The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 814-822

ON A CLASS OF SUBSET SELECTION PROCEDURES[1]

BY SHANTI S. GUPTA AND S. PANCHAPAKESAN

Purdue University and Southern Illinois University

A class of procedures is considered for the subset selection problem when the populations are from a stochastically increasing family $\{F_\lambda\}$. A theorem concerning the monotonicity of an integral associated with $\{F_\lambda\}$ which generalizes an earlier result of Lehmann is obtained. This leads to a sufficient condition for the monotonicity of the probability of a correct selection for the procedure considered. It is shown that this condition is relevant to another sufficient condition for the supremum of the expected subset size to occur when the distributions are identical. The main results are applied to the specific cases where (i) λ is a location parameter (ii) λ is a scale parameter and (iii) the case where the density $f_\lambda(x)$ is a convex mixture of a sequence of known density functions. The earlier known results are shown to follow from the general theory.

1. Introduction and summary. Let $\pi_1, \pi_2, \cdots, \pi_k$ be k independent populations. Let Λ be an interval on the real line. Associated with π_i $(i = 1, 2, \cdots, k)$ is a real-valued random variable X_i with an absolutely continuous distribution $F_i \equiv F_{\lambda_i}$, $\lambda_i \in \Lambda$, and density function $f_i \equiv f_{\lambda_i}$. It is assumed that the functional form of F_{λ_i} is known, but not the value of λ_i. Let $\lambda_{[1]} \leqq \lambda_{[2]} \leqq \cdots \leqq \lambda_{[k]}$ be the ordered λ's. The correct pairing of the ordered and the unordered λ's is not known. It is also assumed that F_λ is differentiable in λ and that $\{F_\lambda\}$, $\lambda \in \Lambda$, is a stochastically increasing (SI) family of distributions, that is to say, for $\lambda < \lambda'$, F_λ and $F_{\lambda'}$ are distinct and $F_\lambda(x) \geqq F_{\lambda'}(x)$ for all x. Let $x_1, x_2, \cdots, x_k$ be observations on $X_1, X_2, \cdots, X_k$, respectively. Based on these observations, the goal is to select a nonempty subset of the k populations with the guarantee that the probability of a correct selection, i.e. selection of a subset which includes the population associated with $\lambda_{[k]}$ $(\lambda_{[1]})$, called the best population, is at least a predetermined number P^* $(k^{-1} < P^* < 1)$. If there are more populations than one with $\lambda_i = \lambda_{[k]}$ $(\lambda_i = \lambda_{[1]})$, then we assume that one of them is tagged as the best population. Letting $P(CS\,|\,R)$ denote the probability of a correct selection using the procedure R, the probability requirement can be written as

$$\inf_\Omega P(CS\,|\,R) \geqq P^* , \tag{1.1}$$

where Ω is the space of all k-tuples $(F_1, F_2, \cdots, F_k)$. This requirement (1.1) will be referred to as the P^*-condition.

It should be pointed out that the emphasis is not so much on the generalization of the class of procedures but on the results dealing with the infimum of the probability of a correct selection and the supremum of the expected size of the

Received July 13, 1970; revised June 6, 1971.

[1] This research was supported in part by the Office of Naval Research Contract NONR-1100(26) and the Aerospace Research Laboratories Contract AF 33(615)67C1244 at Purdue University. Reproduction in whole or part is permitted for any purposes of the United States Government.

selected subset. The results of this paper are more general and present a unified view of the subset selection procedures.

The main results of this paper are given in Theorems 2.1 and 2.2. Theorem 2.1 represents a generalization of an earlier result of Lehmann [6]. Theorem 2.2 establishes the monotone behavior of an integral arising in a class of subset selection procedures. The motivation for discussing the somewhat more general class of rules is explained in the paragraph following (2.6).

2. The class of procedures R_h. Let $h \equiv h_{c,d}$, $c \in [1, \infty)$, $d \in [0, \infty)$ be a class of real-valued functions defined on the real line satisfying the following set of conditions (A): For every x belonging to the support of F_λ, (i) $h_{c,d}(x) \geqq x$, (ii) $h_{1,0}(x) = x$, (iii) $h_{c,d}(x)$ is continuous in c and d and (iv) $\lim_{d\to\infty} h_{c,d}(x) = \infty$, c fixed and/or $\lim_{c\to\infty} h_{c,d}(x) = \infty$, d fixed, $x \neq 0$. Then the procedure R_h is defined as follows. R_h: Include the population π_i in the selected subset iff

$$h(x_i) \geqq \max_{1\leqq r\leqq k} x_r \,. \tag{2.1}$$

This procedure is a slight extension of the class h_b of rules proposed and investigated by Gupta [2]. Letting $X_{(r)}$ denote the random variable of the set X_1, $X_2, \cdots, X_k$ which is associated with $\lambda_{[r]}$ and $F_{[r]} \equiv F_{\lambda_{[r]}}$ denote the corresponding cdf, we obtain

$$\begin{aligned} P(CS \mid R_h) &= P(h(X_{(k)}) \geqq X_{(r)},\ r = 1, 2, \cdots, k-1) \\ &= \textstyle\int \{\prod_{r=1}^{k-1} F_{[r]}(h(x))\} f_{[k]}(x)\, dx \,, \end{aligned} \tag{2.2}$$

where $f_{[r]}$ ($r = 1, 2, \cdots, k$) denotes the density corresponding to $F_{[r]}$ and the integral is taken over the support of the distributions which is assumed to be the same for all F_λ, $\lambda \in \Lambda$. Because $\{F_\lambda\}$ is assumed to be an SI family,

$$P(CS \mid R_h) \geqq \textstyle\int F_{[k]}^{k-1}(h(x)) f_{[k]}(x)\, dx \,. \tag{2.3}$$

Define

$$\psi(\lambda; c, d, t+1) = \textstyle\int F_\lambda^{\,t}(h(x)) f_\lambda(x)\, dx \,. \tag{2.4}$$

Then

$$\inf_\Omega P(CS \mid R_h) = \inf_{\lambda\in\Lambda} \psi(\lambda; c, d, k) \,. \tag{2.5}$$

Because of the set of conditions (A) imposed on h, we have for any $\lambda \in \Lambda$,

$$\begin{aligned} &\text{(i)} \quad \psi(\lambda; c, d, k) \geqq \frac{1}{k} \\ &\text{(ii)} \quad \psi(\lambda; 1, 0, k) = \frac{1}{k} \\ &\text{(iii)} \quad \lim_{d\to\infty} \psi(\lambda; c, d, k) = 1 \,, \quad c \text{ fixed, and/or} \\ &\qquad\ \ \lim_{c\to\infty} \psi(\lambda; c, d, k) = 1 \,, \quad d \text{ fixed} \,. \end{aligned} \tag{2.6}$$

In general, the above conditions are not enough to ensure the existence of constants c and d so that the P^*-condition is satisfied. If the condition (2.18) (to be obtained later) is satisfied then $\psi(\lambda; c, d, k)$ is non-decreasing in λ and $\inf_{\lambda \in \Lambda} \psi(\lambda; c, d, k) = \psi(\lambda_0; c, d, k)$. Then, we can evaluate the constants c and d so that P^*-condition is satisfied, provided the set of conditions (A) hold for λ_0 in case $\lambda_0 \notin \Lambda$. Under these conditions, there can be obviously several choices of c and d satisfying the P^*-condition. Some additional condition can be imposed to determine the choice of a pair (c, d). For example, if $\{F_\lambda\}$ is the family of noncentral chi-square distributions all with the same number of degrees of freedom, where λ denotes the noncentrality parameter, then one could use $h(x) = cx + d$ and choose a pair (c, d) satisfying the P^*-condition for which the expected number of populations included in the subset is minimum when $\lambda_{[1]} = \cdots = \lambda_{[k-1]} = 0 < \lambda_{[k]} = \lambda^*$.

Sufficient condition for the monotonicity of $\psi(\lambda; c, d, k)$. We now prove a theorem which leads to a sufficient condition for the monotonicity of $\psi(\lambda; c, d, k)$.

THEOREM 2.1. *Let* $\{F_\lambda\}$, $\lambda \in \Lambda$, *be a family of absolutely continuous distributions on the real line and* $\psi(x, \lambda)$ *be a real-valued function possessing continuous first partial derivatives* ψ_x *and* ψ_λ, *respectively. Then* $E_\lambda \psi(X, \lambda)$ *is non-decreasing in* λ *provided that*

$$f_\lambda(x)\psi_\lambda(x, \lambda) - \psi_x(x, \lambda)\frac{\partial}{\partial \lambda} F_\lambda(x) \geqq 0 . \tag{2.7}$$

Further $E_\lambda \psi(X, \lambda)$ *is strictly increasing in* λ *if* (2.7) *holds with strict inequality on a set of positive Lebesgue measure.*

PROOF. Let us consider $\lambda_1, \lambda_2 \in \Lambda$ such that $\lambda_1 \leqq \lambda_2$ and define

$$A_i(\lambda_1, \lambda_2) = \int \prod_{r=1, r\neq i}^{2} \psi(x, \lambda_r)\, dF_i(x) , \qquad i = 1, 2 \tag{2.8}$$

and

$$B(\lambda_1, \lambda_2) = \sum_{i=1}^{2} A_i(\lambda_1, \lambda_2) , \tag{2.9}$$

where $F_i \equiv F_{\lambda_i}$, $i = 1, 2$. We note that when $\lambda_1 = \lambda_2 = \lambda$, $B(\lambda, \lambda) = 2E_\lambda \psi(X, \lambda)$.

Integrating $A_1(\lambda_1, \lambda_2)$ by parts and using it in (2.9), it is easily seen that

$$\begin{aligned} B(\lambda_1, \lambda_2) = {} & \text{a term independent of } \lambda_1 \\ & + \int \{\psi(x, \lambda_1) f_2(x) - F_1(x)\psi_x(x, \lambda_2)\}\, dx . \end{aligned} \tag{2.10}$$

Hence,

$$\frac{\partial}{\partial \lambda_1} B(\lambda_1, \lambda_2) = \int \left\{\psi_{\lambda_1}(x, \lambda_1) f_2(x) - \frac{\partial}{\partial \lambda_1} F_1(x)\psi_x(x, \lambda_2)\right\} dx \tag{2.11}$$

and this is nonnegative if, for $\lambda_1 \leqq \lambda_2$,

$$\psi_{\lambda_1}(x, \lambda_1) f_{\lambda_2}(x) - \frac{\partial}{\partial \lambda_1} F_{\lambda_1}(x)\psi_x(x, \lambda_2) \geqq 0 \qquad \text{for all } x . \tag{2.12}$$

Now, we consider the configuration $\lambda_1 = \lambda_2 = \lambda$. It can be easily verified that

$$\frac{d}{d\lambda} B(\lambda, \lambda) = \sum_{i=1}^{2} \frac{\partial}{\partial \lambda_i} B(\lambda_1, \lambda_2)\Big|_{\lambda_1 = \lambda_2 = \lambda} \tag{2.13}$$

and

$$\frac{\partial}{\partial \lambda_2} B(\lambda_1, \lambda_2) = \frac{\partial}{\partial \lambda_2} B(\lambda_2, \lambda_1) = \frac{\partial}{\partial \lambda_1} B(\lambda_1, \lambda_2)\Big|_{\lambda_1 \leftrightarrow \lambda_2}, \tag{2.14}$$

where $\lambda_1 \leftrightarrow \lambda_2$ indicates that after differentiation λ_1 and λ_2 are interchanged in the final expression. Thus

$$\frac{d}{d\lambda} B(\lambda, \lambda) = 2 \frac{\partial}{\partial \lambda_1} B(\lambda_1, \lambda_2)\Big|_{\lambda_1 = \lambda_2 = \lambda}. \tag{2.15}$$

Hence $B(\lambda, \lambda)$ is non-decreasing if (2.12) holds. For the theorem, it is easy to see that it suffices if (2.12) holds when $\lambda_1 = \lambda_2 = \lambda$ in a manner consistent with (2.12); in other words, if (2.7) holds. The strict inequality part is now obvious.

REMARK 2.1. If $\psi(x, \lambda) = \psi(x)$ for all $\lambda \in \Lambda$, then (2.7) reduces to

$$\frac{\partial}{\partial \lambda} F_\lambda(x) \frac{d}{dx} \psi(x) \leqq 0 .$$

This is satisfied if $\{F_\lambda\}$ is an SI family of distributions and $\psi(x)$ is non-decreasing in x and hence $E_\lambda \psi(x)$ is non-decreasing in λ, which is a result of Lehmann ([6] page 112). A generalization of Lehmann's result has been stated earlier by Mahamunulu [7] and Alam and Rizvi [1] for the case of independent random variables with distribution functions F_{λ_i} $(i = 1, \cdots, k)$, where $\psi(x_1, \cdots, x_k)$ is non-decreasing in each argument.

REMARK 2.2. In the proof of Theorem 2.1 we have assumed that all the distributions F_λ have the same support. If (a_1, b_1) and (a_2, b_2) are the supports of F_{λ_1} and F_{λ_2}, where $a_i = a(\lambda_i)$ and $b_i = b(\lambda_i)$, $i = 1, 2$, then (2.10) is

$$B(\lambda_1, \lambda_2) = \psi(b_1, \lambda_2) + \int_{a_2}^{b_2} \psi(x, \lambda_1) f_2(x)\, dx - \int_{a_1}^{b_1} F_1(x) \psi_x(x, \lambda_2)\, dx . \tag{2.16}$$

Hence, we have

$$\begin{aligned} \frac{\partial}{\partial \lambda_1} B(\lambda_1, \lambda_2) &= \psi_x(b_1, \lambda_2) \frac{db_1}{d\lambda_1} + \int_{a_2}^{b_2} \psi_{\lambda_1}(x, \lambda_1) f_2(x)\, dx \\ &\quad - \int_{a_1}^{b_1} \frac{\partial}{\partial \lambda_1} F_1(x) \psi_x(x, \lambda_2)\, dx - \frac{db_1}{d\lambda_1} \psi_x(b_1, \lambda_2) . \\ &= \int_{a_2}^{b_2} \psi_{\lambda_1}(x, \lambda_1) f_2(x)\, dx - \int_{a_1}^{b_1} \frac{\partial}{\partial \lambda_1} F_{\lambda_1}(x) \psi_x(x, \lambda_2)\, dx . \end{aligned} \tag{2.17}$$

From this point we can proceed as before and show that the statement of the theorem is still true.

REMARK 2.3. If we assume that $\psi(x, \lambda) \geqq 0$, then, for any positive integer t, we can let $\phi(x, \lambda) = \psi^t(x, \lambda)$ play the role of $\psi(x, \lambda)$ of Theorem 2.1. It follows immediately that $E_\lambda \psi^t(x, \lambda)$ is non-decreasing in λ if (2.7) holds.

We now state the theorem which gives a sufficient condition for the monotonicity of $\psi(x; c, d, k)$.

THEOREM 2.2. *For the procedure R_h defined by* (2.1), $\psi(\lambda; c, d, k)$ *is non-decreasing in λ provided that for all $\lambda \in \Lambda$ and all x,*

$$f_\lambda(x)\frac{\partial}{\partial\lambda}F_\lambda(h(x)) - h'(x)f_\lambda(h(x))\frac{\partial}{\partial\lambda}F_\lambda(x) \geqq 0 , \tag{2.18}$$

where $h'(x) = dh(x)/dx$. Further $\psi(\lambda; c, d, k)$ is strictly increasing in λ if strict inequality holds in (2.18) *on a set of positive Lebesgue measure.*

PROOF. The proof is immediate by letting $\psi(x, \lambda) = F_\lambda(h(x))$ in Theorem 2.1 and using Remark 2.3.

Some remarks on the properties of R_h. Suppose, in the set-up of Theorem 2.1, we consider $\lambda_i \in \Lambda$, $i = 1, \cdots, k$ subject to the condition that $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_k$ and define

$$A_i \equiv A_i(\lambda_1, \cdots, \lambda_k) = \int \prod_{r=1, r\neq i}^{k} \psi(x, \lambda_r)\, dF_i(x) , \qquad i = 1, \cdots, k , \tag{2.19}$$

and

$$B(\lambda_1, \cdots, \lambda_k) = \sum_{i=1}^{k} A_i(\lambda_1, \cdots, \lambda_k) . \tag{2.20}$$

If we assume that $\psi(x, \lambda) \geqq 0$, we can integrate A_1 by parts, use it in (2.20) and show by differentiating w.r.t. λ_1 that $B(\lambda_1, \cdots, \lambda_k)$ is non-decreasing in λ_1 when $\lambda_2, \cdots, \lambda_k$ are kept fixed, provided that (2.12) holds. As a next step we can show that $B(\lambda_1, \cdots, \lambda_k)$ is non-decreasing in λ for the configuration $\lambda_1 = \cdots = \lambda_m = \lambda \leqq \lambda_{m+1} \leqq \cdots \leqq \lambda_k$, $1 \leqq m \leqq k$, when $\lambda_{m+1}, \cdots, \lambda_k$ are kept fixed, provided that (2.12) holds. This shows that, subject to the restriction $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_k$ the supremum of $B(\lambda_1, \cdots, \lambda_k)$ is attained when $\lambda_1 = \cdots = \lambda_k$.

Now the expected subset size, $E(S)$, is given by $E(S) = p_1 + p_2 + \cdots + p_k$, where p_r is the probability that $\pi_{(r)}$ is included in the selected subset using the procedure R_h. By letting $\psi(x, \lambda_{[i]}) = F_{[i]}(h(x))$, we see that $p_i = A_i$ and $E(S) = B(\lambda_{[1]}, \cdots, \lambda_{[k]})$. Hence, we find that the sup $E(S)$ is attained when the distributions are identical, if

$$\frac{\partial}{\partial\lambda_1}F_{\lambda_1}(h(x))f_{\lambda_2}(x) - h'(x)\frac{\partial}{\partial\lambda_1}F_{\lambda_1}(x)f_{\lambda_2}(h(x)) \geqq 0 \qquad \text{for} \quad \lambda_1 \leqq \lambda_2 \quad \text{and all} \quad x . \tag{2.21}$$

The above result has been proved by Gupta and Panchapakesan [4]. It is also to be noted that (2.21) with $h(x) = x + d$, $d \geqq 0$, is essentially the assumption made by Sobel [9] in order that the sup $E(S)$ be attained by using his rule when the distributions are identical.

It should be pointed out that condition (2.21) implies condition (2.18) of Theorem 2.2. However, there exist cases where (2.18) holds but (2.21) does not. This is, for example, the case when we consider the location parameter of the Cauchy distribution with $h(x) = x + d$. It should be made clear that the results given here in Theorem 2.2 and the inequality (2.21) connect, for the first time, the behavior of the probability of a correct selection and the expected subset size.

3. Some special cases. In this section, we apply the general results of Section 2 to selection problems concerning parameters of location, scale and convex mixtures of distributions. These three broad categories cover most of the selection procedures for various distributions discussed earlier.

(a) *Location parameter case.* In this case, $F_\lambda(x) = F(x-\lambda)$, $-\infty < \lambda < \infty$ and we see that (2.21) reduces to

$$h'(x)f_{\lambda_1}(x)f_{\lambda_2}(h(x)) - f_{\lambda_2}(x)f_{\lambda_1}(h(x)) \geqq 0 . \tag{3.1}$$

Since $h(x) \geqq x$, (3.1) is satisfied if $h'(x) \geqq 1$ and $f_\lambda(x)$ has a monotone likelihood ratio (MLR) in x. In problems discussed earlier, the usual choice has been $h(x) = x + d$, $d \geqq 0$. For this choice, it can be seen that $\psi(\lambda; c, d, k)$ is independent of λ so that the construction of the constants d is accomplished by solving $\psi(0; 1, d, k) = P^*$.

(b) *Scale parameter case.* Here we have $F_\lambda(x) = F(x/\lambda)$, $x \geqq 0$, $\lambda \geqq 0$ and (2.21) becomes

$$xh'(x)f_{\lambda_1}(x)f_{\lambda_2}(h(x)) - h(x)f_{\lambda_1}(h(x))f_{\lambda_2}(x) \geqq 0 . \tag{3.2}$$

If $xh'(x) \geqq h(x) \geqq 0$ and $f_\lambda(x)$ has MLR in x, then (3.2) is satisfied. $h(x) = cx$, $c \geqq 1$, has been the usual choice for the procedures studied under this case, in which case $\psi(\lambda; c, d, k)$ is independent of λ. The constants c are constructed to satisfy $\psi(1; c, 0, k) = P^*$.

(c) *Convex mixtures.* Now we apply the general results of (2.18) and (2.21) to the case where $f_\lambda(x)$ is a convex mixture of densities $g_j(x)$. This will include applications to noncentral χ^2, noncentral F and multiple correlation coefficient R^2 (both conditional and unconditional cases). These applications lead to stronger or more stringent conditions for the convex mixtures case in general; in particular for the case of R^2 (unconditional case) the result is proved that the sup $E(S)$ takes place when the distributions are identical.

$$f_\lambda(x) = \sum_{j=0}^\infty w(\lambda, j)g_j(x) \tag{3.3}$$

where $g_j(x)$, $j = 0, 1, \cdots$ is a sequence of density functions and $w(\lambda, j)$ are nonnegative weights such that $\sum_{j=0}^\infty w(\lambda, j) = 1$. We consider weights given by

$$w(\lambda, j) = \frac{a_j\lambda^j}{A(\lambda)j!} , \qquad A(\lambda) \geqq 0 , \quad \lambda \geqq 0 \tag{3.4}$$

and

$$a_{j+1} = (m + lj)a_j , \qquad j = 0, 1, \cdots; \quad l, m \geqq 0 . \tag{3.5}$$

Using (3.4) and successive applications of (3.5), we have

$$A(\lambda) = a_0(1 - \lambda l)^{-m/l} , \tag{3.6}$$

provided that $\lambda < 1/l$.

This special case is of interest. If we set $m = 1$, $l = 0$ and $a_0 = 1$, we get $w(\lambda, j) = e^{-\lambda}\lambda^j/j!$. Thus the densities $g_j(x)$ are weighted by Poisson weights. Familiar examples of $f_\lambda(x)$ in this case are the densities of noncentral chi-square and noncentral F variables with noncentrality parameter λ. Again, if we set $l = 1$ and $a_0 = 1$, we get densities $g_j(x)$ with negative binomial weights. The distribution of R^2, where R is the multiple correlation coefficient, in the so-called unconditional case is an example of the above. Subset selection procedures in these cases have been discussed by Gupta and Studden [5] and Gupta and Panchapakesan [3] with $h(x) = cx$, $c \geqq 1$. The above authors have obtained sufficient conditions for the monotonicity of $\psi(\lambda; c, d, k)$ in λ for those specific cases. We show below that the sufficient conditions obtained by the above authors are stronger than (2.18) and could be derived starting with the sufficient condition (2.18). In fact we will obtain corresponding sufficient conditions in these special cases starting with (2.21) under which the supremum of $E(S)$ is attained when the distributions are identical.

Now, by defining

$$r_\lambda(x) = A(\lambda)f_\lambda(x) \tag{3.7}$$

$$R_\lambda(x) = A(\lambda)F_\lambda(x) \tag{3.8}$$

and

$$Q_\lambda(x) = A(\lambda)\frac{\partial}{\partial\lambda}R_\lambda(x) - R_\lambda(x)\frac{\partial}{\partial\lambda}A(\lambda)\,, \tag{3.9}$$

it can be seen that (2.21) reduces to

$$Q_{\lambda_1}(h(x))r_{\lambda_2}(x) - h'(x)Q_{\lambda_1}(x)r_{\lambda_2}(h(x)) \geqq 0\,. \tag{3.10}$$

Using (3.6) and (3.8) in (3.9), we get after some simplifications

$$Q_\lambda(x) = a_0(1 - \lambda l)^{-1-m/l} \textstyle\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_{j+1}\Delta G_j(x)\,, \tag{3.11}$$

where $\Delta G_j(x) = G_{j+1}(x) - G_j(x)$.

Using (3.11) in (3.10) and cancelling out $a_0(1 - \lambda l)^{-1-m/l}$ we get

$$\begin{aligned}&\Big(\textstyle\sum_{j=0}^{\infty}\frac{\lambda_1^{\,j}}{j!}a_{j+1}\Delta G_j(h(x))\Big)\Big(\textstyle\sum_{j=0}^{\infty}\frac{\lambda_2^{\,j}}{j!}a_j g_j(x)\Big)\\ &\qquad - h'(x)\Big(\textstyle\sum_{j=0}^{\infty}\frac{\lambda_1^{\,j}}{j!}a_{j+1}\Delta G_j(x)\Big)\Big(\textstyle\sum_{j=0}^{\infty}\frac{\lambda_2^{\,j}}{j!}a_j g_j(h(x))\Big) \geqq 0\,.\end{aligned} \tag{3.12}$$

Since (3.12) should hold for $\lambda_1 \leqq \lambda_2$, we set $\lambda_2 = b\lambda_1 (b \geqq 1)$ in (3.12) and simplify it to obtain

$$\textstyle\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}\sum_{\alpha=0}^{i}\binom{i}{\alpha}a_\alpha a_{i-\alpha}T_\alpha(x) \geqq 0\,, \tag{3.13}$$

where

$$\begin{aligned}T_\alpha(x) = {}& b^{i-\alpha}(m + l\alpha)g_{i-\alpha}(x)\Delta G_\alpha(h(x))\\ &- h'(x)b^\alpha(m + l(i - \alpha))g_\alpha(h(x))\Delta G_{i-\alpha}(x)\,.\end{aligned} \tag{3.14}$$

Thus what we have shown at this stage is that for the case of $f_\lambda(x)$ given by (3.3) with weight functions specified by (3.4) and (3.5), the condition (2.21) is equivalent to (3.13).

We see that we can obtain more stringent sufficient condition by saying that the coefficient of λ^i for every i in (3.13) is nonnegative. Thus, (3.13) is implied by the condition that, for every integer $i \geqq 0$,

$$\sum_{\alpha=0}^{i} \binom{i}{\alpha} a_\alpha a_{i-\alpha} T_\alpha(x) \geqq 0 . \tag{3.15}$$

Now, grouping the terms corresponding to α and $i-\alpha$ in the summation of (3.15), we see that (3.15) in its turn is implied by the condition that, for $\alpha = 0, 1, \cdots, [i/2]$ ($[s]$ denotes the largest integer $\leqq s$), $T_\alpha(x) + T_{i-\alpha}(x) \geqq 0$, i.e.

$$\begin{aligned} &b^{i-\alpha}(m + l\alpha)[g_{i-\alpha}(x)\Delta G_\alpha(h(x)) - h'(x)g_{i-\alpha}(h(x))\Delta G_\alpha(x)] \\ &\quad + b^\alpha(m + l(i-\alpha))[g_\alpha(x)\Delta G_{i-\alpha}(h(x)) - h'(x)g_\alpha(h(x))\Delta G_{i-\alpha}(x)] \geqq 0 . \end{aligned} \tag{3.16}$$

For $b = 1$ (that is, $\lambda_1 = \lambda_2$), (3.16) gives the sufficient condition for $\psi(\lambda; c, d, k)$ to be non-decreasing, which will yield the conditions obtained by Gupta and Studden [5] and Gupta and Panchapakesan [3] by making the proper choices of the constants appropriate to the relevant weight functions.

We now illustrate the application of (3.16) by considering the procedure for selecting from multivariate normal distributions in terms of R^2 investigated by Gupta and Panchapakesan [3]. In this case $h(x) = cx$, $c \geqq 1$ and (3.16) becomes

$$\begin{aligned} &b^{i-\alpha}(m + q + \alpha)[g_{i-\alpha}(x)\Delta G_\alpha(cx) - cg_{i-\alpha}(cx)\Delta G_\alpha(x)] \\ &\quad + b^\alpha(m + q + i - \alpha)[g_\alpha(x)\Delta G_{i-\alpha}(cx) - cg_{i-\alpha}(cx)\Delta G_{i-\alpha}(x)] \geqq 0 , \end{aligned} \tag{3.17}$$

where

$$g_j(x) = \frac{\Gamma(q+m+j)}{\Gamma(q+j)\Gamma(m)} x^{q+j-1}(1-x)^{m-1} , \qquad 0 \leqq x \leqq 1 , \quad q > 0 , \quad m > 0 .$$

It can be seen by integration by parts that

$$\Delta G_j(x) = -\frac{\Gamma(q+m+j)}{\Gamma(q+j+1)\Gamma(m)} \frac{x^{q+j}}{(1+x)^{q+m+j}} .$$

Using this result in (3.17) and taking out the common factors (which are all positive), we see, after some simplifications, that (3.17) holds if

$$\left[\left(\frac{c+cx}{1+cx}\right)^{i-\alpha} - \left(\frac{c+cx}{1+cx}\right)^{\alpha}\right]\left[\frac{q+m+\alpha}{q+\alpha} b^{i-\alpha} - \frac{q+m+i-\alpha}{q+i-\alpha} b^\alpha\right] \geqq 0 . \tag{3.18}$$

Noting that $c \geqq 1$, $\alpha \leqq i-\alpha$ and $(q+m+\alpha)/(q+\alpha) \geqq (q+m+i-\alpha)/(q+i-\alpha)$ we see that (3.18) holds. This shows that sup $E(S)$ takes place when all the distributions are identical and it now follows that sup $E(S) = k$.

Concluding remarks. The problem of selecting a nonempty subset containing the population associated with $\lambda_{[1]}$ can be handled in an analogous manner. However, the properties to be satisfied by the function $h(x)$ will need some obvious modifications in order to have the desirable properties of the expression for the probability of a correct selection. Further, in both the cases of $\lambda_{[1]}$ and $\lambda_{[k]}$, there is the general question of determining a good form of $h(x)$ for a given stochastically increasing family. For the problem of selecting the population associated with $\lambda_{[k]}$, we can easily see that the sufficient condition for $\psi(\lambda; c, d, k)$ to be non-decreasing in λ in the special cases of location and scale parameters is given respectively by $h'(x) \geqq 1$ and $xh'(x) \geqq h(x)$, $x \geqq 0$.

Nagel [8] has discussed the construction of subset selection procedures satisfying certain optimality conditions and has in this context defined a just rule. In our set-up, let $x_1, \cdots, x_k$ and $y_1, \cdots, y_k$ be two sets of observations from the populations such that $x_i \leqq y_i$ and $x_j \geqq y_j$ for all $j \neq i$. Then a rule R is just if the probability of selecting π_i based on the observations $y_1, \cdots, y_k$ is at least as large as that of selecting π_i based on $x_1, \cdots, x_k$. It has been shown by Nagel that the procedure R_h is just if $h(x)$ is non-decreasing in x.

REFERENCES

[1] ALAM, K. and RIZVI, M. H. (1966). Selection from multivariate normal populations. *Ann. Inst. Statist. Math.* **18** 307–318.

[2] GUPTA, S. S. (1966). On some selection and ranking procedures for multivariate normal populations using distance functions, in *Multivariate Analysis*, ed. P. R. Krishnaiah. Academic Press, New York, 457–475.

[3] GUPTA, S. S. and PANCHAPAKESAN, S. (1969a). Some selection and ranking procedures for multivariate normal populations, in *Multivariate Analysis* **2**, ed. P. R. Krishnaiah. Academic Press, New York, 475–505.

[4] GUPTA, S. S. and PANCHAPAKESAN, S. (1969b). Selection and ranking procedures, in *The Design of Computer Simulation Experiments*, ed. T. H. Naylor. Duke Univ. Press, 132–160.

[5] GUPTA, S. S. and STUDDEN, W. J. (1970). On some selection and ranking procedures with applications to multivariate populations, in *Essays in Probability and Statistics*, ed. R. C. Bose, *et al.* Univ. of North Carolina Press.

[6] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

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

[8] NAGEL, K. (1970). On subset selection rules with certain optimality properties. Mimeograph Series No. 222, Department of Statistics, Purdue Univ.

[9] SOBEL, M. (1969). Selecting a subset containing at least one of the t best populations, in *Multivariate Analysis* **2**, ed. P. R. Krishnaiah. Academic Press. New York, 515–540.