*Inequalities in Statistics and Probability*
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 199–205

# APPLICATIONS OF A UNIFIED THEORY OF MONOTONICITY IN SELECTION PROBLEMS

BY ROGER L. BERGER[1] and FRANK PROSCHAN[2]
*North Carolina State University and Florida State University*

In this paper, the general monotonicity results concerning selection problems derived by Berger and Proschan (1984) are reviewed. They are then applied to several different formulations of the selection problem. These include comparison with a control and restricted subset selection problems. Several classes of selection rules previously proposed in the literature are shown to possess the monotonicity properties. In addition, a new class of rules for the restricted subset selection formulation is proposed and shown to possess the monotonicity properties.

**1. Introduction.** In this paper we study some monotonicity properties of ranking and selection rules.

Recall that in a selection problem the general goal is to determine which of several populations possesses the largest value of some parameter. Based on random observations from the populations, a selection rule selects a subset of the populations and leads to an assertion such as, "The population with the largest parameter is in the selected subset." (Different formulations of the selection problem entail different assertions resulting from the selection rule.) A reasonable selection rule should be more likely to choose populations with larger parameters rather than populations with smaller parameters. This property of selection rules is called monotonicity.

In this paper we study some general monotonicity properties of a broad class of selection rules in a unified manner. We also discuss applications of these general results to several different formulations of the selection problem.

In symbols, let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random observation with distribution $F(\mathbf{x}; \lambda)$, where the unknown parameter vector $\lambda = (\lambda_1, \ldots, \lambda_n) \in \Lambda \subset \mathcal{R}^n$. The general goal of a selection problem is to decide which of the coordinates of $\lambda$ are the largest or which are larger than a value $\lambda_0$ (possibly unknown). A (nonrandomized) selection rule $S(\mathbf{x})$ is any measurable mapping from the sample space $\mathcal{X}$ of $\mathbf{X}$ into the set of subsets of $\{1, \ldots, n\}$. Having observed $\mathbf{X} = \mathbf{x}$, the selection rule $S$ asserts that the largest parameters are in $\{\lambda_i : i \in S(\mathbf{x})\}$. The subset $S(\mathbf{X})$ may be of fixed or random size depending on the formulation of the selection problem under consideration. See, for example, Bechhofer (1954) (fixed size), Gupta and Sobel (1958) (random size), and Gupta (1965) (random size).

Gupta (1965) calls a selection rule *montone* if

$$\lambda_i \geq \lambda_j \text{ implies } P_\lambda(i \in S(\mathbf{X})) \geq P_\lambda(j \in S(\mathbf{X})). \tag{1.1}$$

Monotonicity is a desirable property of a selection rule since the selected subset is supposed to consist of the large values of $\lambda_i$. On a case by case basis, various authors have shown that their proposed selection rules are monotone. Monotonicity has not been investi-

[1] This work was supported in part by the U.S. Army Research Office Grant No. DAAG 29–82–K–0168.
[2] Research sponsored by the Air Force Office of Scientific Research Grant AFOSR 82–K–0007.

*AMS 1980 subject classifications.* Primary, 62F07; secondary, 62H10.
Key words and phrases: Ranking and selection, decreasing in transposition, monotonicity, comparison with a control, restricted subset selection, just rules.

gated for some formulations of the selections problem even though it is a desirable property in all formulations.

In this paper we review the results in Berger and Proschan (1984) (BP(1984)). These results concern some general notions of monotonicity which include the previously discussed notion of Gupta (1965). BP(1984) show in a simple unified way that a broad class of selection rules (which includes rules proposed for various formulations of the selection problem) possess these monotonicity properties. BP(1984) also discuss the application of these results to selection rules proposed by Bechhofer (1954), Gupta and Sobel (1958), and Gupta (1965). In the present paper, we discuss the application of these results to other formulations of the selection problem and other classes of selection rules considered in the literature. Also, a new class of selection rules for the restricted subset selection problem is proposed and shown to possess these general monotonicity properties.

The monotonicity properties we consider are the following. Let $A = \{a_1, \ldots, a_k\}$ and $b = \{b_1, \ldots, b_k\}$ denote two subsets of $\{1, \ldots, n\}$ with $|A| = |B| = k$, where $|\cdot|$ denotes subset size. Subset $A$ *is better than* $B$ if, for some arrangements $a_{i(1)}, \ldots, a_{i(k)}$ and $b_{j(1)}, \ldots, b_{j(k)}$ of the elements of $A$ and $B$, $\lambda_{a_{i(r)}} \geq \lambda_{b_{j(r)}}$ for every $r = 1, \ldots, k$. If $A$ is better than $B$, then each of the following inequalities would be desirable for a selection rule:

$$P_\lambda[|A \cap S(\mathbf{X})| \geq m] \geq P_\lambda[|B \cap S(\mathbf{X})| \geq m] \text{ for every } m \in \mathscr{R}^1; \tag{1.2}$$

(In words, $P_\lambda$ [at least $m$ elements of $A$ are selected] $\geq P_\lambda$ [at least $m$ elements of $B$ are selected].)

$$P_\lambda(A = S(\mathbf{X})) \geq P_\lambda(B = S(\mathbf{X})); \tag{1.3}$$

and

$$P_\lambda(|A^c \cap S(\mathbf{X})| \leq m) \geq P_\lambda(|B^c \cap S(\mathbf{X})| \leq m) \text{ for every } m \in \mathscr{R}^1; \tag{1.4}$$

where $A^c$ and $B^c$ are the complements of $A$ and $B$, respectively.

Some special cases are of particular interest. By setting $m = k$ in (1.2) we obtain $P_\lambda[A \subset S(\mathbf{X})] \geq P_\lambda[B \subset S(\mathbf{X})]$; i.e., the better subset is more likely than the worse subset to be included in the selected subset. From the special case $m = k = 1$ in (1.2), we obtain the classical montonicity property (1.1). By setting $m = 0$ in (1.3), we obtain $P_\lambda[A \supset S(\mathbf{X})] \geq P_\lambda[B \supset S(\mathbf{X})]$; i.e., the selected subset is more likely to be in the better subset than in the worse subset.

In Section 2 the class of selection rules is discussed. The assumptions regarding the distribution $F(\mathbf{x}; \lambda)$ are discussed in Section 3. The montonicity results from BP(1984) are presented and applied to three formulations of the selection problem in Section 4. In Section 5, an extension of these results to include additional parameters and statistics is presented and applied to another formulation of the selection problem.

**2. A Class of Selection Rules.** In this section, a broad class of selection rules is described. All of the rules in this class will have the general monotonicity properties (1.2), (1.3), and (1.4).

Let $\pi = (\pi_1, \ldots, \pi_n)$ denote a permutation of $(1, \ldots, n)$. For any $\mathbf{x} \in \mathscr{R}^n$, let $\mathbf{x} \circ \pi$ denote $(x_{\pi_1}, \ldots, x_{\pi_n})$. Let $I_C(\cdot)$ denote the indicator function of the set $C$.

A nonrandomized selection rule $S(\mathbf{x})$ can be defined by specifying its individual selection probabilities, $\psi_1(\mathbf{x}), \ldots, \psi_n(\mathbf{x})$. These are defined by $\psi_i(\mathbf{x}) = I_{S(\mathbf{x})}(i)$. We will be interested in selection rules which satisfy the following for every $\mathbf{x} \in \mathscr{R}^n$ and every $i \in \{1, \ldots, n\}$

$$\text{If } \psi_i(\mathbf{x}) = 1 \text{ and } x_j \geq x_i, \text{ then } \psi_j(\mathbf{x}) = 1, \tag{2.1}$$

and

(2.2) $\psi_{\pi_i}(\mathbf{x}) = \psi_i(\mathbf{x} \circ \pi)$ for every permutation $\pi$. Rules satisfying (2.1) have been called *natural* in some selection literature (for example, Eaton, 1967).

Nagel (1970) and Gupta and Nagel (1971) defined and investigated a class of selection rules called just rules. A selection rule is *just* if, for every $i \in \{1, \ldots, n\}$, $\psi_i(\mathbf{x})$ is a nondecreasing function of $x_i$ and a nonincreasing function of $x_j$, $j \neq i$. If a selection rule is just and satisfies (2.2), then the rule satisfies (2.1). To see this, let $\pi$ be the permutation defined by $\pi_i = j; \pi_j = i;$ and $\pi_r = r$, $r = 1, \ldots, n$, $r \neq i$ or $j$. Then if $x_j \geq x_i$,

$$\psi_j(\mathbf{x}) \geq \psi_j(\mathbf{x} \circ \pi) = \psi_{\pi_j}(\mathbf{x} \circ \pi \circ \pi) = \psi_i(\mathbf{x}).$$

The inequality follows from the justness, and the first equality follows from (2.2)

Almost all the selection rules which have been proposed in the literature for the models described in Section 3 are just rules satisfying (2.2). Thus almost all of the selection rules which have been proposed over the last thirty years satisfy the general monotonicity properties (1.2), (1.3), and (1.4); the results in Section 4 will give a simple unified proof of this fact, as well as other consequences.

**3. The Model and Key Mathematical Ideas.** In this section, the concept of a decreasing in transposition (DT) function is introduced. The effect of assuming that the density of $\mathbf{X}$ is DT is discussed.

Let $\pi$ and $\pi'$ be two permutations and $i$ and $j$ two elements of $\{1, \ldots, n\}$ such that $i < j$; $\pi_i < \pi_j$; $\pi'_i = \pi_j$; $\pi'_j = \pi_i$; and $\pi'_r = \pi_r$, $r = 1, \ldots, n$, $r \neq i$ or $j$. We say that $\pi'$ is a *simple transposition* of $\pi$; in symbols $\pi >^t \pi'$.

The concept of a decreasing in transposition function plays a central role in our derivation of the general monotonicity properties. A real valued function $g(\mathbf{x}; \lambda)$ on $\mathcal{R}^{2n}$ is *decreasing in transposition (DT)* if

(3.1) $g(\mathbf{x}; \lambda) = g(\mathbf{x} \circ \pi; \lambda \circ \pi)$ for every $\mathbf{x} \in \mathcal{R}^n$, $\lambda \in \mathcal{R}^n$ and every permutation $\pi$,

and

(3.2) $x_1 \leq \ldots \leq x_n$, $\lambda_1 \leq \ldots \leq \lambda_n$, and $\pi >^t \pi'$ imply $g(\mathbf{x}; \lambda \circ \pi) \geq g(\mathbf{x}; \lambda \circ \pi')$.

Hollander, Proschan, and Sethuraman (1977) (HPS(1977)) present a detailed investigation of DT functions. The DT property is called arrangement increasing by Marshall and Olkin (1979).

We assume that the observation vector $\mathbf{X} = (X_1, \ldots, X_n)$ has a density $g(\mathbf{x}; \lambda)$ with respect to a measure $\sigma(\mathbf{x})$, where $\sigma$ satisfies $\int_A d\sigma(\mathbf{x}) = \int_A d\sigma(\mathbf{x} \circ \pi)$ for each permutation $\pi$ and Borel set $A \subset \mathcal{R}^n$. We assume that $g$ is DT. HPS(1977) list several discrete and continuous densities which are DT. For example, if $g(\mathbf{x}; \lambda) = \pi_{i=1}^n h(x_i; \lambda_i)$ and $h$ is $TP_2$, then $g$ is DT.

By Theorem 4.1 of HPS(1977), if $\mathbf{X}$ has a DT density, then the coordinates of $\mathbf{X}$ are more likely to be in the same order as the coordinates of $\lambda$ than in any other order. Furthermore, the probability of a rank order for $\mathbf{X}$ decreases as the rank order becomes more transposed from the order of $\lambda$. This behavior is typical in selection problems. Usually $X_i$ is an estimate of $\lambda_i$ and so the $X_i$'s are expected to be in approximately the same order as the $\lambda_i$'s. This leads to the use of selection rules satisfying (2.1). Most of the models considered in the selection literature are models with DT densities.

A class of selection problems considered in the literature which do *not* have DT densities are problems involving unequal sample sizes. See Berger and Gupta (1980) for several references. For example, in the problem of selecting the normal population with the largest normal mean, the density of the sample means will not be DT if the population variances are equal but the sample sizes are unequal. The density does not satisfy (3.1). Thus in their present form the results of BP(1984) do not apply to these selection problems.

**4. Monotonicity Properties.** In this section we state the monotonicity results of BP(1984). Then we apply these results to three different formulations of the selection problem.

THEOREM 4.1. *Suppose (a) the density $g(\mathbf{x} ; \lambda)$ of $\mathbf{X}$ is DT, (b) the individual selection probabilities of the selection rule S satisfy (2.1) and (2.2), (c) $A \subset \{1, \ldots, n\}$, $B \subset \{1, \ldots, n\}$, and A is better than B. Then*

$$P_\lambda[|A \cap S(\mathbf{X})| \geq m] \geq P_\lambda[|B \cap S(\mathbf{X})| \geq m] \text{ for every } m \in \mathcal{R}^1, \tag{4.1}$$

$$P_\lambda[A = S(\mathbf{X})] \geq P_\lambda[B = S(\mathbf{X})], \tag{4.2}$$

and

$$P_\lambda[|A^c \cap S(\mathbf{X})| \leq m] \geq P_\lambda[|B^c \cap S(\mathbf{X})| \leq m] \text{ for every } m \in \mathcal{R}^1. \tag{4.3}$$

The proof of Theorem 4.1 is given in BP(1984). It is based on the fact that the indicator functions of the desired events are DT functions of $\mathbf{X}$ and a vector of 1's and 0's indicating which elements of $\{1, \ldots, n\}$ are in $A$ (or $B$) and which are not. By the Composition Theorem of HPS(1977), the probabilities are DT functions of $\lambda$ and the vector of 1's and 0's. The inequalities then follow, since the vector of 1's and 0's for $B$ is a transposition of the corresponding vector for $A$.

We now present some examples of selection rules satisfying the conditions of Theorem 4.1.

*Example 4.1.* (Restricted subset selection). Santner (1975) introduced the restricted subset formulation of the selection problem. In this formulation, a subset of random size is selected. The size of the selected subset must not exceed $m$, a fixed constant satisfying $1 \leq m \leq n$. Santner (1975) proposed and studied a class of restricted subset selection rules. We will propose a class of rules which satisfy the conditions of Theorem 4.1. and thus possess the monotonicity properties (4.1), (4.2), and (4.3).

Santner (1975) proposed this class of restricted subset selection rules. Let $X_{[1]} \leq \ldots \leq X_{[n]}$ denote the ordered values of $X_1, \ldots, X_n$. Let $h^{-1}(z)$ be a nondecreasing real valued function of the real variable $z$ satisfying $h^{-1}(z) \leq z$. Then a rule in Santner's class is defined by:

Include $i$ in the selected subset if and only if

$$X_i \geq \max(X_{[n-m+1]}, h^{-1}(X_{[n]})). \tag{4.4}$$

Actually Santner places more restrictions on the function $h^{-1}$ than we have stated but these conditions are all that are important for our discussion.

We propose the following class of restricted subset selection rules. For any $\mathbf{x} \in \mathcal{R}^n$, let $\mathbf{x}^i = (x_1^i, \ldots, x_{n-1}^i)$ be the vector obtained by deleting $x_i$ and arranging the remaining $n-1$ components of $\mathbf{x}$ in increasing order. Let $p$ be a real valued function defined on $\mathcal{R}_*^{n-1} \equiv \{y \in \mathcal{R}^{n-1}: y_1 \leq \ldots \leq y_{n-1}\}$ which is nondecreasing in each coordinate. Assume that $p(\mathbf{y}) \leq y_{n-1}$ for every $\mathbf{y} \in \mathcal{R}_*^{n-1}$. Finally we assume, as Santner (1975) did, that $g(\mathbf{x} ; \lambda)$ is a density with respect to Lebesgue measure on $\mathcal{R}^n$; thus no coordinates of $\mathbf{X}$ are tied with probability one. A class of restricted subset selection rules is defined by:

Include $i$ in the selected subset if and only if

$$X_i \geq \max(X_{[n-m+1]}, p(\mathbf{X}^i)). \tag{4.5}$$

The class of restricted subset selection rules defined by (4.5) contains the class of rules defined by (4.4). Every rule in the class (4.5) also satisfies (2.1) and (2.2) and has the general monotonicity properties (4.1), (4.2), and (4.3) if the density of $\mathbf{X}$ is DT.

Let $S$ be a rule in the class (4.5). To see that $S$ satisfies (2.2), note that for any permutation $\pi$, $\mathbf{x} \in \mathcal{R}^n$, and $i \in \{1, \ldots, n\}$, $x_{\pi_i} = (\mathbf{x} \circ \pi)_i$ and $\mathbf{x}^{\pi_i} = (\mathbf{x} \circ \pi)^i$. To see that $S$ satisfies (2.1), note that if $x_j \geq x_i$ (say, $x_i = x_{[r]}$ and $x_j = x_{[s]}$ where $r < s$) then $x_t^i = x_t^j$, $t = 1, \ldots, r-1$, $x_t^i \geqslant x_t^j$, $t = r, \ldots, s-1$, and $x_t^i = x_t^j$, $t = s, \ldots, n-1$. Thus, by the monotonicity of $p$, $p(\mathbf{x}^i) \geqslant p(\mathbf{x}^j)$. So if $\psi_i(\mathbf{x}) = 1$, then $x_j \geqslant x_i \geqslant \max(x_{[n-m+1]}, p(\mathbf{x}^i)) \geqslant \max(x_{[n-m+1]}, p(\mathbf{x}^j))$ and $\psi_j(\mathbf{x}) = 1$.

To see that Santner's (1975) class of restricted subset selection rules is a subset of the class (4.5), let $h^{-1}$ be a function which defines a rule in (4.4). For $\mathbf{y} \in \mathcal{R}_*^{n-1}$ define $p(\mathbf{y}) = h^{-1}(y_{n-1})$. By use of the properties of $h^{-1}$, our restrictions on $p$ are easily verified. If $x_i = x_{[n]}$ both the rule defined with $p$ and the rule defined with $h^{-1}$ include $i$ in the selected subset. If $x_i \neq x_{[n]}$, then $x_{n-1}^i = x_{[n]}$ and $p(\mathbf{x}^i) = h^{-1}(x_{[n]})$. Thus Santner's rule from (4.4) with $h^{-1}$ is equivalent to the rule from (4.5) defined with $p$.

Santner (1975) showed that every rule in the class he considered had the classical monotonicity property (1.1). Santner assumed that the coordinates of $\mathbf{X}$ are independent, the density of $x_i$ is $g(x_i\,;\,\lambda_i)$, and the family $g(x\,;\,\lambda)$ is stochastically increasing. Under these same conditions, using a proof very similar to Santner's, we can show that every rule in the class (4.5) has the montonicity property (1.1). In addition, we can conclude, using Theorem 4.1, that any rule in the class (4.5) satisfies the monotonicity properties (4.1), (4.2), and (4.3) if the density of $\mathbf{X}$ is DT. Inequality (4.1) includes property (1.1) as a special case.

*Example 4.2.* (Comparison with a control). Lehmann (1961) formulated the comparison with a control problem in this way. A population is called *positive* if $\lambda_i \geqslant \lambda_0 + \Delta$ and *negative* if $\lambda_i \leqslant \lambda_0$, where $\Delta > 0$ and $\lambda_0$ are fixed constants. The general goal is to select a subset containing positive populations.

Lehmann (1961) derived minimax rules which minimize $\sup_\Lambda R(\lambda, S)$ subject to $\inf_{\Lambda'} T(\lambda, S) \geqslant \gamma$. Here $\gamma$ is a fixed constant, $\Lambda'$ is the subset of $\Lambda$ for which at least one population is positive, $R$ is either of two criteria concerning the number of negative populations selected, and $T$ is any of four criteria concerning the number of positive populations selected.

One application of Lehmann's (1961) results is the following. Assume $X_1, \ldots, X_n$ are independent. Assume $X_i$ is a sufficient statistic computed from a sample from the $i$-th population. Assume the density $g(x_i\,;\,\lambda_i)$ of $X_i$ possesses the monotone likelihood ratio property. Then the rule defined by $\psi_i(\mathbf{x}) = 1, \alpha, 0$ according as $X_i >, =, < C$ is minimax, where $\alpha$ and $C$ are determined by $E_{\lambda_0+\Delta}\,\psi_i(X_i) = \gamma$.

The above assumptions imply that the density of $\mathbf{X}$ is DT. $\psi_i(\mathbf{x})$ will satisfy (2.1) and (2.2) if $\alpha = 0$ or $\alpha = 1$. This will be the case if $g(x\,;\,\lambda)$ is a density with respect to Lebesgue measure. It will also be the case for certain values of $\lambda_0$ and $\Delta$ if $g(x\,;\,\lambda)$ is a Poisson or binomial density. In each of these cases, Theorem 4.1 implies that the minimax rule will satisfy the montonicity properties (4.1), (4.2), and (4.3).

*Example 4.3.* (Just subset selection rules). In Section 2, it was shown that all just rules which satisfy (2.2) also satisfy (2.1). Thus, if the density of $\mathbf{X}$ is DT, any just rule satisfying (2.2) has the monotonicity properties (4.1), (4.2), and (4.3).

Historically, the concept of justness has been used only with the unrestricted subset selection formulation of Gupta (1965). For example, Bjornstad (1981) recently investigated a large class of just rules. But the concept of justness is equally appealing for other formulations of the selection problem. Indeed, the rules considered in Examples 4.1 and 4.2 are just.

**5. Additional Parameters and Statistics.** BP(1984) prove this more general monotonicity result applying to models which include other parameters besides $\lambda$ and other statistics in addition to $\mathbf{X}$. Let $\mathbf{Y}$ be a statistic, possibly a vector, with sample space $\mathcal{Y}$. Let $\nu$ be a parameter, possibly a vector, with a set of possible values denoted by $N$.

THEOREM 5.1 *Assume that* $(\mathbf{X}, \mathbf{Y})$ *has a density* $g(\mathbf{x},\mathbf{y}; \lambda,\nu)$ *with respect to a measure* $\sigma(\mathbf{x}) \times \mu(\mathbf{y})$, *where* $\sigma$ *satisfies* $\int_A d\sigma(\mathbf{x}) = \int_A d\sigma(\mathbf{x} \circ \pi)$ *for each permutation* $\pi$ *and Borel set* $A \subset \mathcal{R}^n$. *Assume that for each* $\mathbf{y} \in \mathcal{Y}$ *and* $\nu \in N$, $g(\mathbf{x},\mathbf{y}; \lambda,\nu)$ *is a DT function of* $\mathbf{x}$ *and* $\lambda$.

*Let* $\psi_1(\mathbf{x}, \mathbf{y}), \dots, \psi_n(\mathbf{x}, \mathbf{y})$ *denote the individual selection probabilities of a nonrandomized selection rule* $S(\mathbf{X}, \mathbf{Y})$. *Assume* (a) *for every* $\mathbf{y} \in \mathcal{Y}$, *if* $\psi_i(\mathbf{x}, \mathbf{y}) = 1$ *and* $x_j \geq x_i$, *then* $\psi_j(\mathbf{x}, \mathbf{y}) = 1$; (b) $\mathbf{y} \in \mathcal{Y}$, $\mathbf{x} \in \mathcal{R}^n$, $i \in 1, \dots, n$, *and* $\pi$ *a permutation imply* $\psi_{\pi_i}(\mathbf{x}, \mathbf{y}) = \psi_i(\mathbf{x} \circ \pi, \mathbf{y})$.

*Let* $A \subset \{1, \dots, n\}$ *and* $B \subset \{1, \dots, n\}$. *If* $A$ *is better than* $B$, *then*

(5.1) $$P_{\lambda,\nu}[|A \cap S(\mathbf{X}, \mathbf{Y})| \geqslant m] \geqslant P_{\lambda,\nu}[|B \cap S(\mathbf{X}, \mathbf{Y})| \geqslant m] \text{ for every } m \in \mathcal{R}^1,$$

(5.2) $$P_{\lambda,\nu}[|A^c \cap S(\mathbf{X}, \mathbf{Y})| \leqslant m] \geqslant P_{\lambda,\nu}[|B^c \cap S(\mathbf{X}, \mathbf{Y})| \leqslant m] \text{ for every } m \in \mathcal{R}^1,$$

*and*

(5.3) $$P_{\lambda,\nu}[A = S(\mathbf{X}, \mathbf{Y})] \geqslant P_{\lambda,\nu}[B = S(\mathbf{X}, \mathbf{Y})].$$

The proof of Theorem 5.1 may be found in BP(1984).

*Example 5.1.* (Comparison with an unknown control). Tong (1969) formulated the problem of comparison with a control in this way. $X_0, X_1, \dots, X_n$ are independent normal random variables with means $\lambda_0, \lambda_1, \dots . \lambda_n$ and common known variance $\sigma^2/N_0$. The parameter $\lambda_0$ is the unknown control value. For $i = 1, \dots, n$, $\lambda_i$ is bad if $\lambda_i \leqslant \lambda_0 + \delta_1$ and $\lambda_i$ is good if $\lambda_1 \geqslant \lambda_0 + \delta_2$, where $\delta_1 < \delta_2$ are known constants. The sample size $N_0$ is chosen so that the probability that all of the good populations are selected but none of the bad populations is selected is at least a preassigned value.

In our notation, $\mathbf{X} = (X_1, \dots, X_n)$, $\lambda = (\lambda_1, \dots, \lambda_n)$, $\mathbf{Y} = X_0$ and $\nu = \lambda_0$. Let $d = (\delta_1 + \delta_2)/2$. Tong (1969) showed that the selection rule which includes $i$ in $S(\mathbf{X}, X_0)$ if and only if $X_i - X_0 > d$ is Bayes, minimax, and admissible among a class of translation invariant rules.

The conditions of Theorem 5.1 are easily verified for this selection rule and model. Thus $S(\mathbf{X}, X_0)$ possesses the general monotonicity properties (5.1), (5.2), and (5.3). For example, if $A$ is the set of good parameters and $B$ is any other set of equal size, then, by (5.3), $A$ is more likely to be the selected set than is $B$.

In other applications, $\nu$ might include nuisance parameters, which have no bearing on which $\lambda_i$'s are preferred, as well as control parameters, like $\lambda_0$. Similarly, $\mathbf{Y}$ might include estimates of nuisance parameters.

**6. Conclusion.** In this paper, we have reviewed the general monotonicity results for selection rules of BP(1984). By examples, we have indicated that almost all nonrandomized selection rules which have been proposed for models with DT densities possess the general monotonicity properties. Thus, results which have previously been derived on a case by case basis may now be obtained using this unified theory; in addition, other results may be obtained.

## REFERENCES

BECHHOFER, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Stat. 25* 16–39.

BERGER, R. L. and GUPTA, S. S. (1980). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. *Scand. J. Statist. 7* 21–26.

BERGER, R. L. and PROSCHAN, F. (1984). Monotonicity in selection problems: a unified approach. To appear in *Ann. Statist. 12*.

BJØRNSTAD, J. F. (1981). A class of Schur procedures and minimax theory for subset selection. *Ann. Statist. 9* 777–791.

EATON, M. L. (1967). Some optimum properties of ranking procedures. *Ann. Math. Statist. 38* 124–137.

GUPTA, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics 7* 225–245.

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

GUPTA, S. S. and SOBEL, M. (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist. 29* 235–244.

HOLLANDER, M., PROSCHAN, F., and SETHURAMAN, J. (1977). Functions decreasing in transposition and their application in ranking problems. *Ann. Statist. 5* 722–733.

LEHMANN, E. L. (1961). Some Model I problems of selection. *Ann. Math. Statist. 32* 990–1012.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

NAGEL, K. (1970). On subset selection rules with certain optimality properties. Ph.D. thesis, Statistics Department, Purdue University, West Lafayette, IN.

SANTNER, T. J. (1975). A restricted subset selection approach to ranking and selection problems. *Ann. Statist. 3* 334–349.

TONG, Y. L. (1969). On partitioning a set of normal populations by their locations with respect to a control. *Ann. Math. Statist. 40* 1300–1324.