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 \( P(\text{CS}) \geq P^* \). 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, \( P(\text{CS}) = 1/2 \) 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 $\pi_1, \ldots, \pi_k$. $\pi_i$ is characterized by a real-valued parameter $\theta_i$. $X_i$ is the observation from $\pi_i$. $X_1, \ldots, X_k$ 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 $\theta$, and continuous in $x$ for each $\theta$ in the parameter-space $\Theta \subset \mathbb{R}$. This class of distribution functions is denoted by $3\subset$. Let $\Omega = \Theta^k$. Let $\theta_1 \leq \cdots \leq \theta_k \leq X_1 \leq \cdots \leq X_k$ denote the ordered $\theta_i$'s and $X_1$'s. $X(i)^{\prime}$ correspond to $U(i)$. $U(k)$ is defined to be the best population. The decision-space $\mathcal{X}$ is the set of all non-empty subsets of $(\pi_1, \ldots, \pi_k)$. A subset selection-rule $\delta$ is for each observed $x = (x_1, \ldots, x_k)$ a probability-measure $\delta(a|x)$ over $a \in \mathcal{X}$. For a procedure $\delta$, the individual selection probabilities are given by:

$$
\phi_i(x) = P(\text{selecting } \pi_i \text{ with rule } \delta|x) = \sum_{a \in \mathcal{X}} \delta(a|x)
$$
let \( \psi(i) \) correspond to \( \alpha[i] \). Then the classical \( P^* \) condition is

\[
\inf_{\theta \in \Omega} P_{\theta}(CS|\delta) = \inf_{\theta \in \Omega} E_{\theta} \psi^\delta(X) = P^*
\]

(1)

Let us for convenience denote \( \inf_{\theta \in \Omega} P_{\theta}(CS|\delta) \) by \( P^*(\delta) \). The range of possible values of \( P^*(\delta) \) will depend upon the class \( & \) under consideration.

Suppose \( \inf_{\delta \in \&} P^*(\delta) = \alpha \), then the principle is that \( P^* \) should be at least \( \alpha \), because no rules in \( \& \) can achieve a \( P^* \) less than \( \alpha \). Consider for example the class \( &_G \) of Gupta's rules (see Gupta (1965)).

\[
\psi_i = 1 \text{ iff } x_i > x_{[k]} - d, \quad d > 0.
\]

Here \( \inf_{\delta \in \&} P^*(\delta) \) is \( 1/k \), achieved by the rule corresponding to \( d=0 \).

Hence for \( &_G \), \( P^* \) 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 \( \&_I \) of permutation-invariant procedures, let \( g \) be a permutation of \( (1,\ldots,k) \) such that \( g_i \) is the new position of element \( i \) under permutation \( g \). Then \( gx \) is defined by \( (gx)_i = x_{g^{-1}(i)} \), and \( ga = \{ g_i : i \in a \} \) for \( a \in \& \).

**Definition 1.** \( \delta \in \&_I \) if for each permutation \( g \)

\[
\delta(ga|gx) = \delta(a|x) \quad \forall a \in \& \text{, } \forall x.
\]

**Definition 2.** \( \delta \) is said to be ordered if

\[
\psi^\delta_i(x) \leq \psi^\delta_j(x) \quad \text{when } x_i < x_j
\]

\( \&_0 \) denotes the class of ordered procedures.
Definition 3. $\delta$ is called monotone if for each $i$,
\[ \psi_i^\delta(x) \leq \psi_i^\delta(y) \quad \text{if} \quad x_i \leq y_i \quad \text{and} \quad x_j \geq y_j \quad \forall j \neq i. \]

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

$\mathcal{I}$, $\mathcal{O}$, $\mathcal{M}$ are the three basic classes of procedures we consider.

If $\mathcal{F}$ is one of the three or a combination of these, the basic question to answer is whether or not the following statement is true:
\[ \inf_{\delta \in \mathcal{F}} \mathbb{P}^*(\delta) = \frac{1}{k} \quad (2) \]

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) $\delta \in \mathcal{I} \not\subseteq \delta \in \mathcal{O}$ (i.e. $\mathcal{I}$ - $\mathcal{O}$ is non-empty.)

(ii) $\delta \in \mathcal{O} \not\subseteq \delta \in \mathcal{I}$ (i.e. $\mathcal{O}$ - $\mathcal{I}$ is non-empty.)

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

Here $\mathcal{I}, \mathcal{M} = \mathcal{I} \cap \mathcal{M}$ and $\mathcal{I}, \mathcal{O} = \delta^{\mathcal{I}, \mathcal{O}}$.

Proof.

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

(ii). Consider the following rule $\delta$:

If $X_i = X_{[k]}$: select $\pi_1$.

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

$\delta \in \mathcal{O}$, but $\delta \notin \mathcal{I}$.
(iii). First we note that $\delta \in \mathcal{E}_I$ implies that $\psi_i^\delta(x) = \psi_{g_i(x)}^\delta \quad \forall (g, i, x)$. Assume $\delta \in \mathcal{E}_{I, M}$. Let $x$ be such that $x_i = x_j$.

We shall show that $\psi_i^\delta(x) \leq \psi_j^\delta(x)$.

Let $y$ be the permutation with $g_i = j$, $g_j = i$, $g_k = k$ \quad $\forall k \neq i, j$, and let $y = gx$.

Then $y_i = x_j > x_i$ and $y_j < x_j$ \quad $\forall k \neq i$. Hence $\psi_i^\delta(x) \leq \psi_i^\delta(y)$ from Definition 3, and $\psi_j^\delta(x) = \psi_{g_j}^\delta(gx) = \psi_i^\delta(y) \geq \psi_j^\delta(x)$, which proves the first statement. Let now $k > 3$, and consider the following rule.

$$\delta: \text{ select } \pi_i = x_i = \min \left( x_i, \frac{X[k]}{X} \right) \ ; \ X = \frac{1}{k} \sum_{i=1}^{k} x_i$$

is clearly in $\mathcal{E}_{I, 0}$. We shall show that $\delta$ is not monotone.

Let $x_i = 3/2$ for $i \leq k - 1$ and $x_k = 2$, and let $y_i = 0$ for $i < k - 2$, $y_{k - 1} = 3/2$, $y_k = 2$.

Here $x_{k-1} = y_{k-1}$ and $y_j < x_j$ \quad $\forall j \neq k - 1$. It is readily seen that $\psi_{k-1}(x) = 1$ and $\psi_{k-1}(y) = 0$. Q.E.D.

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

Theorem 1.

(a) $\inf_{\delta \in \mathcal{E}_I} P^*(\delta) = 0$, provided $F(x) > 1$ as $\Theta = \inf \Theta$.

(b) $\inf_{\delta \in \mathcal{E}_0} P^*(\delta) = \left( \frac{1}{k} \right)^2$ \quad $\forall F \in \mathcal{F}_C$.

(c) $\inf_{\delta \in \mathcal{E}_M} P^*(\delta) = 0$ \quad $\forall F \in \mathcal{F}_C$.

Proof.

(a) Consider the rule

$$\delta: \text{ select } \pi_i = x_i = \frac{X[k]}{X}$$

Then $\delta \in \mathcal{E}_I$ so we may assume $\theta_i = \Theta[i]$. 

(b) $\inf_{\delta \in \mathcal{E}_0} P^*(\delta)$ follows from the results in (a) and (b) above.
Then
\[
P_\hat{\theta}(CS|\delta) = P(X_k \leq X_i ; \forall i \leq k-1) = \int \prod_{i=1}^{k-1} P(X_i \geq x) dF_{\hat{\theta}_k}(x)
\]
\[
= \int \prod_{j=1}^{k-1} (1-F_{\hat{\theta}_j}(x)) dF_{\hat{\theta}_k}(x).
\]

Let \(\theta_j \to \inf \cap \forall \ j \leq k-1\) and keep \(\theta_k\) fixed. Then \(F_{\hat{\theta}_j}(x), \forall x\) and from Lebesgue convergence theorem \(P_{\hat{\theta}}(CS|\delta) \to 0\). Hence \(P^*(\delta) = 0\).

(b) We observe that \(\neq \phi\) implies \(\sum_{a \in \Omega} \delta(a|x) = 1 \forall x\) and hence
\[
\sum_{i=1}^{k} \delta_\phi(x) \geq \sum_{a \in \Omega} \delta(a|x) = 1. \tag{3}
\]

Let \(\delta \in \mathcal{E}_0\). Then \(x_i = x[k] \Rightarrow \psi_i(x) \geq 1/k\) from (3).

Hence:
\[
P_{\hat{\theta}}(CS|\delta) = \mathbb{E}_{\hat{\theta}}(\psi_1(\delta) \geq \prod_{i=1}^{k} dF_{\hat{\theta}_i}(x) \geq \frac{1}{k} P_{\hat{\theta}}(X(k) = x[k])
\]
\[
\{X : X(k) = x[k]\}
\]

Now \(\inf_{\theta \in \Omega} P_{\hat{\theta}}(X(k) = x[k]) = P_{\theta_1 = \ldots = \theta_k}(X(k) = x[k]) = 1/k\).

Hence:
\[
P_{\hat{\theta}}(CS|\delta) \geq (1/k)^2 \quad \forall \theta \in \Omega.
\]

The lower bound is achieved by the following rule in \(\mathcal{E}_0\):

If \(x_i = x[k]\) and \(i < k\): \(\delta(a_i|x) = 1\)

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

Here \(a_i = (\pi_i)\).

(c) is obvious e.g. let \(\delta\) be given by:
\[
\psi_i(x) = 1 \Rightarrow x_i \geq x[k-1] \text{ for } i \leq k-1
\]
\[
\psi_k(x) = 0 \quad \forall x.
\]
At \( \epsilon_M \) and if \( \eta_k = \max \eta_i \), then \( P_0(\text{CS} | \delta) = 0 \). 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 \( \delta \) is unbiasedness.

**Definition 4.** \( \delta \) is said to be unbiased if

\[
\forall i < j \Rightarrow E_\delta \psi_\delta(i) \leq E_\delta \psi_\delta(j)
\]

(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

\[
E_\delta(S | \delta) = \sum_{i=1}^{k} E_\delta \psi_\delta(i)
\]

**Lemma 2.**

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

**Proof.** Let \( \delta \in \Omega \), arbitrary, and assume that \( \delta \) is unbiased. Then

\[
k \cdot P_0(\text{CS} | \delta) = k E_\delta \psi_\delta(k) \geq \sum_{i=1}^{k} E_\delta \psi_\delta(i) = E_\delta(S | \delta) \geq 1, \text{ since } S \geq 1.
\]

Hence \( P^*(\delta) \geq 1/k \). 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 \( \mathcal{L}_1, \mathcal{L}_0, \mathcal{L}_M \). 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. \( \delta \) is non-randomized if for each \( x \) there exists \( a \in A \), such that \( \delta(a | x) = 1 \). For a given class \( \mathcal{L} \), let \( \mathcal{L}^n \) denote the class of non-randomized procedures in \( D \). E.g. \( \mathcal{L}_0^n \) is the
class of non-randomized ordered procedures. Our basic question is now answered by the following results.

Theorem 2.
(a) \( \inf_{\delta \in \mathcal{I}, M} P^*(\delta) = 1/k \), \( \forall F_0 \in \mathcal{F}_c \).
(b) \( \inf_{\delta \in \mathcal{I}_0^\mathbb{N}} P^*(\delta) = \inf_{\delta \in \mathcal{I}_0, 0} P^*(\delta) = 1/k \); \( \forall F_0 \in \mathcal{F}_c \).

Here \( \mathcal{I}_M, 0 = \mathcal{I}_M \cap \mathcal{I}_0 \).

(c) \( \inf_{\delta \in \mathcal{I}_0, 0} P^*(\delta) < 1/k \) for some \( F_0 \in \mathcal{F}_c \).

(d) \( \inf_{\delta \in \mathcal{I}_M, 0} P^*(\delta) \leq 1/2k \), \( \forall F_0 \in \mathcal{F}_c \).

Proof.
(a) Let \( \delta \in \mathcal{I}_{1, M} \). Then from Nagel (1970):
\[
P^*(\delta) = \inf_{\Theta \in \Omega_0} P_{\Theta} (CS|\delta) , \quad \text{where } \Omega_0 = \{ \emptyset : \emptyset = \ldots = \emptyset_k \}.
\]
For \( \Theta \in \Omega_0 \): \( E_{\Theta}(\psi^\delta_0) = \ldots = E_{\Theta}(\psi^\delta_k) = P_{\Theta}(CS|\delta) \) since \( \delta \in \mathcal{I}_1 \).

Hence: \( E_{\Theta}(S|\delta) = kP_{\Theta}(CS|\delta) \). Since \( S \geq 1 \), it follows that \( \inf_{\mathcal{I}_{1, M}} P^*(\delta) \geq 1/k \).

The lower bound is obtained by the rule that selects \( n_1 \) if and only if \( X_1 = X_{[k]} \).

(b) Let now \( \delta \in \mathcal{I}_0^\mathbb{N} \). Then \( X(k) = X_{[k]} = X(k) > X(i) \) \( \forall i \leq k-1 \) with \( F_0 \)-probability 1, which implies \( \psi(k)(X) = 1 \) with probability 1. Hence
\[
P_{\Theta}(CS|\delta) = E_{\Theta}(\psi^\delta_0) = P_{\Theta}(X(k) = X_{[k]}) \geq 1/k .
\]

The lower bound is attained by the same rule as in (a).
(c) Let $F_0, C, C_C$ be such that there exists $0 < a^i, b^i$ for which

$$F_0^1(a) = .49 \quad 1 - F_0^1(b) = .5$$
$$F_0^{i+1}(a) = .01 \quad 1 - F_0^{i+1}(b) = .5$$

Let $\omega = 0 \ldots 0_{k-1} = 0^i, \Omega_k = 0^n$.

Consider the following rule.

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

for $i = 1, \ldots k$

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

It is readily seen that $\delta \in I, 0$.

$$P_0(CS|\delta) = E_0^\delta = P(a < X_i < b, \forall i \leq k-1 \text{ and } X_k > b) + \frac{1}{k}(1 - P(A))$$

Here:

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

This gives:

$$E_0^\delta = \frac{1}{k} - P(A) - \frac{1}{2}(.01)^k - 1$$

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

(d) Consider the following procedure $\delta$ given by:

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

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

Here $a[i] = \{ a[i] \}$. 

\[ \delta \in \mathcal{\mathcal{B}}_{I,0}. \quad \text{Let } \Omega_k = \{ \Theta_k : \Theta_k = \Theta_{-k} \}. \]

Then

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

Q.E.D.

Remarks.

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

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

3. Since \( \mathcal{\mathcal{B}}_{I,M} \subseteq \mathcal{\mathcal{B}}_{I,0} \) 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 \( \mathcal{\mathcal{J}}_{\mathcal{C}} \) be the class of all stochastically increasing discrete \( F_\Theta(x), \Theta \in \Theta \subseteq \mathbb{R}. \) The results for \( \mathcal{\mathcal{B}}_{I}, \mathcal{\mathcal{B}}_{I,M}, \mathcal{\mathcal{B}}_{I,0}, \mathcal{\mathcal{B}}_{M} \) are essentially the same as before. It can also be shown that for the classes \( \mathcal{\mathcal{B}}_{I,0}^{\mathbb{N}}, \mathcal{\mathcal{B}}_{0}^{\mathbb{N}}, \mathcal{\mathcal{B}}_{M,0}^{\mathbb{N}}, \mathcal{\mathcal{B}}_{0,0} \) we now get that \( \inf P^*(\delta) = 0 \) for some \( F_\Theta \in \mathcal{\mathcal{C}}_{\mathcal{B}} \). This differs from the results for \( \mathcal{\mathcal{C}} \).

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

By definition, \( \delta \) is a no-data rule if it is independent of \( x \), i.e.

\[
\delta(a|x) = \delta(a) \quad \forall a \in \mathcal{\mathcal{A}}, \forall x, \quad \text{so that}
\]


\[
\sum_{a \in \mathbb{A}} \delta(a) = 1.
\]

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

If \(|a|\) denotes the size of \(a\), then

\[
\delta(ga) = \delta(a) \quad \forall g, \forall a \in \mathbb{A} : \delta(a) = \delta(a') \quad \text{if } |a| = |a'|.
\]

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

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

Let \(p = (p_1, \ldots, p_k)\). \(p\) characterizes a permutation-invariant no-data rule, since for any \(a\) with size \(i\)

\[
\delta(a) = \frac{p_i}{\binom{k}{i}}.
\]

One way to select according to this rule in practice is first to select a subset size according to \(p\). 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 Becher and Santner (1979)) that

\[
P_0(CS|p) = \sum_{i=1}^{k} \frac{p_i i}{k} \quad \text{; independent of } \theta
\]

and

\[
E_0(S|p) = \sum_{i=1}^{k} ip_i \quad \text{; independent of } \theta.
\]

**Lemma 3.**

If \(\delta\) is a permutation-invariant no-data rule, then

\[
P(CS|\delta) \geq \frac{1}{k}
\]

**Proof.**

\[
P(CS|\delta) = \sum_{i=1}^{k} \frac{p_i i}{k} \geq \frac{1}{k} \sum_{i=1}^{k} p_i = 1/k
\]

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 $p$ with $p_1 = 1$.

Now, for any $P^* > 1/k$ there exists a no-data rule $p$ with $P(\text{CS}|p) = P^*$.

This can be seen as follows. If $1/k < P^* < 1/2$, let for example $p$ be given by:

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

Then

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

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

Then

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

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

$$\inf_{\delta \in \Omega} P_0(\text{CS}|\delta) \geq P^*$$

are minimax for the risk $E_0(S|\delta)$.

The criterion used by Bechhofer and Santner (1979), which is to choose $P^*$ greater than or equal to $P(\text{CS}|p^0)$ where $p^0$ is minimax therefore seems hard to understand.
REFERENCES


