An Essentially Complete Class of Two-Stage Selection Procedures with Screening at the First Stage

Shanti S. Gupta[1]
Purdue University

Klaus-J. Miescke[2]
Mainz University and Purdue University

Mimeograph Series #81-53

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #81-53

November 1981
(Revised June 1982)

---

[1] Research supported by ONR contract N00014-75-C-0455 at Purdue University, West Lafayette, IN U.S.A.

[2] Research supported by NSC contract VE 81015 at Academia Sinica, Taipei, Republic of China.

<u>AMS 1970 subject classification</u>: <u>Primary</u> 62F07, <u>secondary</u> 62F05, 62F15, 62L99.

An Essentially Complete Class of Two-Stage Selection Procedures With Screening at the First Stage

by Shanti S. Gupta [1]
Purdue University

and Klaus-J. Miescke [2]
Mainz University and Purdue University

[1] Research supported by ONR contract N00014-75-C-0455 at Purdue University, West-Lafayette, U. S. A.

[2] Research supported by NSC contract VE 81015 at Academia Sinica, Taipei, Republic of China.

AMS 1970 subject classification: Primary 62F07, secondary 62F05, 62F15, 62L99.

Key words and Phrases: Multiple decision procedures, two-stage selection procedures, screening procedures, sequential procedures, Bayesian procedures, essentially complete class.

SUMMARY

Let $\pi_1, \ldots, \pi_k$ be given populations associated with unknown real parameters $\theta_1, \ldots, \theta_k$. The goal is to find a population with a sufficiently large parameter in two stages with screening out inferior populations at the first stage. Both the control and the non-control situations are considered simultaneously. Let $\mathcal{D}_I$ denote the class of permutation invariant randomized procedures (ϕ, ψ, δ), where at Stage 1, ϕ and ψ decide how many populations and then which ones, respectively, are selected, and where at Stage 2, after additional samples have been drawn from the selected populations, δ makes the final decision. Let ψ^* and δ^* denote the natural decisions, i.e. which are associated with the largest sufficient statistics. Under the assumption of a common discrete or continuous type strongly unimodal exponential family, it is shown that with respect to every reasonable loss function, procedures of the type (ϕ, ψ^*, δ^*) form an essentially complete class within $\mathcal{D}_I$.

1. Introduction.

Let $\pi_1, \ldots, \pi_k$ be k given populations associated with unknown parameters $\theta_1, \ldots, \theta_k \in \Omega$, where $\Omega \subseteq \mathbb{R}$ is an unbounded or bounded interval. Let the goal be to find a population with a sufficiently large parameter after having screened out inferior populations at the first stage. If all (all but one) populations are screened out at Stage 1, the procedure stops and decides finally in favor of none (this one). Otherwise, at Stage 2 additional observations are taken from all populations which have been selected at Stage 1, and a final decision is made among them. The result of this paper is applicable to both the control (where a control value $\theta_0 \in \Omega$ is given) and the non-control cases. For references of papers dealing with two-stage procedures of the type described above see Gupta and Miescke [1981].

Assume that samples $\{X_{ij}\}_{j=1,\ldots,n}$ and $\{Y_{ij}\}_{j=1,\ldots,m}$ can be drawn from π_i at Stage 1 and Stage 2, respectively, $i = 1, \ldots, k$, which are mutually independent. Let the observations from π_i be real-valued and have a density $c(\theta_i)\exp(\theta_i x)b(x)$, $x \in \mathbb{R}$, $\theta_i \in \Omega$, w.r.t. μ, the Lebesgue measure on $\mathbb{R}$ or the counting measure on $\mathbb{Z}$, resp., $i = 1, \ldots, k$. The function $b(x)$, $x \in \mathbb{R}$, and thus the underlying exponential family is assumed to be common for all k populations. Let $U_i = X_{i1} + \ldots + X_{in}$ and $V_i = Y_{i1} + \ldots + Y_{im}$ be the sufficient statistics for θ_i with respect to the samples of π_i at the two stages, and let their densities with respect to μ be denoted by f_{θ_i} and g_{θ_i}, respectively, $i = 1, \ldots, k$. Finally, let $W_i = U_i + V_i$ be the overall sufficient statistic for θ_i, $i = 1, \ldots, k$. For notational convenience, let $\underline{U} = (U_1, \ldots, U_k)$, $\underline{V} = (V_1, \ldots, V_k)$ etc. in the following. For later considerations, note that

$$(1) \qquad f_\theta(u) = c_n(\theta)\exp(\theta u)b_n(u), \quad u \in \mathbb{R}, \quad \theta \in \Omega ,$$

$$g_\theta(v) = c_m(\theta)\exp(\theta v)b_m(v), \quad v \in \mathbb{R}, \quad \theta \in \Omega ,$$

where $c_n(\theta) = c(\theta)^n$, $c_m(\theta) = c(\theta)^m$, $\theta \in \Omega$, and where b_n and b_m denote the n-fold and m-fold convolution of b with respect to μ .

We now give a precise definition of a randomized two-stage procedure.

<u>Definition 1.</u> <u>Two-stage procedure (ϕ, ψ, δ).</u>

<u>Stage 1</u>: After having observed $\underline{U} = \underline{u}$, two decisions have to be made: how many populations should be selected and then, which ones. Let ϕ and ψ be the corresponding decision functions. Thus let $\phi = \{\phi_i \mid i = 0, 1, \ldots, k\}$, where $\phi_i : \mathbb{R}^k \to [0, 1]$ is measurable, $i = 0, 1, \ldots, k$, and $\sum_{i=0}^{n} \phi_i \equiv 1$. Moreover, let $\psi = \{\psi_{s,t} \mid s \subseteq \{1, \ldots, k\}, |s| = t, t \geq 1\}$, where $\psi_{s,t}: \mathbb{R}^k \to [0, 1]$ is measurable, $s \subseteq \{1, \ldots, k\}$, $|s| = t$, $t \geq 1$, and $\sum_{s,|s|=t} \psi_{s,t} \equiv 1$, $t \geq 1$, and where $|s'|$ denotes the size of a subset s' of $\{1, \ldots, k\}$.

If a decision based on $\phi(\underline{u})$ states that $|s| = 0$ then the procedure stops, and no population is finally selected. If a decision based on $\phi(\underline{u})$ states the $|s| = 1$ then the procedure stops also, and a final decision is made based on $\{\psi_{\{i\},1}(\underline{u}) \mid i = 1, \ldots, k\}$. In all other cases the procedure proceeds to Stage 2.

Stage 2: If at Stage 1, populations π_j with $j \in \tilde{s} = \{i_1, \ldots, i_t\}$, $i_1 < \ldots < i_t$, $t \geq 2$, have been selected, then, after having observed $V_{i_1} = v_{i_1}, \ldots, V_{i_t} = v_{i_t}$, the final decision will be made based on $\{\delta_{j,\tilde{s}}(\underline{u}, v_{i_1}, \ldots, v_{i_t}) | j \in \tilde{s}\}$. Here, $\delta = \{\delta_{i,s} | i \in s, s \subseteq \{1, \ldots, k\}, |s| \geq 2\}$, where $\delta_{i,s} : \mathbb{R}^k \times \mathbb{R}^{|s|} \to [0, 1]$ is measurable, $i \in s$, $s \subseteq \{1, \ldots, k\}$, $|s| \geq 2$, and $\sum_{i \in s} \delta_{i,s} \equiv 1$, $|s| \geq 2$.

Let $\mathscr{D}$ denote the class of all such two-stage procedures.

Definition 2. A procedure $(\phi, \psi, \delta) \in \mathscr{D}$ is called (permutation) *invariant*, if the following three conditions are fulfilled:

ϕ is invariant: For every $i \in \{1, \ldots, k\}$, $\underline{u} \in \mathbb{R}^k$, and permutation σ of $(1, \ldots, k)$, $\phi_0(\underline{u}) = \phi_0(u_{\sigma(1)}, \ldots, u_{\sigma(k)})$ and $\phi_i(\underline{u}) = \phi_i(u_{\sigma(1)}, \ldots, u_{\sigma(k)})$.

ψ is invariant: For every $t \in \{1, \ldots, k\}$, $s \subseteq \{1, \ldots, k\}$ with $|s| = t$, $\underline{u} \in \mathbb{R}^k$, and permutation σ of $(1, \ldots, k)$, $\psi_{\sigma(s),t}(\underline{u}) = \psi_{s,t}(u_{\sigma(1)}, \ldots, u_{\sigma(k)})$, where $\sigma(s) = \{\sigma(j) | j \in s\}$.

δ is invariant: For every $s = \{i_1, \ldots, i_t\}$ with $i_1 < \ldots < i_t$, $t \geq 2$, every $\sigma(s) = \{j_1, \ldots, j_t\}$ with $j_1 < \ldots < j_t$, σ permutation of $(1, \ldots, k)$, and every $(v_{j_1}, \ldots, v_{j_t}) \in \mathbb{R}^t$, $\delta_{\sigma(i),\sigma(s)}(\underline{u}, v_{j_1}, \ldots, v_{j_t}) = \delta_{i,s}(u_{\sigma(1)}, \ldots, u_{\sigma(k)}, v_{\sigma(i_1)}, \ldots, v_{\sigma(i_t)})$.

Let $\mathscr{D}_I$ denote the class of all invariant $(\phi, \psi, \delta) \in \mathscr{D}$.

Loss Assumptions: Let $L(\underline{\theta}, s, i)$ denote the loss at $\underline{\theta} \in \Omega^k$ if subset

$s \subseteq \{1, \ldots, k\}$ with $|s| \geq 1$ is selected at Stage 1 and the final decision is made in favor of $i \in s$. Let this loss function be invariant, i.e. $L(\underline{\theta}, \sigma(s), \sigma(i)) = L((\theta_{\sigma(1)}, \ldots, \theta_{\sigma(k)}), s, i)$, $i \in s \subseteq \{1, \ldots, k\}$, $\underline{\theta} \in \Omega^k$, for every permutation σ of $(1, \ldots, k)$. Assume that at every fixed $\underline{\theta} \in \Omega^k$ with $\theta_1 \leq \theta_2$, and $\tilde{s} \subseteq \{3, \ldots k\}$ with $0 \leq |\tilde{s}| \leq k-2$, the following four conditions are satisfied:

(L1) $L(\underline{\theta}, \{2\}, 2) \leq L(\underline{\theta}, \{1\}, 1)$

(L2) $L(\underline{\theta}, \tilde{s} \cup \{1,2\}, 2) \leq L(\underline{\theta}, \tilde{s} \cup \{1,2\}, 1)$

(L3) $L(\underline{\theta}, \tilde{s} \cup \{2\}, i) \leq L(\underline{\theta}, \tilde{s} \cup \{1\}, i)$, $i \in \tilde{s}$,

(L4) $L(\underline{\theta}, \tilde{s} \cup \{2\}, 2) \leq L(\underline{\theta}, \tilde{s} \cup \{1\}, 1)$, $|\tilde{s}| \geq 1$.

Note that we make no assumptions about the loss function $L_0(\underline{\theta})$, say, for selecting no population at Stage 1 at $\underline{\theta} \in \Omega^k$. All reasonable loss functions in the control as well as in the non-control setting should have the properties assumed above.

Example: The following loss function has been adopted by Gupta and Miescke (1982) in the control case. $L_0(\underline{\theta}) = 0$, $L(\underline{\theta}, \{j\}, j) = \ell(\theta_0 - \theta_j)$, $j = 1, \ldots, k$, $L(\underline{\theta}, s, i) = c|s| + \ell(\theta_0 - \theta_i)$, $i \in s$, $s \subseteq \{1, \ldots, k\}$ with $|s| \geq 2$, where $\theta_0 \in \Omega$ is a given control value, ℓ is non-decreasing with $\ell(0) = 0$, and $c \geq 0$ is the cost for every population which enters Stage 2.

In this paper our purpose is to show that every $(\phi, \psi, \delta) \in \mathscr{D}_I$ is dominated, uniformly in terms of risk, by (ϕ, ψ, δ^*), and, moreover, under the additional assumption of a strongly unimodal discrete or continuous type exponential family, by (ϕ, ψ^*, δ^*) , where ψ^* and δ^* are the natural decision functions which are defined below.

Definition 3. For every $t \in \{1, \dots, k\}$, $s \subseteq \{1, \dots, k\}$ with $|s| = t$, $\underline{u} \in \mathbb{R}^k$, $\psi^*_{s,t}(\underline{u}) = |B_t(\underline{u})|^{-1}(0)$ if $s \in (\notin)\ B_t(\underline{u})$, where $B_t(\underline{u}) = \{s' | s' \subseteq \{1, \dots, k\}, |s'| = t, \max\{u_i | i \notin s'\} \leq \min\{u_j | j \in s'\}\}$. Similarly, for every $s = \{i_1, \dots, i_t\}$ with $t \geq 2$, $i \in s$, $\underline{u} \in \mathbb{R}^k$, $(v_{i_1}, \dots, v_{i_t}) \in \mathbb{R}^t$ and $w_{i_j} = u_{i_j} + v_{i_j}$, $j = 1, \dots, t$, $\delta^*_{i,s}(\underline{u}, v_{i_1}, \dots, v_{i_t}) = |C_s(w_{i_1}, \dots, w_{i_t})|^{-1}(0)$ if $i \in (\notin) C_s(w_{i_1}, \dots, w_{i_t})$, where $C_s(w_{i_1}, \dots, w_{i_t}) = \{j | w_j = \max\{w_r | r \in s\}, j \in s\}$. Note that $\delta^*_{i,s}$ is only a function of $(w_{i_1}, \dots, w_{i_t})$.

The results by Eaton (1967) will play a fundamental role in our considerations and will be used repeatedly. Instead of repeating at every new occasion let us point out now that in all relevant situations Eaton's "property M" is given. The argument is always the same and can be found e.g. at the end of Section 2 in Eaton's paper. Also, the specific loss functions under concern will always be invariant in his sense, but we do not assume that they are non-negative.

2. The results.

The risk of a procedure $(\phi, \psi, \delta) \in \mathscr{D}$ at $\underline{\theta} \in \Omega^k$ is given by

$$(2) \qquad R(\underline{\theta}, (\phi,\psi,\delta)) = L_0(\underline{\theta})E_{\underline{\theta}}\phi_0(\underline{U}) + \sum_{i=1}^{k} L(\underline{\theta},\{i\},i)E_{\underline{\theta}}[\phi_1(\underline{U})\psi_{\{i\},1}(\underline{U})]$$

$$+ \sum_{t=2}^{k} E_{\underline{\theta}}[\phi_t(\underline{U}) \sum_{s=\{i_1,\dots,i_t\}} \psi_{s,t}(\underline{U}) \sum_{j \in s} L(\underline{\theta},s,j)\delta_{j,s}(\underline{U},V_{i_1},\dots,V_{i_t})] .$$

Since for every $(\phi, \psi, \delta) \in \mathscr{D}_I$, $R(\underline{\theta}, (\phi, \psi, \delta)) = R((\theta_{\sigma(1)}, \ldots, \theta_{\sigma(k)}), (\phi, \psi, \delta))$, $\underline{\theta} \in \Omega^k$, for every permutation σ of $(1, \ldots, k)$, it is convenient to compare procedures in $\mathscr{D}_I$ in terms of their Bayes risks with respect to permutation symmetric priors. Thus let $\underline{\Theta} = (\Theta_1, \ldots, \Theta_k)$ be from now on the random parameter vector with any but fixed permutation symmetric (prior) probability distribution τ defined on the Borel sets in Ω^k. It has to be assumed now that $L(\underline{\theta}, s, i)$ for every fixed $i \in s$, $s \subseteq \{1, \ldots, k\}$, is measurable and integrable properly. Note that for most of our results we need to consider only priors τ which have finite supports, where these two conditions mentioned above are met automatically. The Bayes risk for $(\phi, \psi, \delta) \in \mathscr{D}$ under τ is given by

$$(3) \qquad r(\tau, (\phi, \psi, \delta)) = \int_{\Omega^k} R(\underline{\theta}, (\phi, \psi, \delta)) d\tau(\underline{\theta}) = E\, R(\underline{\Theta}, (\phi, \psi, \delta)) .$$

Remark 1. Note that for two procedures in $\mathscr{D}_I$ we have $R(\underline{\theta}, (\phi_1, \psi_1, \delta_1)) \leq R(\underline{\theta}, (\phi_2, \psi_2, \delta_2))$, $\underline{\theta} \in \Omega^k$, if and only if $r(\tau', (\phi_1, \psi_1, \delta_1)) \leq r(\tau', (\phi_2, \psi_2, \delta_2))$ for every symmetric prior τ' with support $(\tau') = \{(\theta_{\sigma(1)}, \ldots, \theta_{\sigma(k)}) \mid \sigma$ permutation of $(1, \ldots, k)\}$, $\underline{\theta} \in \Omega^k$.

Since we will compare procedures in $\mathscr{D}_I$ which are only different in the ψ-and δ-components, the natural way to do this is to look at the conditional posterior risks, given $\underline{U} = \underline{u}$, for every fixed $\underline{u} \in \mathbb{R}^k$. Thus let $\underline{u} \in \mathbb{R}^k$ be fixed, which in view of the invariance of the problem can be assumed to satisfy $u_1 \leq u_2 \leq \ldots \leq u_k$. Now, for $(\phi, \psi, \delta) \in \mathscr{D}_I$, this conditional risk, given $\underline{U} = \underline{u}$, is given by

$$(4)\qquad E\{L(\underline{\Theta}, (\phi,\psi,\delta)) | \underline{U} = \underline{u}\} = \phi_0(\underline{u}) E\{L_0(\underline{\Theta}) | \underline{U} = \underline{u}\}$$

$$+ \phi_1(\underline{u}) \sum_{i=1}^{k} \psi_{\{i\},1}(\underline{u}) E\{L(\underline{\Theta}, \{i\}, i) | \underline{U} = \underline{u}\} + \sum_{t=2}^{k} \phi_t(\underline{u}) \sum_{s=\{i_1,\ldots,i_t\}} \psi_{s,t}(\underline{u})$$

$$\cdot E\{\sum_{j\in s} E\{\delta_{j,s}(\underline{u}, V_{i_1}, \ldots, V_{i_t}) L(\underline{\Theta}, s, j) | \underline{U} = \underline{u}, V_{i_1}, \ldots, V_{i_t}\} | \underline{U} = \underline{u}\},$$

which, after some standard computations, is seen to be equal to

$$(5)\qquad \phi_0(\underline{u}) \int_{\Omega^k} L_0(\underline{\theta}) \prod_{q=1}^{k} f_{\theta_q}(u_q) d\tau(\underline{\theta}) \beta(\underline{u})^{-1}$$

$$+ \phi_1(\underline{u}) \sum_{i=1}^{k} \psi_{\{i\},1}(\underline{u}) \int_{\Omega^k} L(\underline{\theta}, \{i\}, i) \prod_{q=1}^{k} f_{\theta_q}(u_q) d\tau(\underline{\theta}) \beta(\underline{u})^{-1}$$

$$+ \sum_{t=2}^{k} \phi_t(\underline{u}) \sum_{s=\{i_1,\ldots,i_t\}} \psi_{s,t}(\underline{u}) \int_{\mathbb{R}^t} \sum_{j\in s} \delta_{j,s}(\underline{u}, v_{i_1}, \ldots, v_{i_t})$$

$$\int_{\Omega^k} L(\underline{\theta}, s, j) \prod_{q=1}^{k} f_{\theta_q}(u_q) \prod_{r\in s} g_{\theta_r}(v_r) d\tau(\underline{\theta}) \prod_{p\in s} d\mu(v_p) \beta(\underline{u})^{-1} ,$$

where $\beta(\underline{u}) = \int_{\Omega^k} \prod_{q=1}^{k} f_{\theta_q}(u_q) d\tau(\underline{\theta})$.

<u>Lemma 1</u>. *Let* $(\phi, \psi, \delta) \in \mathscr{D}_I$ *be fixed. Then* $r(\tau, (\phi, \tilde{\psi}, \delta)) \leq r(\tau, (\phi, \psi, \delta))$ *and thus* $R(\underline{\theta}, (\phi, \tilde{\psi}, \delta)) \leq R(\underline{\theta}, (\phi, \psi, \delta))$, *uniformly in* $\underline{\theta} \in \Omega^k$ *, where* $\tilde{\psi}$ *is the same as* ψ *except that* $\tilde{\psi}_{\{i\},1} = \psi^*_{\{i\},1}$, $i = 1, \ldots, k$.

<u>Proof:</u> The first assertion follows from (L1) and Lemma 4.1 of Eaton (1967), and the second one holds true in view of Remark 1.

From now on we have to study the term following $\psi_{s,t}(\underline{u})$ in (5) in more detail. Before we make use of the assumption that we are dealing with an exponential family, it is crucial for our further considerations to note that for every fixed $s = \{i_1, \dots, i_t\}$ with $t \geq 2$ the following holds by Fubini's Theorem.

$$(6) \quad \int_{\mathbb{R}^t} \sum_{j\in s} \delta_{j,s}(\underline{u},v_{i_1},\dots,v_{i_t}) \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} f_{\theta_q}(u_q) \prod_{r\in s} g_{\theta_r}(v_r) d\tau(\underline{\theta}) \prod_{p\in s} d\mu(v_p)$$

$$= \int_{\mathbb{R}^k} \sum_{j\in s} \delta_{j,s}(\underline{u},v_{i_1},\dots,v_{i_t}) \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} f_{\theta_q}(u_q) g_{\theta_q}(v_q) d\tau(\underline{\theta}) \prod_{p=1}^{k} d\mu(v_p).$$

Remark 2. It should be noticed that in (6) we have introduced the v_i's with $i \notin s$ as dummy variables. This is necessary in order to prove our Lemma 3 below with the help of Eaton's (1967) results. Lemma 2, however, could also be proved without this trick.

For fixed $s = \{i_1, \dots, i_t\}$ with $t \geq 2$, the term following $\psi_{s,t}(\underline{u})$ in (5) is seen, after inserting the exponential density function (1), to be equal to

$$(7) \quad \int_{\mathbb{R}^k} \sum_{j\in s} \delta_{j,s}(\underline{u}, v_{i_1}, \dots, v_{i_t}) \int_{\Omega^k} L(\underline{\theta}, s, j) \prod_{q=1}^{k} \exp(\theta_q(u_q + v_q))$$

$$c_{n+m}(\theta_q) d\tau(\underline{\theta}) \prod_{r=1}^{k} b_m(v_r) d\mu(v_r) \tilde{\beta}(\underline{u})^{-1},$$

where $\tilde{\beta}(\underline{u}) = \int_{\Omega^k} \prod_{q=1}^{k} \exp(\theta_q u_q) c_n(\theta_q) d\tau(\underline{\theta})$ and $c_{n+m}(\theta) = c(\theta)^{n+m}$, $\theta \in \Omega$.

A change of variables $w_q = u_q + v_q$, $q = 1, \ldots, k$, reduces (7) to

$$(8) \quad \int_{\mathbb{R}^k} \sum_{j\in s} \delta_{j,s}(\underline{u}, w_{i_1}-u_{i_1} \ldots, w_{i_t}-u_{i_t}) \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} \exp(\theta_q w_q)$$

$$c_{n+m}(\theta_q) d\tau(\underline{\theta}) \prod_{r=1}^{k} b_m(w_r-u_r) d\mu(w_r) \tilde{\beta}(\underline{u})^{-1} .$$

Lemma 2. *Let* $(\phi, \psi, \delta) \in \mathscr{D}_I$ *be fixed. Then* $r(\tau, (\phi, \tilde{\psi}, \delta^*)) \leq r(\tau, (\phi, \psi, \delta))$ *and thus* $R(\underline{\theta}, (\phi, \tilde{\psi}, \delta^*)) \leq R(\underline{\theta}, (\phi, \psi, \delta))$, *uniformly in* $\underline{\theta} \in \Omega^k$.

Proof: In view of Lemma 1, we can assume that $\psi = \tilde{\psi}$ holds. The measure $d\hat{\tau}(\underline{\theta}) = \prod_{q=1}^{k} c_{n+m}(\theta_q) d\tau(\underline{\theta})$ is seen to be a permutation invariant σ-finite (c is continuous) measure on the Borel sets in Ω^k . Thus it follows from (L2) and Lemma 4.1 of Eaton (1967) that for every fixed $\underline{w} \in \mathbb{R}^k$,

$$(9) \quad \sum_{j\in s} \delta_{j,s}(\underline{u}, w_{i_1}-u_{i_1} \ldots, w_{i_t}-u_{i_t}) \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} \exp(\theta_q w_q) d\hat{\tau}(\underline{\theta})$$

is minimized if $\delta_{j,s} = \delta^*_{j,s}$, $j \in s$. This proves the first assertion, and the second one holds true in view of Remark 1.

Remark 3. As we have pointed out already at the end of Section 1, note that the optimal decision function $\delta^*_{j,s}$, $j \in s$, makes use only of the information contained in $(w_{i_1}, \ldots, w_{i_t})$.

Now, for $\delta_{j,s} = \delta^*_{j,s}$, $j \in s$, formula (8) reduces to

$$(10) \qquad \int_{\mathbb{R}^k} \min_{j\in s} \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} \exp(\theta_q w_q) d\hat{\tau}(\underline{\theta}) \prod_{r=1}^{k} b_m(w_r-u_r) d\mu(w_r) \tilde{\beta}(\underline{u})^{-1} .$$

From now on, let only $\underline{u} \in \mathbb{R}^k$ with $u_1 \leq u_2 \leq \ldots \leq u_k$ and $t \in \{2, \ldots, k\}$ be fixed and let $s \subseteq \{1, \ldots, k\}$ be variable subject to $|s| = t$. For notational convenience, let

$$(11) \qquad \mathcal{L}_s(\underline{w}) = \min_{j\in s} \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{q=1}^{k} \exp(\theta_q w_q) d\hat{\tau}(\underline{\theta}), \quad \underline{w} \in \mathbb{R}^k , \quad |s| = t .$$

<u>Lemma 3</u>. *Let* $\tilde{s} \subseteq \{1, \ldots, k\}$ *with* $1 \leq |\tilde{s}| \leq k-2$ *be fixed and let for* $p,q \in \{1, \ldots, k\}\setminus\tilde{s}$, $s_p = \tilde{s} \cup \{p\}$ *and* $s_q = \tilde{s} \cup \{q\}$. *Then for every* $\underline{w} \in \mathbb{R}^k$ *with* $w_p \leq w_q$, $\mathcal{L}_{s_q}(\underline{w}) \leq \mathcal{L}_{s_p}(\underline{w})$.

<u>Proof</u>: Consider $H_{s,j}(\underline{w}) = \int_{\Omega^k} L(\underline{\theta},s,j) \prod_{r=1}^{k} \exp(\theta_r w_r) d\hat{\tau}(\underline{\theta})$, with $s \subseteq \{1, \ldots, k\}$ and $|s| = |\tilde{s}| + 1$, $j \in s$, $\underline{w} \in \mathbb{R}^k$. As we have mentioned already, $\hat{\tau}$ is a permutation invariant σ-finite measure on the Borel sets of Ω^k . Moreover, for $\underline{w} \in \mathbb{R}^k$ with $w_p \leq w_q$, we have the following.

<u>(i)</u> $H_{s_q,j}(\underline{w}) \leq H_{s_p,j}(\underline{w})$, $j \in \tilde{s}$. This follows essentially in the same way as Lemma 4.1 of Eaton (1967). This is because for every fixed $j \in \tilde{s}$, we are concerned with a fixed size $|\tilde{s}| + 1$ subset selection problem, where (L3) assures that the assumptions concering the loss function in Eaton (1967) are satisfied.

(ii) $H_{s_q,q}(\underline{w}) \leq H_{s_p,p}$. The proof of this inequality is also very similar to that of Lemma 4.1 of Eaton (1967). This time we are concerned with a fixed size 1 subset selection problem, where now (L4) assures that the assumptions concerning the loss function in Eaton (1967) are satisfied.

Combining now (i) and (ii), the proof of Lemma 3 is completed.

For the last step in our considerations, we assume from now on that the underlying exponential family is strongly unimodal, i.e. that the densities in this family are log-concave. This is equivalent to function b being log-concave, as can be seen immediately. At this point let us recall that the measure μ is either the Lebesgue measure on $\mathbb{R}$ or the counting measure on $\mathbb{Z}$, or any other lattice in $\mathbb{R}$. This is crucial for the proof of our main result below, even in the case of $m = n = 1$.

Theorem. *Assume that the underlying exponential family is strongly unimodal and that the measure* μ *is either the Lebesgue measure on* $\mathbb{R}$ *or the counting measure on* $\mathbb{Z}$. *Let* $(\phi, \psi, \delta) \in \mathscr{D}_I$ *be fixed. Then* $r(\tau, (\phi, \psi^*, \delta^*)) \leq r(\tau, (\phi, \psi, \delta))$ *and thus* $R(\underline{\theta}, (\phi, \psi^*, \delta^*)) \leq R(\underline{\theta}, (\phi, \psi, \delta))$, *uniformly in* $\underline{\theta} \in \Omega^k$.

Proof: In the continuous as well as in the discrete case, strong unimodality of a member of the exponential family of densities is preserved under its convolutions w.r.t. μ . For details and references, see Barndorff-Nielsen (1978), chap. 6. Thus $b_m(x)$, $x \in \mathbb{R}$ or $x \in \mathbb{Z}$, respectively, is log-concave. By Lehmann (1959), p.330, this implies that $b_m(w-u)$ has monotone likelihood ratio in w with respect to u, $u,w \in \mathbb{R}$ or $u,w \in \mathbb{Z}$, respectively.

At the fixed $\underline{u} \in \mathbb{R}^k$ with $u_1 \leq u_2 \leq \ldots \leq u_k$ and $t \in \{2, \ldots, k\}$, in view of (5)-(8), (10) and (11), an optimal $\psi_{s,t}(\underline{u})$ equals 0, unless s with $|s| = t$ minimizes

$$(12) \qquad \int_{\mathbb{R}^k} \mathcal{L}_s(\underline{w}) \prod_{r=1}^{k} b_m(w_r - u_r) d\mu(w_r) .$$

Here we are concerned with a fixed size t subset selection problem. By Lemma 3, $\mathcal{L}_s(\underline{w})$ has the properties of the loss function assumed in Eaton (1967). Thus the proof is completed by an application of Lemma 4.2 in Eaton (1967) and in view of Remark 1.

<u>Corollary 1</u>. *Under the assumptions of the Theorem, the class of procedures* $(\phi, \psi^*, \delta^*) \in \mathcal{D}_I$ *constitutes an essentially complete class in* $\mathcal{D}_I$.

The proof of Corollary 1 as well as that of the first part of Corollary 2 is obvious. The second part follows from Blackwell and Girshick (1954), sec. 8.6. and the fact that the group of permutations is finite.

<u>Corollary 2</u>. *Under the assumptions of the Theorem, let* $(\phi, \psi, \delta) \in \mathcal{D}_I$ *be fixed. If* (ϕ, ψ, δ) *is minimax in* $\mathcal{D}_I$ *then* (ϕ, ψ^*, δ^*) *has the same property and, moreover, both procedures are minimax in* $\mathcal{D}$.

Remark 4. If one is looking for a Bayesian procedure with respect to a symmetric prior, then if there exists any, there will also be one of the type $(\phi, \psi^*, \delta^*) \in \mathscr{D}_I$. Thus the problem reduces to optimization of ϕ which, admittedly, will usually be still a difficult task. In the normal case (with unknown means and a common known variance) under a symmetric product normal prior a Bayes procedure has been studied by Gupta and Miescke (1982) in the known control case with respect to the loss function given in the Example in Section 1.

REFERENCES

[1] Barndorff-Nielsen, O. (1978): Information and Exponential Families in Statistics. Wiley, New York.

[2] Blackwell, D. and Girshick, M. A. (1954): Theory of Games and Statistical Decisions. Wiley, New York.

[3] Eaton, M. L. (1967): Some optimum properties of ranking procedures. Ann. Math. Statist. 38, 124-137.

[4] Gupta, S. S. and Miescke, K. J. (1981): On the least favorable configurations in certain two-stage selection procedures. In G.Kallianpur, P.R.Krishnaiah and J.K.Ghosh (eds.): Statistics and Probability: Essays in Honor of C.R.Rao. North-Holland, Amsterdam, 295-305.

[5] Gupta, S. S. and Miescke, K. J. (1982): On the problem of finding a best population with respect to a control in two stages. In S. S. Gupta and J. Berger (eds.): Statistical Decision Theory and Related Topics III, Proceedings of the Third Symposium at Purdue University, U. S. A., 24pp., to appear.

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

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER Mimeograph Series #81-53

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)*
An Essentially Complete Class of Two-Stage Selection Procedures with Screening at the First Stage

5. TYPE OF REPORT & PERIOD COVERED Technical

6. PERFORMING ORG. REPORT NUMBER Mimeo. Series #81-53

7. AUTHOR(s)
Shanti S. Gupta and Klaus-J. Miescke

8. CONTRACT OR GRANT NUMBER(s) ONR N00014-75-C-0455

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Purdue University
Department of Statistics
West Lafayette, IN 47906

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Washington, DC

12. REPORT DATE November 1981

13. NUMBER OF PAGES 14

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*

15. SECURITY CLASS. *(of this report)* Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*
Approved for public release, distribution unlimited.

17. DISTRIBUTION ST ENT *(of t abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY TES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*
Multiple decision procedures, two-stage selection procedures, screening procedures, sequential procedures, Bayesian procedures, essentially complete class.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Let $\pi_1,\ldots,\pi_k$ be given populations associated with unknown real parameters $\theta_1,\ldots,\theta_k$. The goal is to find a population with a sufficiently large parameter in two stages with screening out inferior populations at the first stage. Both, the control and the non-control situations are considered simultaneously. Let $\mathcal{D}_I$ denote the class of permutation invariant randomized procedures (ϕ, ψ, δ), where at Stage 1, ϕ and ψ decide how many populations and then which ones, respectively,

DD FORM 1 JAN 73 1473

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

are selected, and where at Stage 2, after additional samples have been drawn from the selected populations, δ makes the final decision. Let ψ^* and δ^* denote the natural decisions, i.e. which are associated with the largest sufficient statistics. Under the assumption of a common discrete or continuous type strongly unimodal exponential family it is shown that with respect to every reasonable loss function, procedures of the type (ϕ, ψ^*, δ^*) form an essentially complete class within $\mathcal{D}_I$.