225

ON A CLASS OF SUBSET SELECTION PROCEDURES*

by

Shanti S. Gupta and S. Panchapakeṣan

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 225

June 1970

---

*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.

On A Class of Subset Selection Procedures

Shanti S. Gupta and S. Panchapakesan

Purdue University

1. Introduction and Summary. Let $\pi_1, \pi_2, \ldots, \pi_k$ be k independent populations. Let Λ be an interval on the real line. Associated with π_i ($i = 1,2,\ldots,k$) is a real valued random variable X_i with an absolutely continuous distribution $F_i \equiv F_{\lambda_i}$, $\lambda_i \varepsilon \Lambda$, 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 $\lambda_{[1]} \leq \lambda_{[2]} \leq \ldots \leq \lambda_{[k]}$ 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 $\{F_\lambda\}$, $\lambda \varepsilon \Lambda$, is a stochastically increasing (SI) family of distributions, that is to say, for $\lambda < \lambda'$, F_λ and $F_{\lambda'}$ are distinct and $F_\lambda(x) \geq F_{\lambda'}(x)$ for all x. Let $x_1, x_2, \ldots, x_k$ be observations on $X_1, X_2, \ldots, X_k$, respectively. Based on these observations, the goal is to select a non-empty 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* ($\frac{1}{k} < P^* < 1$). If there are more than one populations 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

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

*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.

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

In the next section a class of procedures R_h is defined using a function $h \equiv h_{c,d}$, $c \geq 1$, $d \geq 0$, defined on the real line, for selection of the population associated with $\lambda_{[k]}$. This class of procedures is a natural generalization of a class of procedures proposed and studied by Gupta [2]. A main result of this section is Theorem 2.2 which generalizes an earlier result of Lehmann [6]. This theorem is used to evaluate the infimum of the probability of a correct selection. Section 3 discusses some properties of this class of procedures. The succeeding section briefly deals with the problem of selecting the population associated with $\lambda_{[1]}$.

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) $h_{c,d}(x) \geq x$, (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

$$h(x_i) \geq \max_{1\leq r\leq k} x_r . \tag{2.1}$$

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

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

where $f_{[r]}$ $(r = 1, 2, \ldots, k)$ 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,

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

Define

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

Then

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

Because of the set of conditions (A) imposed on h, we have

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

It is easy to see from the above that constants c and d can be chosen to satisfy the P^*-condition.

Sufficient Condition for the Monotonicity of $\psi(\lambda;\, c,d,k)$.

We state without proof as a preliminary result the following theorem which is essentially the result of Lehmann [6, p. 112].

Theorem 2.1. Let $\{F_\lambda\}$ be an SI family of distributions on the real line. Then $E_\lambda\psi(X)$ is non-decreasing in λ for any non-decreasing function ψ, where E_λ denotes the expectation w.r.t. F_λ.

A generalization of the above theorem has been stated by Mahamunulu [7] and Alam and Rizvi [1] for the case of independent and identically distributed random variables $X_1, X_2, \ldots, X_k$ with distribution function F_λ, where $\psi(x_1, x_2, \ldots, x_k)$ is non-decreasing in each argument. But what we presently seek is a generalization of Theorem 2.1 in a different direction stated in the following theorem.

Theorem 2.2. 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 the first partial derivatives ψ_x and ψ_λ w.r.t. x and λ respectively. Then $E_\lambda(X,\lambda)$ is non-decreasing in λ provided that

(2.7) $$\left| \frac{\partial(F,\psi)}{\partial(x,\lambda)} \right| \geq 0 ,$$

where

(2.8) $$\left| \frac{\partial(F,\psi)}{\partial(x,\lambda)} \right| = \begin{vmatrix} \frac{\partial}{\partial x} F_\lambda(x) & \psi_x(x,\lambda) \\ \frac{\partial}{\partial \lambda} F_\lambda(x) & \psi_\lambda(x,\lambda) \end{vmatrix} .$$

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

In order to prove this theorem we introduce some notations and establish two lemmas. Let

(2.9) $$A(\lambda) = \int \psi(x,\lambda)\, d\, F_\lambda(x) \equiv E_\lambda\, \psi(X,\lambda) .$$

Let us consider $\lambda_1, \lambda_2 \in \Lambda$ such that $\lambda_1 \leq \lambda_2$ and define

(2.10) $$A_i(\lambda_1,\lambda_2) = \int \prod_{\substack{r=1 \\ r\neq i}}^{2} \psi(x,\lambda_r)\, dF_i(x) , \quad i = 1,2$$

and

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

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

Lemma 2.1. $B(\lambda_1,\lambda_2)$ is non-decreasing in λ_1, when λ_2 is kept fixed, provided that, for $\lambda_1 \leq \lambda_2$,

$$\psi_{\lambda_1}(x,\lambda_1)\, f_{\lambda_2}(x) - \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(x)\, \psi_x(x,\lambda_2) \geq 0 \tag{2.12}$$

Proof. Integrating $A_1(\lambda_1,\lambda_2)$ by parts and using it in (2.11), it is easily seen that

$$B(\lambda_1,\lambda_2) = \text{a term independent of } \lambda_1 + \int \{\psi(x,\lambda_1)\, f_2(x) - F_1(x)\, \psi_x(x,\lambda_2)\}\, dx. \tag{2.13}$$

Hence

$$\frac{\partial}{\partial\lambda_1} B(\lambda_1,\lambda_2) = \int \{\psi_{\lambda_1}(x,\lambda_1)\, f_2(x) - \frac{\partial}{\partial\lambda_1} F_1(x)\, \psi_x(x,\lambda_2)\}\, dx \tag{2.14}$$

and this is non-negative if (2.12) holds.

Lemma 2.2. If $\lambda_1 = \lambda_2 = \lambda$, then $B(\lambda,\lambda)$ is non-decreasing in λ provided that (2.12) holds.

Proof. We note the following properties of $B(\lambda_1,\lambda_2)$ which can be easily verified.

$$\frac{d}{d\lambda} B(\lambda,\lambda) = \sum_{i=1}^{2} \frac{\partial}{\partial\lambda_i} B(\lambda_1,\lambda_2) \Bigg|_{\lambda_1 = \lambda_2 = \lambda} \tag{2.15}$$

$$\frac{\partial}{\partial\lambda_2} B(\lambda_1,\lambda_2) = \frac{\partial}{\partial\lambda_2} B(\lambda_2,\lambda_1) = \frac{\partial}{\partial\lambda_1} B(\lambda_1,\lambda_2) \Bigg|_{\lambda_1 \leftrightarrow \lambda_2}, \tag{2.16}$$

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

$$\frac{d}{d\lambda} B(\lambda,\lambda) = 2\, \frac{\partial}{\partial\lambda_1} B(\lambda_1,\lambda_2) \Bigg|_{\lambda_1 = \lambda_2 = \lambda} \tag{2.17}$$

Thus $\frac{d}{d\lambda} B(\lambda,\lambda) \geq 0$ if $\frac{\partial}{\partial\lambda_1} B(\lambda_1,\lambda_2) \geq 0$ for $\lambda_1 \leq \lambda_2$. Now appealing to Lemma 2.1, the proof is completed.

We are now equipped to prove Theorem 2.2.

Proof of Theorem 2.2. By Lemma 2.2, $B(\lambda,\lambda) = 2A(\lambda)$ is non-decreasing in λ if (2.12) holds. But for the theorem, it is easy to see that it suffices if (2.12) holds when $\lambda_1 = \lambda_2 = \lambda$. Thus the sufficient condition reduces to (2.7). The strict inequality part is now obvious.

Remark 2.1. In the proof of Lemma 2.1 we have assumed that all the distributions F_λ have the same support. But the lemma is true even if the support changes with λ. If (a_1, b_1) and (a_2, b_2) are the supports of F_{λ_1} and F_{λ_2}, it can be easily verified by integrating $A_1(\lambda_1,\lambda_2)$ by parts and differentiating w.r.t. λ_1 that

$$(2.18)\quad \frac{\partial}{\partial\lambda_1} A_1(\lambda_1,\lambda_2) = -\int_{a_1}^{b_1} \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(x)\ \psi_x(x,\lambda_2)\ dx\ .$$

Hence we have

$$(2.19)\quad \frac{\partial}{\partial\lambda_1} B(\lambda_1,\lambda_2) = \int_{a_2}^{b_2} \psi_{\lambda_1}(x,\lambda_1)\ f_{\lambda_2}(x)\ dx - \int_{a_1}^{b_1} \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(x)\ \psi_x(x,\lambda_2)\ dx\ ,$$

which is non-negative if (2.12) holds. It can also be seen that (2.15) and (2.16) are still true. Thus Theorem 2.2 is true when the supports are not the same.

Remark 2.2. If $\psi(x,\lambda) = \psi(x)$ for all $\lambda \in \Lambda$, then $E_\lambda\ \psi(x)$ is non-decreasing in λ if $\frac{\partial}{\partial\lambda} F_\lambda(x) \cdot \frac{d}{dx} \psi(x) \leq 0$. If we further assume that $\{F_\lambda\}$ is an SI family of distributions, then $E_\lambda\ \psi(x)$ is non-decreasing in λ if $\frac{d}{dx}\psi(x) \geq 0$, i.e. if $\psi(x)$ is non-decreasing in x, which is Lehmann's result.

Corollary 2.1. Let $\{F_\lambda\}$ and $\psi(x,\lambda)$ be as in the hypothesis of Theorem 2.2 with the additional condition that $\psi(x,\lambda) \geq 0$. Then, for any positive integer t, $E_\lambda\ \psi^t(x,\lambda)$ is non-decreasing in λ provided that (2.7) holds and is strictly increasing in λ if strict inequality holds in (2.7) on a set of positive Lebesgue measure.

Proof. Let $\phi(x,\lambda) = \psi^t(x,\lambda)$ play the role of $\psi(x,\lambda)$ in Theorem 2.2. Then $E_\lambda\ \psi^t(x,\lambda)$ is non-decreasing in λ if $\left| \frac{\partial(F,\phi)}{\partial(x,\lambda)} \right| \geq 0$. Since $\left| \frac{\partial(F,\phi)}{\partial(x,\lambda)} \right| =$ $t\ \psi^{t-1}(x,\lambda) \left| \frac{\partial(F,\psi)}{\partial(x,\lambda)} \right|$ and $\psi(x,\lambda)$ is non-negative the conclusion of the corollary is obvious.

Remark 2.3. Starting with integration of $A_2(\lambda_1,\lambda_2)$ by parts and then differentiating $B(\lambda_1,\lambda_2)$ partially w.r.t. λ_2, it can be shown that $A(\lambda)$ is non-increasing in λ if $\left| \frac{\partial(F,\psi)}{\partial(x,\lambda)} \right| \leq 0$.

Now we generalize the results of Lemmas 2.1 and 2.2 in order to use them subsequently when we discuss the expected subset size. Consider $\lambda_i \in \Lambda$, $i = 1,\ldots,k$ subject to the condition that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$. Define

$$A_i(\lambda_1,\ldots,\lambda_k) = \int \prod_{\substack{r=1 \\ r\neq i}}^{k} \psi(x,\lambda_i)\ d\ F_i(x),\ i = 1,\ldots,k\ , \tag{2.20}$$

$$B(\lambda_1,\ldots,\lambda_k) = \sum_{i=1}^{k} A_i(\lambda_1,\ldots,\lambda_k) \tag{2.21}$$

and

$$A(\lambda,k) = \int \psi^{k-1}(x,\lambda)\ dF_\lambda\ (x)\ . \tag{2.22}$$

Then $B(\lambda,\ldots,\lambda) = k\ A(\lambda,k)$. Integrating $A_1(\lambda_1,\ldots,\lambda_k)$ by parts and using it in (2.21) and then differentiating w.r.t. λ_1 we get

$$(2.23)\quad \frac{\partial}{\partial\lambda_1} B(\lambda_1,\dots,\lambda_k) =$$

$$\sum_{\alpha=2}^{k} \int \prod_{\substack{r=2\\ r\neq\alpha}}^{k} \psi(x,\lambda_r)\ \{\ \psi_{\lambda_1}(x,\lambda_1)\ f_\alpha(x) - \frac{\partial}{\partial\lambda_1} F_1(x)\ \psi_x(x,\lambda_\alpha)\ \}\ dx\ .$$

Thus we get the following lemma.

Lemma 2.3. $B(\lambda_1,\dots,\lambda_k)$ is non-decreasing in λ_1, when $\lambda_2,\dots,\lambda_k$ are kept fixed, provided that $\psi(x,\lambda) \geq 0$ and (2.12) holds.

If $\lambda_1 = \dots = \lambda_m = \lambda \leq \lambda_{m+1} \leq \dots \leq \lambda_k,\ 1 \leq m \leq k$,

then by a reasoning similar to that employed in the proof of Lemma 2.2, we can show that

$$(2.24)\quad \frac{\partial}{\partial\lambda} B(\lambda,\dots,\lambda,\ \lambda_{m+1},\dots,\lambda_k) = m\, \frac{\partial}{\partial\lambda_1} B(\lambda_1,\dots,\lambda_k) \Big|_{\lambda_1 = \dots = \lambda_m = \lambda}\ .$$

Thus we obtain the following lemma.

Lemma 2.4. If $\lambda_1 = \dots = \lambda_m = \lambda \leq \lambda_{m+1} \leq \dots \leq \lambda_k$, $1 \leq m \leq k$, then $B(\lambda,\dots,\lambda,\ \lambda_{m+1},\dots,\ \lambda_k)$ is non-decreasing in λ when $\lambda_{m+1},\dots,\lambda_k$ are kept fixed, provided that $\psi(x,\lambda)$ is non-negative and (2.12) holds.

As a consequence of Lemma 2.4, we obtain the following theorem.

Theorem 2.3. Let $B(\lambda_1,\dots,\lambda_k)$ be defined by (2.21) and (2.22). Then the supremum of $B(\lambda_1,\dots,\lambda_k)$ over $\lambda_1,\dots,\lambda_k \in \Lambda$ subject to the condition $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$ takes place at $\lambda_1 = \dots = \lambda_k$ provided that $\psi(x,\lambda)$ is non-negative and (2.12) holds.

We conclude this section by stating the following theorem which gives a sufficient condition for the monotonicity of $\psi(\lambda;\ c,d,k)$.

Theorem 2.4. For the procedure R_h defined by (2.1), $\psi(\lambda;\ c,d,k)$ is non-decreasing in λ provided that

(2.25) $$\frac{\partial}{\partial\lambda} F_\lambda(h(x))\, f_\lambda(x) - h'(x)\, f_\lambda(h(x)) \frac{\partial}{\partial\lambda} F_\lambda(x) \geq 0 ,$$

where $h'(x) = \frac{d}{dx} h(x)$ and $\psi(\lambda;\, c,d,k)$ is strictly increasing in λ if strict inequality holds in (2.25) on a set of positive Lebesgue measure.

Proof. The proof is immediate by letting $\psi(x,\lambda) = F_\lambda(h(x))$ in Theorem 2.2 and Corollary 2.1.

3. Properties of Procedure R_h. Let $\pi_{(r)}$ $(r = 1,\ldots,k)$ be the population associated with $\lambda_{[r]}$ and p_r be the probability that $\pi_{(r)}$ is included in the selected subset using the procedure R_h.

Theorem 3.1. The procedure R_h is monotone, i.e., for $1 \leq i < j \leq k$, $p_i \leq p_j$, provided that $h(x)$ is non-decreasing in x.

Proof. We need only show that $p_1 \leq p_2$. Let $\psi(x) = \prod_{r=3}^{k} F_{[r]}(x)$. Then

$p_1 = \int \psi(h(x))\, F_{[2]}(h(x))\, dF_{[1]}(x) \leq \int \psi(h(x))\, F_{[2]}(h(x))\, dF_{[2]}(x)$ because of Theorem 2.1 and the fact that $h(x)$ is non-decreasing in x. Since $F_{[2]}(h(x)) \leq F_{[1]}(h(x))$, $p_1 \leq \int \psi(h(x))\, F_{[1]}(H(x))\, dF_{[2]}(x) = p_2$.

Let S denote the size of the subset selected and S' be the number of non-best populations included in the subset. Then

(3.1) $$E(S) \equiv E(S|R_h) = p_1 + \ldots + p_k$$

and

(3.2) $$E(S') \equiv E(S'|R_h) = p_1 + \ldots + p_{k-1} .$$

Now by taking $\psi(x,\lambda_{[i]}) = F_{[i]}\,(h(x))$, we see that $p_i = A_i(\lambda_{[1]},\ldots, \lambda_{[k]})$ and $E(S|R_h) = B(\lambda_{[1]},\ldots, \lambda_{[k]})$. Hence, as the consequence of Theorem 2.3, we obtain the following result.

Theorem 3.2. For the procedure R_h, the supremum of $E(S|R_h)$ is attained at $\lambda_1 = \lambda_2 = \ldots = \lambda_k$ if, for $\lambda_1 \leq \lambda_2$, (3.3) $$\frac{\partial}{\partial\lambda_1} F_{\lambda_1}(h(x))\, f_{\lambda_2}(x) - h'(x)\, \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(x)\, f_{\lambda_2}(h(x)) \geq 0 .$$

Lemma 3.1. $E(S')$ is non-decreasing in $\lambda_{[1]}$, the other λ's kept fixed, provided that (3.3) holds.

Proof. It is easy to see that $p_{[k]}$ is non-increasing in $\lambda_{[1]}$. $E(S)$ is non-decreasing in $\lambda_{[1]}$ if (3.3) holds. Thus $E(S') = E(S) - p_{[k]}$ is non-decreasing in $\lambda_{[1]}$ if (3.3) holds.

From the proofs of Lemmas 2.3 and 2.4, it is clear that $E(S') \equiv \sum_{i=1}^{k-1} A_i(\lambda_{[1]},\ldots, \lambda_{[k]})$ is non-decreasing in λ, where $\lambda_{[1]} = \ldots = \lambda_{[m]} = \lambda \leq \lambda_{[m+1]} \leq \cdots \leq \lambda_{[k]}$, $1 \leq m \leq k-1$, provided that (3.3) is satisfied. This leads to the following theorem.

Theorem 3.3. For the procedure R_h, the supremum of $E(S')$ is attained at $\lambda_1 = \ldots = \lambda_k$ if, for $\lambda_1 \leq \lambda_2$, (3.3) is satisfied.

Remark 3.1. It is to be noted that the sufficient condition (2.25) is included in the condition (3.3). In many cases (3.3) can be verified to be true.

Some Special Cases.

(a) λ is a location parameter, i.e., $F_\lambda(x) = F(x-\lambda)$, $-\infty < \lambda < \infty$. In this case (3.3) reduces to

(3.4) $$h'(x)\, f_{\lambda_1}(x)\, f_{\lambda_2}(h(x)) - f_{\lambda_2}(x)\, f_{\lambda_1}(h(x)) \geq 0 .$$

Since $h(x) \geq x$, if $f_\lambda(x)$ has a monotone likelihood ratio (MLR) in x and $h'(x) \geq 1$, then (3.4) is satisfied. In particular, $h(x) = x + d$, $d \geq 0$, is the usual choice.

(b) λ is a scale parameter, i.e., $F_\lambda(x) = F(\frac{x}{\lambda})$, $x \geq 0$, $\lambda \geq 0$. In this case, (3.3) becomes

$$(3.5) \qquad x\, h'(x)\, f_{\lambda_1}(x)\, f_{\lambda_2}(h(x)) - h(x)\, f_{\lambda_1}(h(x))\, f_{\lambda_2}(x) \geq 0 .$$

If $f_\lambda(x)$ has MLR in x and $x\, h'(x) \geq h(x) \geq 0$, then (3.5) is satisfied. In particular, $h(x) = cx$, $c \geq 1$, has been the usual choice.

(c) $f_\lambda(x)$ is given by

$$(3.6) \qquad f_\lambda(x) = \sum_{j=0}^{\infty} w(\lambda,j)\, g_j(x) ,$$

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

$$(3.7) \qquad w(\lambda,j) = \frac{a_j \lambda^j}{A(\lambda)\; j!} , \quad A(\lambda) \geq 0, \qquad \lambda \geq 0$$

and

$$(3.8) \qquad a_{j+1} = (q + pj)a_j , \quad j = 0, 1, \ldots ; \quad p,q \geq 0 .$$

Using (3.7) and successive applications of (3.8), we have

$$(3.9) \qquad A(\lambda) = a_0\, (1-\lambda p)^{-q/p} ,$$

provided that $\lambda < \frac{1}{p}$. Let us define

$$(3.10) \qquad r_\lambda(x) = A(\lambda)\, f_\lambda(x)$$

and

$$(3.11) \qquad R_\lambda(x) = A(\lambda)\, F_\lambda(x) .$$

Then (3.3) reduces to

$$(3.12) \qquad Q_{\lambda_1}(h(x))\, r_{\lambda_2}(x) - h'(x)\, Q_{\lambda_1}(x)\, r_{\lambda_2}(h(x)) \geq 0 ,$$

where

$$(3.13) \qquad Q_\lambda(x) = A(\lambda) \frac{\partial}{\partial\lambda} R_\lambda(x) - R_\lambda(x) \frac{\partial}{\partial\lambda} A(\lambda) .$$

Using (3.9) and (3.11), we can write (3.13) as

$$Q_\lambda(x) = a_o(1-\lambda p)^{-1-q/p} \left[(1-\lambda p) \sum_{j=1}^{\infty} \frac{\lambda^{j-1}}{(j-1)!} a_j G_j(x) - q \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_j G_j(x) \right] . \tag{3.14}$$

The expression inside the square brackets of (3.14)

$$= \sum_{j=0}^{\infty} \frac{j}{j!} a_{j+1}\, G_{j+1}(x) - \sum_{j=1}^{\infty} (q+jp) \frac{\lambda^j}{j!} a_j G_j(x) - q\, a_o\, G_o\, (x)$$

$$= \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_{j+1}\, \Delta G_j(x) \ ,$$

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

Hence, (3.3) holds if

$$\Big(\sum_{j=0}^{\infty} \frac{\lambda_1^j}{j!} a_{j+1}\, \Delta G_j(h(x))\Big) \Big(\sum_{j=0}^{\infty} \frac{\lambda_2^j}{j!} a_j g_j(x)\Big) - h'(x) \Big(\sum_{j=0}^{\infty} \frac{\lambda_1^j}{j!} a_{j+1}\, \Delta G_j(x)\Big) \Big(\sum_{j=0}^{\infty} \frac{\lambda_2^j}{j!} a_j\, g_j(h(x))\Big) \geq 0 \ . \tag{3.15}$$

Since (3.15) should hold for $\lambda_1 \leq \lambda_2$, we set $\lambda_2 = b\ \lambda_1$, $b \geq 1$, and rewrite (3.15) in its equivalent form

$$\Big(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_{j+1}\, \Delta G_j(h(x))\Big) \Big(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_j b^j g_j(x)\Big) - h'(x) \Big(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_{j+1}\, \Delta G_j(x)\Big) \Big(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} a_j\, b^j g_j(h(x))\Big) \geq 0 \ . \tag{3.16}$$

Now this can be simplified and using (3.8) rewritten as

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \sum_{\alpha=0}^{i} \binom{i}{\alpha} a_\alpha\, a_{i-\alpha}\, T_\alpha(x) \geq 0 \ , \tag{3.17}$$

where

$$(3.18)\quad T_\alpha(x) = b^{i-\alpha}(q + p\alpha)\, g_{i-\alpha}(x)\, \Delta G_\alpha(h(x)) - h'(x)\, b^{\alpha}(q + p(i-\alpha))\, g_\alpha(h(x))\, \Delta G_{i-\alpha}(x).$$

Obviously, for (3.17) to hold, it is sufficient that for every integer $i \geq 0$,

$$(3.19)\quad \sum_{\alpha=0}^{i} \binom{i}{\alpha} a_\alpha\, a_{i-\alpha}\, T_\alpha(x) \geq 0 .$$

Grouping the terms corresponding to α and $i-\alpha$ in (3.19), a more stringent condition for (3.17) to hold is that, for $\alpha = 0, 1, \ldots, [\frac{i}{2}]$ ([s] denotes the largest integer $\leq s$), $T_\alpha(x) + T_{i-\alpha}(x) \geq 0$, i.e.

$$(3.20)\quad b^{i-\alpha}(q + p\alpha)\,[g_{i-\alpha}(x)\, \Delta G_\alpha(h(x)) - h'(x)\, g_{i-\alpha}(h(x))\, \Delta G_\alpha(x)] + b^{\alpha}(q + p(i-\alpha))\,[g_\alpha(x)\, \Delta G_{i-\alpha}(h(x)) - h'(x)\, g_\alpha(h(x))\, \Delta G_{i-\alpha}(x)] \geq 0 .$$

Remark 3.2. A sufficient condition for $\psi(\lambda;\, c,\, d,\, k)$ to be non-decreasing in λ in the three special cases (a) through (c) is given respectively by

$$(3.21)\quad h'(x) \geq 1$$

$$(3.22)\quad x\, h'(x) \geq h(x)$$

and

$$(3.23)\quad (q+p\alpha)\,[g_{i-\alpha}(x)\, \Delta G_\alpha(h(x)) - h'(x)\, g_{i-\alpha}(h(x))\, \Delta G_\alpha(x)] + (q + p(i-\alpha))\,[g_\alpha(x)\, \Delta G_{i-\alpha}(h(x)) - h'(x)\, g_\alpha(h(x))\, \Delta G_{i-\alpha}(x)] \geq 0.$$

Under the case (c), if we set $q = 1$, $p = 0$ and $a_o = 1$, we get $w(\lambda,j) = \frac{e^{-\lambda}\lambda^j}{j!}$. Thus the densities $g_j(x)$ are weighted by Poisson weights. Familiar examples of $f_\lambda(x)$ in this case are the densities of a non-central chi-square or non-central

F variable with non-centrality parameter λ. The sufficient condition (3.23) in this case with $h(x) = cx$, $c \geq 1$, has been obtained by Gupta and Studden [4]. Again, if we set $p = 1$ and $a_o = 1$, 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. The sufficient condition (3.23) for this case with $h(x) = cx$, $c \geq 1$, has been obtained by Gupta and Panchapakesan [3]. The sufficient condition in these special cases of Poisson and negative binomial weights have been obtained by the above authors by a direct method in absence of the general sufficient condition (2.25).

The case of binomial weights can be brought under (c). We have

$w(\lambda,j) = \binom{N}{j} \lambda^j (1-\lambda)^{N-j}$, $j = 0,\ldots, N$; $0 \leq \lambda \leq 1$. If we let $\mu = \lambda/(1-\lambda)$,

$w(\lambda,j) = \frac{\mu^j}{j!} \frac{a_j}{A(\mu)}$, where

$$(3.24) \quad a_j = \begin{cases} 1 & , j = 0 \\ N(N-1) \ldots (N-j+1), & j = 1,\ldots,N \\ 0 & , j = N+1,\ldots \end{cases}$$

and $A(\mu) = (1 + \mu)^N$. The density in this case becomes

$$(3.25) \qquad f_\mu(x) = \sum_{j=0}^{N} \frac{\mu^j}{j!} \frac{a_j}{A(\mu)} g_j(x) \quad .$$

We see that $$\frac{a_{j+1}}{a_j} \simeq \begin{cases} q + pj & , j = 0, 1,\ldots,N \\ 0 & , j = N+1,\ldots \end{cases}$$

where $q = N$ and $p = -1$. Though we assumed earlier that p and q are non-negative, all we need is that $q + jp$ be non-negative for all j in the finite mixture. Since

μ is an increasing function of λ, we can easily see that the sufficient condition (3.19) reduces to

$$(3.26)\quad \sum_{\alpha=\max(0,\ i+1-N)}^{\min(i,N)} \binom{i}{\alpha} M_\alpha(x) \geq 0 \text{ for } i = 0, 1, \ldots, 2N-1$$

where

$$(3.27)\quad M_\alpha(x) = a_\alpha\, a_{i-\alpha+1}\, [g_\alpha(x)\ \Delta G_{i-\alpha}(h(x)) - h'(x)\ g_\alpha(h(x))\ \Delta G_{i-\alpha}(x)]\ .$$

We conclude this section by stating a lemma regarding the MLR property of $f_\lambda(x)$ when it is given by (3.6).

Lemma 3.2. Let $f_\lambda(x)$ be a density given by (3.6). Then $f_\lambda(x)$ is totally positive of order 2 (TP_2), i.e., for $\lambda_1 < \lambda_2$ and $x_1 < x_2$, $\begin{vmatrix} f_{\lambda_1}(x_1) & f_{\lambda_1}(x_2) \\ f_{\lambda_2}(x_1) & f_{\lambda_2}(x_2) \end{vmatrix} \geq 0$ provided that $g_j(x)$ and $w(\lambda,j)$ are TP_2.

Proof. The proof is a consequence of the basic composition formula of Polyá and Szegö (see Karlin [5], p. 17), which in the present case is

$$\begin{vmatrix} f_{\lambda_1}(x_1) & f_{\lambda_1}(x_2) \\ f_{\lambda_2}(x_1) & f_{\lambda_2}(x_2) \end{vmatrix} = \sum_{j_1<j_2} \begin{vmatrix} g_{j_1}(x_1) & g_{j_1}(x_2) \\ g_{j_2}(x_1) & g_{j_2}(x_2) \end{vmatrix} \begin{vmatrix} w(\lambda_1,j_1) & w(\lambda_1,j_2) \\ w(\lambda_2,j_1) & w(\lambda_2,j_2) \end{vmatrix}$$

for $x_1 < x_2$ and $\lambda_1 < \lambda_2$.

4. Selection of the Population Associated with $\lambda_{[1]}$. In this case we will only briefly mention the modifications made and state the results without proofs. Let $H \equiv H_{c,d}$; $c\ \varepsilon[1,\infty)$, $d\ \varepsilon[0,\infty]$ be a real valued function defined on the real line

satisfying the following set of conditions (B): For every x belonging to the support of F , (i) $H_{c,d}(x) \leq x$, (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 $\lim_{c\to\infty} H_{c,d}(x) = 0$, d fixed.

Then the class of procedures R_H for selecting a subset containing the population associated with $\lambda_{[1]}$ is defined as follows.

R_H: Include the population Π_i in the selected subset iff

$$H(x_i) \leq \min_{1\leq r\leq k} x_r \ . \tag{4.1}$$

This procedure selects a non-empty subset because of the condition B-(i). The probability of a correct selection is given by

$$P(CS|R_H) = \int \prod_{r=2}^{k} \bar{F}_{[r]}\,(H(x))\, dF_{[1]}\,(x) \ , \tag{4.2}$$

where $\bar{F}_{[r]}\,(x) = 1 - F_{[r]}(x)$. Because of the stochastic ordering of F_λ, we have

$$\inf_{\Omega} P(CS|R_H) = \inf_{\lambda\varepsilon\Lambda} \phi(\lambda;\, c,d,k) \ , \tag{4.3}$$

where

$$\phi(\lambda;\, c,d,k) = \int \bar{F}_{\lambda}^{\,k-1}\,(H(x))\, dF_\lambda(x) \ . \tag{4.4}$$

Because of the conditions (B), we get

$$(4.5)\quad \begin{cases} \text{(i)}\ \ \phi(\lambda;\, c,d,k) \geq \frac{1}{k} \\ \text{(ii)}\ \ \phi(\lambda;\, 1,0,k) = \frac{1}{k} \\ \text{(iii)}\ \lim_{d\to\infty} \phi(\lambda;\, c,d,k) = 1 \text{ and/or} \\ \text{(iv)}\ \lim_{c\to\infty} \phi(\lambda;\, c,d,k) = [1 - F_\lambda(0)]^{k-1} \ . \end{cases}$$

If (iii) holds, then for every λ, c and k, we can choose d such that the P*-condition is satisfied. If (iv) alone holds, then for every λ, d and k, we can choose c in order to satisfy the P*-condition provided that $[1 - F_\lambda(0)]^{k-1} \geq P^*$ for all admissible λ and P*. Since P* can be chosen as close to 1 as we desire, we must have $F_\lambda(0) = 0$. Thus, if (iv) holds but not (iii), then to obtain values of the constants whatever P* be, the random variables must be non-negative. A sufficient condition for the monotonicity of $\phi(\lambda; c,d,k)$ is given in the following theorem.

Theorem 4.1. For the procedure R_H, $\phi(\lambda; c,d,k)$ is non-decreasing in λ provided that

$$H'(x)\, f_\lambda(H(x)) \frac{\partial}{\partial\lambda} F_\lambda(x) - f_\lambda(x) \frac{\partial}{\partial\lambda} F_\lambda(H(x)) \geq 0 \,, \tag{4.6}$$

where $H'(x) = \frac{d}{dx} H(x)$ and $\phi(\lambda; c,d,k)$ is strictly increasing in λ if strict inequality holds in (4.6) on a set of positive Lebesgue measure.

The procedure R_H is monotone, i.e., for $1 \leq i < j \leq k$, $p_i \geq p_j$. Further $E(S') = p_2 + \ldots + p_k = E(S) - p_1$. As in the case of R_h, we obtain the following results.

Theorem 4.2. For the procedure R_H, E(S) is non-decreasing in $\lambda_{[1]}$ when other λ's are kept fixed provided that, for $\lambda_1 \leq \lambda_2$,

$$H'(x)\, f_{\lambda_2}(H(x)) \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(x) - f_{\lambda_2}(x) \frac{\partial}{\partial\lambda_1} F_{\lambda_1}(H(x)) \geq 0 \,. \tag{4.7}$$

Theorem 4.3. E(S) attains its supremum over Ω at a point where $\lambda_1 = \lambda_2 = \ldots = \lambda_k$ provided that (4.7) holds.

It can also be shown that for $\lambda_{[1]} \leq \lambda_{[2]} = \ldots = \lambda_{[m]} = \lambda \leq \lambda_{[m+1]} \leq \ldots \leq \lambda_{[k]}$ for $2 \leq m \leq k$, E(S') is non-decreasing in λ provided that (4.7) holds. Because of the stochastic ordering, E(S') is non-decreasing in $\lambda_{[1]}$. Thus we have the following result concerning E(S').

Theorem 4.4. For the procedure R_H, sup $E(S')$ over Ω takes place at a point where $\lambda_1 = \dots = \lambda_k$ provided that (4.7) is satisfied.

Remarks 4.1. If λ is a location parameter, then (4.7) is satisfied if $f_\lambda(x)$ has MLR in x and $H'(x) \le 1$. In the case of a scale parameter λ, (4.7) is satisfied if $f_\lambda(x)$ has MLR in x and $H(x) \ge x\, H'(x) \ge 0$. In the case of $f_\lambda(x)$ given by (3.6) and (3.7), the condition (4.7) is satisfied if, for $\alpha = 0, 1, \dots, [\frac{i}{2}]$,

$$(4.8) \qquad b^{i-\alpha}(q + p\alpha)\,[H'(x) g_{i-\alpha}(H(x))\, \Delta G_\alpha(x) - g_{i-\alpha}(x)\, \Delta G_\alpha(H(x))]$$

$$+\; b^{\alpha}(q + p(i-\alpha))\,[H'(x) g_\alpha(H(x))\, \Delta G_{i-\alpha}(x) - g_\alpha(x)\, \Delta G_{i-\alpha}(H(x))] \ge 0 \;.$$

Remark 4.2. The condition (4.7) in the case of R_H corresponds to (3.3) in the case of R_h. Suppose we use R_h with $h(x) = x + d$ or $h(x) = cx$ (in the case of non-negative random variables) and use R_H correspondingly with $H(x) = x-d$ or $H(x) = \frac{x}{c}$. Then it is easy to see by change of variables that the conditions (3.3) and (4.7) are the same. Consequently, the sufficient conditions for $\psi(\lambda;\, c,d,k)$ and $\phi(\lambda;\, c,d,k)$ to be non-decreasing in λ are the same.

5. Concluding Remarks. Most subset selection procedures discussed in the literature fall under the class of procedures R_h and R_H. The general results of the preceding sections are directly applicable to specific procedures discussed by Gupta [2], and Gupta and Panchapakesan [3]. For details on these applications reference can be made to Panchapakesan [9].

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 $x_1,\dots,x_k$ and $y_1,\dots,y_k$ are two sets of observations from the populations such that $x_i \le y_i$ and $x_j \ge y_j$ for all $j \ne i$. Then a rule R is just if the probability of selecting π_i based on the observations $y_1,\ldots,y_k$ is at least as large as that of selecting π_i based on $x_1,\ldots,x_k$. It has been shown by Nagel that the procedure $R_h(R_H)$ is just if $h(x)$ $(H(x))$ 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. Multivariate Analysis (ed. P.R. Krishnaiah), Academic Press, N.Y., 457-475.

[3] Gupta, S.S. and Panchapakesan, S. (1969). Some selection and ranking procedures for multivariate normal populations. Multivariate Analysis II (ed. P.R. Krishnaiah), Academic Press, N.Y., 475-505.

[4] Gupta, S.S. and Studden, W.J. (1970). On some selection and ranking procedures with applications to multivariate populations. Essays in Probability and Statistics (ed. R.C. Bose et al), University of North Carolina Press, Chapel Hill, N. Carolina.

[5] Karlin, S. (1968). Total Positivity Vol. I. Stanford University Press, Stanford, California.

[6] Lehmann, E.L. (1959). Testing Statistical Hypotheses, John Wiley, N.Y.

[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 University.

[9] Panchapakesan, S. (1969). Some contributions to multiple decisions (selection and ranking) procedures. Mimeograph Series No. 192, Department of Statistics, Purdue University.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Purdue University	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

On A Class Of Subset Selection Procedures

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report, June 1970

5. AUTHOR(S) *(Last name, first name, initial)*

Gupta, Shanti S. and Panchapakesan, S.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1970	21	9

8a. CONTRACT OR GRANT NO. NONR 1100 (26) and AF 33(615)67C1244 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Mimeo Series # 225 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Also supported by the Aerospace Research Laboratories, Wright Patterson AFB, Ohio	Office of Naval Research Washington, D.C.

13. ABSTRACT

Let π_i, $i = 1,\ldots,k$, be a continuous population with associated d.f. F_{λ_i}, $\lambda_i \in \Lambda$, an interval on the real line. The family $\{F_\lambda\}$ is assumed to be stochastically increasing in λ. Section 2 defines a class of procedures R_h for selecting a non-empty subset of the k populations such that the probability of a correct selection (PCS), i.e. selection of a subset which includes the population with the largest λ_i, is at least P*, a pre-assigned level. A generalization (Theorem 2.2) of a result of Lehmann is used to obtain a sufficient condition for the monotonicity of a certain integral leading to the evaluation of the infimum of PCS over the parameter space. Results concerning the supremum of the expected subset size and other properties of R_h form the content of the next section. Section 3 also contains more specific results when the density $f_\lambda(x)$ is a convex mixture of a sequence of known density functions. The problem of selection of the population associated with the smallest λ_i is briefly discussed in Section 4.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification