x}) \ge \sum_{a \in \mathscr{A}} \delta(a|\underline{x}) = 1. \tag{3}$$

Let $\delta \in \mathscr{D}_0$. Then $x_i = x_{[k]} \Rightarrow \psi_i(\underline{x}) \ge 1/k$ from (3)

Hence:

$$P_{\underline{\theta}}(CS|\delta) = E_{\underline{\theta}}(\psi_{(k)}|\delta) \ge \int_{\{\underline{x}:\; x_{(k)}=x_{[k]}\}} \psi_{(k)}(\underline{x}) \prod_{i=1}^{k} dF_{\theta_i}(x_i) \ge \frac{1}{k} P_{\underline{\theta}}(X_{(k)} = X_{[k]})$$

Now $\inf_{\underline{\theta} \in \Omega} P_{\underline{\theta}}(X_{(k)} = X_{[k]}) = P_{\theta_1 = \cdot = \theta_k}(X_{(k)} = X_{[k]}) = 1/k$,

since F_{θ} is stochastically increasing in .

Hence: $P_{\underline{\theta}}(CS|\delta) \ge (1/k)^2 \quad \forall\, \underline{\theta} \in \Omega$.

The lower bound is achieved by the following rule in $\mathscr{D}_0$:

If $x_i = x_{[k]}$ and $i<k$: $\delta(a_i|\underline{x}) = 1$

If $x_k = x_{[k]}$: $\delta(a_i|\underline{x}) = 1/k$ for $i=1,\ldots k$.

Here $a_i = \{\pi_i\}$.

(c) is obvious e.g. let δ be given by:

$$\psi_i^{\delta}(\underline{x}) = 1 \Leftrightarrow x_i \ge x_{[k-1]} \text{ for } i \le k-1$$

$$\psi_k^{\delta}(\underline{x}) = 0 \quad \forall\, \underline{x}$$

$\delta \in \mathcal{D}_M$ and if $\theta_k < \max_{1 \le i \le k-1} \theta_i$ then $P_{\underline{\theta}}(CS|\delta) = 0$. Q.E.D.

From Theorem 1, we see that none of the three properties, permutation-invariant, monotone or ordered, alone insures (2).

A desirable property of a procedure δ is unbiasedness.

Definition 4. δ is said to be unbiased if

$$i<j \Rightarrow E_{\underline{\theta}}\psi^{\delta}_{(i)} \le E_{\underline{\theta}}\psi^{\delta}_{(j)}$$

(Some authors, e.g. Gupta (1965) and Nagel (1970) use the terminology "monotone" for this property.)

Let S be the size of the selected subset. Then

$$E_{\underline{\theta}}(S|\delta) = \sum_{i=1}^{k} E_{\underline{\theta}}\psi^{\delta}_{(i)}$$

Lemma 2.

$$P^*(\delta) < 1/k \Rightarrow \delta \text{ is not unbiased.}$$

Proof. Let $\underline{\theta} \in \Omega$, arbitrary, and assume that δ is unbiased. Then

$$k \cdot P_{\underline{\theta}}(CS|\delta) = kE_{\underline{\theta}}\psi^{\delta}_{(k)} \ge \sum_{i=1}^{k} E_{\underline{\theta}}\psi^{\delta}_{(i)} = E_{\underline{\theta}}(S|\delta) \ge 1 \text{ , since } S \ge 1.$$

Hence $P^*(\delta) \ge 1/k$.

Q.E.D.

So, if one only wants to consider unbiased procedures at least (2) must be satisfied. From Theorem 1 we see that there are biased procedures in each of the classes $\mathcal{D}_I$, $\mathcal{D}_O$, $\mathcal{D}_M$. It turns out (see Theorem 2 below) that stronger results can be obtained for some combinations of the three classes. Also restricting attention to non-randomized procedures can lead to different results. δ is non-randomized if for each x there exists $a \in \mathscr{A}$, such that $\delta(a|\underline{x}) = 1$. For a given class $\mathcal{D}$, let $\mathcal{D}^n$ denote the class of non-randomized procedures in D. E.g. $\mathcal{D}_O^n$ is the class of non-randomized ordered procedures. Our basic question is now answered by the following results.

Theorem 2.

(a) $\inf_{\delta \in \mathcal{D}_{I,M}} P^*(\delta) = 1/k \; , \quad \forall\, F_\theta \in \mathcal{F}_C .$

(b) $\inf_{\delta \in \mathcal{D}_0^n} P^*(\delta) = \inf_{\delta \in \mathcal{D}_{I,0}^n} P^*(\delta) = \inf_{\delta \in \mathcal{D}_{M,0}^n} P^*(\delta) = 1/k ; \quad \forall\, F_\theta \in \mathcal{F}_C .$

Here $\mathcal{D}_{M,0} = \mathcal{D}_M \cap \mathcal{D}_0$.

(c) $\inf_{\delta \in \mathcal{D}_{I,0}} P^*(\delta) < 1/k$ for some $F_\theta \in \mathcal{F}_C$

(d) $\inf_{\delta \in \mathcal{D}_{M,0}} P^*(\delta) \leq 1/2k \; , \quad \forall\, F_\theta \in \mathcal{F}_C .$

Proof.

(a) Let $\delta \in \mathcal{D}_{I,M}$. Then from Nagel (1970):

$$P^*(\delta) = \inf_{\underline{\theta} \in \Omega_0} P_{\underline{\theta}}(CS|\delta) \; , \text{ where } \Omega_0 = \{\underline{\theta} : \theta_1 = \ldots = \theta_k\}.$$

For $\underline{\theta} \in \Omega_0$: $E_{\underline{\theta}}(\psi_1^\delta) = \ldots = E_{\underline{\theta}}(\psi_k^\delta) = P_{\underline{\theta}}(CS|\delta)$ since $\delta \in \mathcal{D}_I$.
Hence: $E_{\underline{\theta}}(S|\delta) = kP_{\underline{\theta}}(CS|\delta)$. Since $S \geq 1$, it follows that $\inf_{\mathcal{D}_{I,M}} P^*(\delta) \geq 1/k$.

The lower bound is obtained by the rule that selects π_i if and only if $X_i = X_{[k]}$.

(b) Let now $\delta \in \mathcal{D}_0^n$. Then $X_{(k)} = X_{[k]} \Rightarrow X_{(k)} > X_{(i)} \quad \forall\, i \leq k-1$ with F_θ-probability 1, which implies $\psi_{(k)}(\underline{X}) = 1$ with probability 1. Hence

$$P_{\underline{\theta}}(CS|\delta) = E_{\underline{\theta}}\psi_{(k)}^\delta \geq P_{\underline{\theta}}(X_{(k)} = X_{[k]}) \geq 1/k \; .$$

The lower bound is attained by the same rule as in (a).

(c) Let $F_\theta \in \mathcal{F}_c$ be such that there exists $\theta' < \theta''$, $a<b$ for which

$$F_{\theta'}(a) = .49 \qquad 1-F_{\theta'}(b) = .5$$

$$F_{\theta''}(a) = .01 \qquad 1-F_{\theta''}(b) = .5$$

Let $\theta_1 = \ldots = \theta_{k-1} = \theta'$, $\theta_k = \theta''$.

Consider the following rule.

$$\delta(a_i|\underline{x}) = 1 \quad \text{iff} \quad x_i > b \text{ and } a < x_j \leq b \quad \forall\, j \neq i$$

$$\text{for } i=1,\ldots k$$

$$\delta(a_1|\underline{x}) = \ldots = \delta(a_k|\underline{x}) = 1/k, \text{ otherwise.}$$

It is readily seen that $\delta \in \mathcal{D}_{I,0}$.

$$P_{\underline{\theta}}(CS|\delta) = E_{\underline{\theta}}\psi_k^\delta = P(a < X_i \leq b, \forall i \leq k-1 \text{ and } X_k > b) + \frac{1}{k}(1-P(A))$$

Here:

$$P(A) = P[\bigcup_{i=1}^{k}(X_i > b \text{ and } a < X_j \leq b \quad \forall\, j \neq i)] = (.01)^{k-2}\,\frac{k-1}{2} + (.01)^{k-1}\cdot\frac{1}{2}$$

This gives:

$$E_{\underline{\theta}}\psi_k^\delta = \frac{1}{k} - \left(\frac{P(A)}{k} - \frac{1}{2}(.01)^{k-1}\right)$$

$$= \frac{1}{k} - (.24)(.01)^{k-2}\left(\frac{k-1}{k}\right) < \frac{1}{k} \quad .$$

(d) Consider the following procedure δ given by:

If $x_i = x_{[k]}$ and $i<k$, then $\delta(a_i|\underline{x}) = 1$

If $x_k = x_{[k]}$, then $\delta(a_{[k-1]}|\underline{x}) = \delta(a_{[k]}|\underline{x}) = \frac{1}{2}$.

Here $a_{[i]} = \{\pi_{[i]}\}$.

$\delta \in \mathscr{D}_{M,0}$. Let $\Omega_k = \{\underline{\theta}: \theta_k = \theta_{[k]}\}$.
Then

$$\inf_{\underline{\theta} \in \Omega_k} P_{\underline{\theta}}(CS|\delta) = \inf_{\underline{\theta} \in \Omega_k} E_{\underline{\theta}} \psi_k^{\delta} = \frac{1}{2} \inf_{\underline{\theta} \in \Omega_k} P(X_k = X_{[k]}) = \frac{1}{2k} .$$

Q.E.D.

Remarks.

1. From the proofs of Theorem 1 (a), (c) and Theorem 2(a) we see that for the classes $\mathscr{D}_I$, $\mathscr{D}_M$ and $\mathscr{D}_{I,M}$ the same results hold when restricting attention to non-randomized procedures.

2. It does not necessarily follow from (2) that all $\delta \in \mathscr{D}$ are unbiased. However, for the class $\mathscr{D}_{I,M}$, Nagel (1970) showed that all $\delta \in \mathscr{D}_{I,M}$ are also unbiased.

3. Since $\mathscr{D}_{I,M} \subset \mathscr{D}_{I,0}$ we see from Theorem 2(a), that it is essentially required that a procedure is permutation-invariant, ordered and monotone for (2) to hold, although for non-randomized procedures it is enough that the procedure is ordered.

We conclude this section with a few observations about the discrete distribution-case. Let $\mathscr{F}_d$ be the class of all stochastically increasing discrete $F_\theta(x)$, $\theta \in \Theta \subset \mathbb{R}$. The results for $\mathscr{D}_I$, $\mathscr{D}_{I,M}$, $\mathscr{D}_{I,0}$, $\mathscr{D}_M$ are essentially the same as before. It can also be shown that for the classes $\mathscr{D}_0^n$, $\mathscr{D}_0$, $\mathscr{D}_{M,0}^n$, $\mathscr{D}_{M,0}$ we now get that $\inf P^*(\delta) = 0$ for some $F_\theta \in \mathscr{F}_d$. This differs from the results for $\mathscr{F}_c$.

3. P(CS) for Permutation-Invariant "No-Data" Rules.

By definition, δ is a no-data rule if it is independent of x, i.e. $\delta(a|\underline{x}) = \delta(a) \quad \forall\, a \in \mathscr{A}, \forall\, \underline{x}$, so that

$$\sum_{a \in \mathscr{A}} \delta(a) = 1.$$

δ is permutation-invariant if $\delta(ga) = \delta(a)$, $\forall g, \forall a$.

If $|a|$ denotes the size of a, then

$$\delta(ga) = \delta(a) \ \forall g, \ \forall a \in \mathscr{A} \Leftrightarrow \delta(a) = \delta(a') \qquad \text{if } |a| = |a'|$$

Let p_i be the probability that a subset of size i selected, i.e.

$$p_i = \sum_{\{a:|a|=i\}} \delta(a) = \binom{k}{i}\delta(\{1,\dots i\}) \quad , \quad \text{if } \delta \text{ is permutation-invariant.}$$

Let $\underline{p} = (p,\dots p_k)$. $\underline{p}$ characterizes a permutation-invariant no-data rule, since for any a with size i

$$\delta(a) = p_i/\binom{k}{i} \quad .$$

One way to select according to this rule in practice is first to select a subset size according to $\underline{P}$. Then given size i, one chooses a randomly, i.e. each subset of size i have probability $\binom{k}{i}^{-1}$ of being selected. It is readily seen (also shown by Bechhofer and Santner (1979)) that

$$P_{\underline{\theta}}(CS|\underline{p}) = \sum_{i=1}^{k} p_i \frac{i}{k} \quad ; \quad \text{independent of } \underline{\theta}$$

and

$$E_{\underline{\theta}}(S|\underline{p}) = \sum_{i=1}^{k} i p_i \quad ; \quad \text{independent of } \underline{\theta} \ .$$

Lemma 3.

If δ is a permutation-invariant no-data rule, then

$$P(CS|\delta) \geq 1/k$$

Proof.

$$P(CS|\delta) = \sum_{i=1}^{k} p_i \frac{i}{k} \geq \frac{1}{k} \sum_{i=1}^{k} p_i = 1/k \qquad \text{Q.E.D.}$$

Hence, there are no permutation-invariant no-data rule which can achieve $P^*<1/k$, showing that Gupta's statement is incorrect for this class.

We also see that the lower bound $1/k$ is achieved by the rule $\underline{p}$ with $p_1=1$.

Now, for any $P^*\geq 1/k$ there exists a no-data rule $\underline{p}$ with $P(CS|\underline{p})=P^*$. This can be seen as follows. If $1/k\leq P^*<1/2$, let for example $\underline{p}$ be given by:

$$p_1 = \frac{k+2-2kP^*}{k} \quad , \quad p_2 = \ldots = p_k = \frac{2kP^*-2}{k(k-1)}$$

Then

$$P(CS|\underline{p}) = P^*.$$

If $P^*\geq 1/2$, let $p_1 + \ldots + p_{k-1} = \frac{2(1-P^*)}{k-1}$ and $p_k = 2P^*-1$.

Then

$$P(CS|\underline{p}) = P^* .$$

Since $E(S|\underline{p}) = kP(CS|\underline{p})$, it follows from Berger (1979), (under weak regularity conditions) that for each $\frac{1}{k} \leq P^* \leq 1$ there exists a permuation-invariant no-data rule that subject to the condition

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|\delta) \geq P^*$$

~~are~~ is minimax for the risk $E_{\underline{\theta}}(S|\delta)$.

The ~~criterium~~ criterion used by Bechhofer and Santner (1979), which is to choose P^* greater than or equal to $P(CS|\underline{p}^o)$ where $\underline{p}^o$ is minimax therefore seems hard to understand.

REFERENCES

Berger, R. L. (1979). Minimax subset selection for loss measured by subset size. (To appear Ann. Statist. 7 No.6).

Bechhofer, R. E. and Santner, T. J. (1979). A note on the lower bound for the P(CS) of Gupta's subset selection procedures. Technical Report #401, Cornell University.

Gibbons, J. D., Olkin, I. and Sobel, M. (1977). Selecting and ordering populations: A new statistical methodology. John Wiley & Sons, Inc.

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

Nagel, K. (1970). On subset selection rules with certain optimality properties. Mimeograph Series #222, Department of Statist., Purdue University.