

ON A SEQUENTIAL SUBSET SELECTION PROCEDURE
FOR EXPONENTIAL FAMILY DISTRIBUTIONS *

by

TaChen Liang
Wayne State University

Technical Report #88-20C

Department of Statistics
Purdue University

May 1988
Revised July 1989

* This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grant DMS-8606964 at Purdue University.

On A Sequential Subset Selection Procedure
For Exponential Family Distributions *

by

TaChen Liang
Wayne State University

Abstract

This paper deals with the problem of selecting the best population among k populations belonging to the same class of exponential family distributions through a sequential subset selection approach. We desire that the best population should be selected and each selected population should be good. Based on the modified likelihood ratio of the conditional frequency function of some statistics, an elimination-type sequential subset selection procedure is proposed. This sequential procedure achieves the selection goal with guaranteed probability at least P^* for some prespecified value P^* . At each stage, this procedure also provides some statistical inference about an upper bound on the measure of separation from the unknown best population to each remaining contending population. Finally, a modified sequential procedure to select a good population is also studied. This modified sequential procedure achieves the goal of selecting a good population with guaranteed probability at least P^* .

AMS 1980 Subject Classification: 62F07, 62L10

Key words and phrases: Best population, good population, correct selection, sequential selection procedure, subset selection, conditional likelihood function.

* This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grant DMS-8606964 at Purdue University.

On A Sequential Subset Selection Procedure
For Exponential Family Distributions

1. Introduction

Consider designing and analyzing an experiment for comparing k populations π_1, \dots, π_k . Suppose that observations can be obtained from the k populations sequentially. It is often desirable to terminate sampling from a population as soon as there is statistical evidence that it is not the best population, and this population is eliminated from further consideration. Selection through sequential comparison with elimination provides a significant advantage. To achieve a certain accuracy, it requires, on the average, substantially fewer samples than the fixed sample size procedures.

In sequential selection and ranking procedures, contributions have been made to select the best population by using the indifference zone approach. The simplest formulation of the indifference zone approach is the situation where one may wish to select only a single population and guarantee with a prespecified probability that the selected population is the best population provided some other condition on the parameters is satisfied, usually an indifference zone. However, in the real situation, it is hard or not possible to specify such a condition. Thus, a reasonable and useful approach is to derive some sequential selection procedure to select a small subset containing the best population. Further, it is desired that each of the selected populations should be not far from the best population. Therefore, some statistical inference is needed to assert that each selected population is within some prespecified distance from the best population.

In this paper, we assume that the random observations from population π_i has a frequency function $f(x|\theta_i)$ of the form

$$f(x|\theta_i) = \exp\{Q(\theta_i)P(x) + R(x) - \beta(\theta_i)\}. \quad (1.1)$$

Here, $P(x)$ and $R(x)$ do not involve the parameter θ , while $Q(\theta)$ and $\beta(\theta)$ do not depend on x . The function $Q(\theta)$ is assumed to be a continuous, strictly increasing function of θ .

Define $\delta_{ij} = Q(\theta_i) - Q(\theta_j)$ as a measure of separation from π_i to π_j . This particular

separation measure has been considered by Bechhofer, Kiefer and Sobel (1968). In many cases, δ_{ij} has a natural interpretation. For the normal distributions model, $\delta_{ij} = \theta_i - \theta_j$ is the difference between the two population means of populations π_i and π_j . For the case of the negative exponential distributions with means $\frac{1}{\theta_i}$, $\delta_{ij} = \theta_i - \theta_j$, and for the binomial distributions with the parameters of probability of success θ_i , $\delta_{ij} = \log \left[\left(\frac{\theta_i}{1-\theta_i} \right) / \left(\frac{\theta_j}{1-\theta_j} \right) \right]$ is the log odds ratio of the two binomial populations π_i and π_j . We call π_i the best population if π_i is the unique population such that $Q(\theta_i) = \max_{1 \leq j \leq k} Q(\theta_j)$. If more than one population has this property, one of them is tagged and considered as the best population. Since the function $Q(\theta)$ is increasing in θ , the population associated with $\theta_{(k)} \equiv \max_{1 \leq j \leq k} \theta_j$ is the best population. We let (k) denote the index of the best population and denote the best population by $\pi_{(k)}$. We note that a sequential procedure for selecting the best population from among exponential family distributions through the indifference zone approach has been derived by Hoel and Mazumdar (1968). Perng and Groves (1977) have also derived two sequential procedures for partitioning a set of one-parameter exponential populations with respect to a control. Since it is assumed that there is no prior information about the possible configurations of the measures of separation $\delta_{ij}, i, j = 1, \dots, k, i \neq j$, it seems to be hard to set up an indifference zone. Thus, a subset selection approach is more appropriate here. However, for the usual subset selection approach (for example, see Gupta (1965)), the size of the selected subset is a random number, and the quality of each of the selected populations cannot be guaranteed. In the fixed sample size case, Gupta and Santner (1973) and Santner (1975) introduced restricted subset selection procedures in which the size of the selected subset is at most m , where $1 \leq m \leq k - 1$. For the specified constants $\delta_2 > \delta_1 > 0$, Desu (1970) called a population π_i good if $\theta_{(k)} - \theta_i \leq \delta_1$, and bad if $\theta_{(k)} - \theta_i \geq \delta_2$. He studied the problem of selecting a subset of the k populations so that none of the bad populations are included in the selected subset. With the same selection goal, Santner (1976) proposed a two-stage selection procedure to select only good populations. Lam (1986) has studied a dual selection goal in which all the good populations should be included in the selected subset. In this paper, we try to control the quality of the selected populations.

For a prespecified positive constant δ^* , we say that population π_i is good if $Q(\theta_{(k)}) - Q(\theta_i) \leq \delta^*$ and bad otherwise. Let S denote the selected subset and $CS(\delta^*)$ denote the event that $\pi_{(k)} \in S$ and $\delta_{(k)i} \leq \delta^*$ for all $\pi_i \in S$. We desire a sequential subset selection procedure \mathcal{P} such that

$$P_{\theta} \{CS(\delta^*)|\mathcal{P}\} \geq P^* \quad \text{for all } \theta \in \Omega \quad (1.2)$$

where $P^*(k^{-1} < P^* < 1)$ is a preassigned probability level and $\Omega = \{\theta = (\theta_1, \dots, \theta_k) | f(x|\theta_i)$ is well defined, $i = 1, \dots, k\}$ is the parameter space.

For this particular measure of separation δ_{ij} , we consider an appropriate transformation of the random observations taken from any two populations. With this transformation, the likelihood function of the new statistics can be factored into two parts, one of which, obtained by a conditional argument, and termed the conditional likelihood function, is a function only of the parameter of interest. Based on this conditional likelihood function, a sequential subset selection procedure is derived. This sequential subset selection procedure achieves the selection goal described above. At each stage, it also provides some statistical inference about the bounds of separation between each remaining population and the unknown best population. Finally a modified sequential procedure to select a good population is also studied. This modified sequential procedure achieves the goal of selecting a good population with guaranteed probability at least P^* .

2. Some Properties Associated with Exponential Family Distributions

Let X_{in} denote the n th observation taken from population π_i . It is assumed that the random observations $X_{in}, n = 1, 2, \dots$ are independently distributed with the common frequency function $f(x|\theta_i), i = 1, \dots, k$. For each pair $(i, j), 1 \leq i, j \leq k, i \neq j$, let

$$W_{ij}(n) = (P(X_{in}) - P(X_{jn}))/2, \quad Z_{ij}(n) = (P(X_{in}) + P(X_{jn}))/2. \quad (2.1)$$

Then, $(W_{ij}(n), Z_{ij}(n))$ have a joint frequency function

$$g(w, z|\theta_i, \theta_j) = c \exp\{\delta_{ij}w + (Q(\theta_i) + Q(\theta_j))z + a(z, w) - \beta(\theta_i) - \beta(\theta_j)\} I_A(w, z), \quad (2.2)$$

and given $Z_{ij}(n) = z$, the conditional frequency function of $W_{ij}(n)$ is

$$g(w|z, \delta_{ij}) = \exp\{\delta_{ij}w + a(z, w) - \Psi(z, \delta_{ij})\}I_A(w, z), \quad (2.3)$$

where

$$\left\{ \begin{array}{l} \text{(a)} \quad a(z, w) = R(z + w) + R(z - w), \\ \text{(b)} \quad \Psi(z, \delta) = \log \int_{w \in A_z} \exp\{\delta w + a(z, w)\}dw, \\ \text{(c)} \quad \text{The set } A \text{ is the support of the random vector } (W_{ij}(n), Z_{ij}(n)), \\ \quad \text{and } A_z = \{w | (w, z) \in A\}, \text{ and} \\ \text{(d)} \quad c \text{ is a constant such that } g(w, z | \theta_i, \theta_j) \text{ is a frequency function.} \end{array} \right. \quad (2.4)$$

Note that for each z , $(w, z) \in A$ iff $(-w, z) \in A$. Thus, $w \in A_z$ if and only if $-w \in A_z$. Also, $a(z, w) = a(z, -w)$. Hence, by the definition of $\Psi(z, \delta)$, $\Psi(z, \delta) = \Psi(z, -\delta)$.

From (2.3), it is clear that $g(w|z, \delta_{ij})$ has an exponential family distribution and hence has the monotone likelihood ratio property. Therefore, $E_\delta[h(W_{ij}(n))|Z_{ij}(n)]$ is nondecreasing in δ whenever $h(w)$ is any nondecreasing function of w . Also, from the usual theory for the exponential family distributions, for given $Z_{ij}(n) = z$, the conditional mean and variance of $W_{ij}(n)$, for any δ_0 , are:

$$\left\{ \begin{array}{l} E_{\delta_0}[W_{ij}(n)|Z_{ij}(n) = z] = \frac{\partial}{\partial \delta} \Psi(z, \delta) \Big|_{\delta=\delta_0} \equiv \Psi_\delta(z, \delta_0), \\ \text{Var}_{\delta_0}(W_{ij}(n)|Z_{ij}(n) = z) = \frac{\partial^2}{\partial \delta^2} \Psi(z, \delta) \Big|_{\delta=\delta_0} \equiv \Psi_{\delta\delta}(z, \delta_0) > 0. \end{array} \right. \quad (2.5)$$

The following lemma can be obtained easily based on the above discussion.

Lemma 2.1. For each z belonging to the domain of the random variable $Z_{ij}(n)$,

- (a) $\Psi(z, \delta) = \Psi(z, -\delta)$ from all δ .
- (b) $\Psi(z, \delta)$ is a convex function of δ .
- (c) $\Psi(z, \delta)$ is strictly increasing (decreasing) in δ for all $\delta \in [0, \infty)$

$$(\delta \in (-\infty, 0]), \text{ and } \Psi(z, 0) = \inf_{\delta} \Psi(z, \delta).$$

- (d) $\Psi_\delta(z, \delta_0)$ is strictly increasing in δ_0 , $\Psi_\delta(z, 0) = 0$ and $\Psi_\delta(z, -\delta_0) = -\Psi_\delta(z, \delta_0)$.

(e) Let a_1, a_2, b_1 and b_2 be points such that $a_1 \leq a_2 < b_2, a_1 < b_1 \leq b_2$.

$$\text{Then, } \frac{\Psi(z, b_1) - \Psi(z, a_1)}{b_1 - a_1} \leq \frac{\Psi(z, b_2) - \Psi(z, a_2)}{b_2 - a_2}.$$

(f) For each fixed z , define

$$I(b|z) = b\Psi_\delta(z, b) - \Psi(z, b) + \Psi(z, 0). \quad (2.6)$$

Then, $I(b|z) \geq 0$ for all b and $I(b|z) > 0$ if $b \neq 0$.

3. A Sequential Subset Selection Procedure $\mathcal{P}(H, \delta^*)$

Let $\delta^* > 0$ be the prespecified value used to define the event $CS(\delta^*)$. Let $H(y)$ be a distribution function satisfying Condition A.

Condition A: $H(y)$ is a distribution function defined on the interval $[0, \delta^*]$ such that for some constant $c_0 \in (0, \delta^*)$, the interval $[0, c_0]$ is contained in the support of H .

For each $a \geq 0, n \geq 1$ and each pair (i, j) , define

$$L_{ij}(n, a) = \frac{\int_0^{\delta^*} \prod_{m=1}^n g(W_{ij}(m)|Z_{ij}(m), y) dH(y)}{\prod_{m=1}^n g(W_{ij}(m)|Z_{ij}(m), -a)}, \quad (3.1)$$

where $g(w|z, \delta)$ is the conditional frequency function defined in (2.3). We now define a sequential subset selection procedure $\mathcal{P}(H, \delta^*)$ as follows:

Let $S_0 = \{\pi_1, \dots, \pi_k\}$. For each $n \geq 1$, define

$$S_n = \{\pi_i \in S_{n-1} | L_{ji}(n, 0) < \frac{k-1}{1-P^*} \text{ for all } \pi_j \in S_{n-1} - \{\pi_i\}\}. \quad (3.2)$$

That is, S_n is the set of contending populations up to stage n . At stage n , population $\pi_i \in S_n$ is labelled as good if $L_{ji}(n, \delta^*) \geq \frac{k-1}{1-P^*}$ for all $\pi_j \in S_n - \{\pi_i\}$. Let $|S_n|$ denote the size of the set S_n . If either $|S_n| = 1$ or all the populations in S_n have been labelled as good then the procedure terminates and we take $S = S_n$ as the selected subset; otherwise, we go to the next stage. The procedure is continued in this way.

Termination with Probability One

In order to apply the sequential selection procedure $\mathcal{P}(H, \delta^*)$, we need to assert that this procedure terminates with probability one. As described above, the distribution function $H(y)$ is chosen to satisfy Condition A. It is also assumed that Condition B holds true, where

Condition B: $\Psi_{\delta\delta}(z, b) \leq M(b)$ a.e. $(Z_{ij}(n))$ for each $b > 0$, where $M(\cdot)$ is bounded on the interval $[a, \infty)$ for each $a > 0$. That is, there is a finite function $q(\cdot)$ such that $M(b) \leq q(a)$ for all $b \geq a > 0$.

It is not hard to verify that many exponential family distributions, including normal, exponential, binomial, satisfy Condition B.

To prove the procedure $\mathcal{P}(H, \delta^*)$ terminating with probability one, it suffices to show that for any two populations, say π_1 and π_2 , with probability one the event E occurs, where E is the event that either one of them will be eliminated (in comparison with the other) or both of them are labelled as good. Without loss of generality, it is assumed that $\theta_1 \geq \theta_2$. For simplicity, in the following of this section, we let $k = 2$ and $\underline{\theta} = (\theta_1, \theta_2)$.

Defining stopping times T_1 and T_2 as follows:

$$T_1 = \max\{T_{ij}, i, j = 1, 2, i \neq j\} \quad (3.3)$$

where

$$T_{ij} = \min \left\{ n \mid L_{ij}(n, \delta^*) \geq \frac{k-1}{1-P^*} \right\}, \quad (3.4)$$

and

$$T_2 = \min \left\{ n \mid L_{12}(n, 0) \geq \frac{k-1}{1-P^*} \right\}. \quad (3.5)$$

That is, T_1 is the random time at which both π_1 and π_2 have been labelled as good (assuming no elimination) and T_2 is the random time at which π_2 is eliminated by π_1 (assuming that π_2 cannot eliminate π_1 and there is no labelling for good). Note that $\{T_1 < \infty \text{ or } T_2 < \infty\} \subset E$. Thus, it suffices to prove that $P_{\underline{\theta}}\{T_1 < \infty \text{ or } T_2 < \infty\} = 1$ for all possible configurations of the parameter $\underline{\theta}$.

We have the following two lemmas. The proofs are given in the Appendix.

Lemma 3.1. Under Conditions A and B, $P_{\underline{\theta}}\{T_2 < \infty\} = 1$ if $\theta_1 > \theta_2$.

Lemma 3.2. Under Condition A, $P_{\underline{\theta}}\{T_1 < \infty\} = 1$ if $\theta_1 = \theta_2$.

Thus, one can see under Conditions A and B that $P_{\underline{\theta}}\{T_1 < \infty \text{ or } T_2 < \infty\} = 1$ for all possible configurations of the parameter $\underline{\theta}$. Hence, under Conditions A and B, the sequential subset selection procedure $\mathcal{P}(H, \delta^*)$ terminates with probability one.

4. Probability of a Correct Selection

For each $n \geq 1$, let $\mathcal{F}_{ij}(n)$ denote the σ -field generated by the random variables $(W_{ij}(m), Z_{ij}(m), m = 1, 2, \dots, n)$. Then, we have the following lemma.

Lemma 4.1. $E_{\underline{\theta}}[L_{ij}(n, \delta_{ji}) | \mathcal{F}_{ij}(n-1)] = L_{ij}(n-1, \delta_{ji})$ for all $n \geq 1$ and $\underline{\theta} \in \Omega$.

Proof: Note that

$$\begin{aligned} & E_{\underline{\theta}}[L_{ij}(n, \delta_{ji}) | \mathcal{F}_{ij}(n-1)] \\ &= E_{\underline{\theta}} \left[\int_0^{\delta^*} \exp \left\{ (y + \delta_{ji}) \sum_{m=1}^n W_{ij}(m) - \sum_{m=1}^n [\Psi(Z_{ij}(m), y) - \Psi(Z_{ij}(m), \delta_{ji})] \right\} dH(y) | \mathcal{F}_{ij}(n-1) \right] \\ &= \int_0^{\delta^*} \exp \left\{ (y + \delta_{ji}) \sum_{m=1}^{n-1} W_{ij}(m) - \sum_{m=1}^{n-1} [\Psi(Z_{ij}(m), y) - \Psi(Z_{ij}(m), \delta_{ji})] \right\} A_{ij}(\underline{\theta}, y, n) dH(y), \end{aligned}$$

where $A_{ij}(\underline{\theta}, y, n) = E_{\underline{\theta}}[\exp\{(y + \delta_{ji})W_{ij}(n) - \Psi(Z_{ij}(n), y) + \Psi(Z_{ij}(n), \delta_{ji})\}]$ and the second equality is obtained from Fubini's theorem and by the independence between $(W_{ij}(n), Z_{ij}(n))$ and $\mathcal{F}_{ij}(n-1)$. Now,

$$A_{ij}(\underline{\theta}, y, n) = E_{\underline{\theta}} [E_{\delta_{ij}}[\exp\{(y + \delta_{ji})W_{ij}(n) - \Psi(Z_{ij}(n), y) + \Psi(Z_{ij}(n), \delta_{ji})\} | Z_{ij}(n)]] = 1$$

and therefore, the result follows.

Theorem 4.1. Let $\{S_n\}$ be the sequence of the sets of contending populations determined by the sequential subset selection procedure $\mathcal{P}(H, \delta^*)$ through (3.2). Then, for all $\underline{\theta} \in \Omega$,

$$P_{\underline{\theta}}\{L_{i(k)}(n, \delta_{(k), i}) < \frac{k-1}{1-P^*} \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \geq P^*,$$

provided that Conditions A and B hold.

Proof: Note that for any $\theta \in \Omega$.

$$\begin{aligned}
& P_{\theta} \{L_{i(k)}(n, \delta_{(k),i}) < \frac{k-1}{1-P^*} \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \\
&= 1 - P_{\theta} \{L_{i(k)}(n, \delta_{(k),i}) \geq \frac{k-1}{1-P^*} \text{ for some } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for some } n \geq 1\} \\
&\geq 1 - \sum_{\substack{i=1 \\ i \neq (k)}}^k P_{\theta} \{L_{i(k)}(n, \delta_{(k),i}) \geq \frac{k-1}{1-P^*} \text{ for some } n \geq 1\}. \tag{4.1}
\end{aligned}$$

By Lemma 4.1, $\{L_{i(k)}(n, \delta_{(k),i}), \mathcal{F}_{i(k)}(n); n \geq 1\}$ forms a nonnegative martingale. Thus, from a lemma of Robbins and Siegmund (1973), for $i \neq (k)$,

$$P_{\theta} \{L_{i(k)}(n, \delta_{(k),i}) \geq \frac{k-1}{1-P^*} \text{ for some } n \geq 1\} \leq \frac{1-P^*}{k-1}. \tag{4.2}$$

Then by (4.1) and (4.2), the result follows.

For each pair (i, j) and $a \geq 0$, let $B_{ij}(n, a)$ denote the event that $L_{ij}(n, a) < \frac{k-1}{1-P^*}$. That is $B_{ij}(n, a) = \{L_{ij}(n, a) < \frac{k-1}{1-P^*}\}$. The following result is very helpful for obtaining some sequential estimate of an upper bound on $\delta_{(k),i}, i = 1, 2, \dots, k$.

Lemma 4.2. Let a and b be any two nonnegative values such that $b \geq a \geq 0$. Then,

$$B_{ij}(n, b) \subset B_{ij}(n, a) \text{ for all } n \geq 1. \tag{4.3}$$

Proof: Suppose that the statement of (4.3) is not true. Then, there exists some $n \geq 1$ such that

$$L_{ij}(n, b) < \frac{k-1}{1-P^*} \leq L_{ij}(n, a). \tag{4.4}$$

By (3.1), we have

$$L_{ij}(n, b) = L_{ij}(n, a) \exp\left\{(b-a) \sum_{m=1}^n W_{ij}(m) + \sum_{m=1}^n [\Psi(Z_{ij}(m), b) - \Psi(Z_{ij}(m), a)]\right\}. \tag{4.5}$$

(4.4) and (4.5) together imply that

$$\sum_{m=1}^n W_{ij}(m) < \frac{-1}{b-a} \sum_{m=1}^n [\Psi(Z_{ij}(m), b) - \Psi(Z_{ij}(m), a)]. \tag{4.6}$$

Then, by (4.4) and (4.6),

$$\begin{aligned}
1 &< \frac{k-1}{1-P^*} \\
&\leq L_{ij}(n, a) \\
&= \int_0^{\delta^*} \exp \left\{ (y+a) \sum_{m=1}^n W_{ij}(m) + \sum_{m=1}^n [\Psi(Z_{ij}(m), a) - \Psi(Z_{ij}(m), y)] \right\} dH(y) \quad (4.7) \\
&\leq \int_0^{\delta^*} \exp \left\{ -(y+a) \sum_{m=1}^n Q_{ij}(m, y, a, b) \right\} dH(y),
\end{aligned}$$

where

$$Q_{ij}(m, y, a, b) = \frac{\Psi(Z_{ij}(m), y) - \Psi(Z_{ij}(m), a)}{y - (-a)} - \frac{\Psi(Z_{ij}(m), -a) - \Psi(Z_{ij}(m), -b)}{(-a) - (-b)}.$$

By Lemma 2.1 and the fact that $-b < -a \leq y$ for each $y \in [0, \delta^*]$, we have, for each $m = 1, 2, \dots, n$, $Q_{ij}(m, y, a, b) \geq 0$. With this fact and from (4.7), we obtain $1 < \frac{k-1}{1-P^*} \leq L_{ij}(n, a) < 1$, which is a contradiction. Therefore, $B_{ij}(n, b) \subset B_{ij}(n, a)$ for all $n \geq 1$ when $b \geq a \geq 0$.

An immediate result of Lemma 4.2 is:

$$\bigcap_{m=1}^n B_{ij}(m, b) \subset \bigcap_{m=1}^n B_{ij}(m, a) \text{ for all } n \geq 1 \text{ whenever } b \geq a \geq 0. \quad (4.8)$$

For each $n \geq 1$, $\pi_i, \pi_j \in S_{n-1}$, $i \neq j$, define $C_{ij}(n)$ and $D_{ij}(n)$ as follows:

$$C_{ij}(n) = \left\{ a \geq 0 \mid L_{ij}(n, a) < \frac{k-1}{1-P^*} \right\}. \quad (4.9)$$

$$D_{ij}(n) = \begin{cases} \sup C_{ij}(n) & \text{if } C_{ij}(n) \neq \phi, \\ 0 & \text{if } C_{ij}(n) = \phi, \end{cases} \quad (4.10)$$

where ϕ denotes the empty set. Also, let $D_{ii}(n) = 0$. From Lemma 4.2, if $D_{ij}(n) > 0$, then, $L_{ij}(n, a) < \frac{k-1}{1-P^*}$ for all $a \in [0, D_{ij}(n))$ and $L_{ij}(n, b) \geq \frac{k-1}{1-P^*}$ for all $b > D_{ij}(n)$.

For each $n \geq 2$, if $\pi_i \in S_{n-1}$, define

$$D_i(n) = \min_{1 \leq m \leq n} \left(\max_{\pi_j \in S_{m-1}} D_{ij}(m) \right). \quad (4.11)$$

If $\pi_i \notin S_{n-1}$, let $n_i = \max\{m \mid \pi_i \in S_{m-1}\}$ and $D_i(n) = D_i(n_i)$.

By the definition of $D_i(n)$, for each $i = 1, \dots, k$, $\{D_i(n)\}$ is a nonincreasing sequence and bounded below by zero. The value $D_i(n)$ will be used at stage n as an estimator of an upper bound of $\delta_{(k)i}$. We note that the technique used to define $D_i(n)$ is from Hsu and Edwards (1983) for the location parameter model.

Theorem 4.2. Let $L_{ij}(n, a)$, S_n and $D_i(n)$, $n \geq 1$ be those defined in (3.1), (3.2) and (4.1), respectively. Then

$$\begin{aligned} & \{L_{i(k)}(n, \delta_{(k)i}) < \frac{k-1}{1-P^*} \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \\ & \subset \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq D_i(n) \text{ for all } \pi_i \in S_{n-1} \text{ for all } n \geq 1\}. \end{aligned}$$

Proof: By Lemma 4.2 and the fact that $\delta_{(k)i} \geq 0$, we have

$$\begin{aligned} & \{L_{i(k)}(n, \delta_{(k)i}) < \frac{k-1}{1-P^*} \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \\ & \subset \{L_{i(k)}(n, 0) < \frac{k-1}{1-P^*} \text{ and } \delta_{(k)i} \leq D_{i(k)}(n) \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \\ & \subset \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq D_{i(k)}(n) \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} \quad (4.12) \\ & \subset \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq \max_{\pi_j \in S_{n-1}} D_{ij}(n) \text{ for all } \pi_i \in S_{n-1} \text{ for all } n \geq 1\} \\ & = \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq D_i(n) \text{ for all } \pi_i \in S_{n-1} \text{ for all } n \geq 1\}. \end{aligned}$$

An immediate consequence of Theorems 4.1 and 4.2 is that for all $\theta \in \Omega$, under Conditions A and B.

$$P_{\theta} \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq D_i(n) \text{ for all } \pi_i \in S_{n-1} \text{ for all } n \geq 1\} \geq P^*. \quad (4.13)$$

This result provides a sequential comparison inference, with confidence level at least P^* , as follows: simultaneously, at each stage n , the best population is not eliminated and the separation from the unknown best population to each contending population, say π_i , up to stage n , is not larger than the value $D_i(n)$ for all $n \geq 1$. Another consequence of Theorems 4.1 and 4.2 is that when the sequential selection procedure $\mathcal{P}(H, \delta^*)$ terminates, the event of a correct selection $CS(\delta^*)$ is guaranteed with probability at least P^* . We state this result as a theorem as follows:

Theorem 4.3. Let $\mathcal{P}(H, \delta^*)$ be the sequential subset selection procedure defined in Section 3. Then, $P_{\theta}\{CS(\delta^*)|\mathcal{P}(H, \delta^*)\} \geq P^*$ holds for all $\theta \in \Omega$.

Proof: Note that when the sequential selection procedure $\mathcal{P}(H, \delta^*)$ terminates, then either $|S| = 1$ or all the populations in S must have been labelled as good at some stage. Let N be the stopping time of the sequential selection procedure $\mathcal{P}(H, \delta^*)$ and when $|S| \geq 2$, for each $\pi_i \in S$, let N_i denote the first time at which π_i was labelled as good. Hence, $L_{ij}(N_i, \delta^*) \geq \frac{k-1}{1-P^*}$ for all $\pi_j \in S_{N_i} - \{\pi_i\}$. By the definition of $D_{ij}(n)$ and Lemma 4.2, $D_{ij}(N_i) \leq \delta^*$ for all $\pi_j \in S_{N_i} - \{\pi_i\}$, and thus, $D_{i(k)}(N_i) \leq \delta^*$ if $\pi_{(k)} \in S_{N_i} - \{\pi_i\}$. Also, note that $S = S_N$ and when $|S| \geq 2$, $N_i \leq N$ for all $\pi_i \in S$. Now

$$\begin{aligned} & \{\pi_{(k)} \in S \text{ and } \delta_{(k)i} \leq D_{i(k)}(n) \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\} & (4.14) \\ \subset & \{\pi_{(k)} \in S \text{ and } |S| = 1\} \cup \{\pi_{(k)} \in S, |S| \geq 2 \text{ and } \delta_{(k)i} \leq D_{i(k)}(N_i) \text{ for all } \pi_i \in S - \{\pi_{(k)}\}\} \\ & \text{(since } S \subset S_{n-1} \text{ for all } n \text{ and } N_i \leq N \text{ for all } \pi_i \in S) \\ \subset & \{\pi_{(k)} \in S \text{ and } |S| = 1\} \cup \{\pi_{(k)} \in S, |S| \geq 2 \text{ and } \delta_{(k)i} \leq \delta^* \text{ for all } \pi_i \in S - \{\pi_{(k)}\}\} \\ = & CS(\delta^*). \end{aligned}$$

Then, by (4.14), (4.12) and Theorem 4.1, we have that for all $\theta \in \Omega$,

$$P_{\theta}\{CS(\delta^*)|\mathcal{P}(H, \delta^*)\} \geq P^*.$$

5. A Sequential Procedure for Selecting a Good Population

Consider a selection problem among k populations. In most applications, an experimenter is usually content with selecting a good population. Kao and Lai (1980) have studied the problem of selecting a good population among k normal populations through a sequential approach. Edwards (1987) has studied a sequential procedure for selecting a good population for location-parameter models. For our problem, with this selection goal, what we want is in fact a selection procedure having the property that

$$P_{\theta}\{\pi_i \text{ is selected and } \delta_{(k)i} \leq \delta^*\} \geq P^* \text{ for all } \theta \in \Omega. \quad (5.1)$$

Also, the procedure should stop as soon as we are confident that a good population has been found. A sequential selection procedure achieving this selection goal can be obtained from $\mathcal{P}(H, \delta^*)$ with a little modification on its stopping strategy. We now describe it as follows.

Let $\{S_n\}$ be the sequence of sets of remaining populations defined in (3.2). For each $\pi_i \in S_n$, π_i is labelled as good if $L_{ij}(n, \delta^*) \geq \frac{k-1}{1-P^*}$ for all $\pi_j \in S_n - \{\pi_i\}$. The procedure terminates at stage n as soon as one of the populations in S_n has been labelled as good. If there is only one population in S_n having been labelled as good, we then select this population as a good population. If more than one population have been labelled as good, we usually select the one with the smallest $D_i(n)$ value. We denote this modified sequential selection procedure by \mathcal{P}' .

The following theorem guarantees the desired confidence level of selecting a good population by applying the selection procedure \mathcal{P}' .

Theorem 5.1. Let \mathcal{P}' be the sequential selection procedure defined above. Then, $P_\theta\{\pi_i \text{ is selected and } \delta_{(k)i} \leq \delta^* | \mathcal{P}'\} \geq P^*$ for all $\theta \in \Omega$, provided Conditions A and B hold.

Note: The argument for the proof of this theorem is similar to that of Theorem 7 of Kao and Lai (1980). For completeness, the proof is given as follows.

Proof: Let $A(\theta) = \{L_{i(k)}(n, \delta_{k(i)}) < \frac{k-1}{1-P^*} \text{ for all } \pi_i \in S_{n-1} - \{\pi_{(k)}\} \text{ for all } n \geq 1\}$. Then by Theorem 4.1, $P_\theta\{A(\theta)\} \geq P^*$ for all $\theta \in \Omega$. We note that on the event $A(\theta)$, by Lemma 4.2, $L_{i(k)}(n, 0) < \frac{k-1}{1-P^*}$ for all $\pi_i \in S_{n-1} - \{\pi_{(k)}\}$ for all $n \geq 1$. Therefore, $\pi_{(k)}$ can never be eliminated in comparison with other populations at any stage. Let $B(\theta) = \{\pi_i | \delta_{(k)i} > \delta^*\}$. That is, $B(\theta)$ is the set of bad populations. It suffices to show that on the event $A(\theta)$, any population in $B(\theta)$ cannot be labelled as good when the procedure \mathcal{P}' terminates.

Let M be the stopping time of the sequential selection procedure \mathcal{P}' . On $A(\theta)$, since $\pi_{(k)}$ can never be eliminated prior to the stopping time M , then $\pi_{(k)} \in S_n$ for all $1 \leq n \leq M$. Moreover, for each $\pi_i \in B(\theta)$, $\delta_{(k)i} > \delta^* > 0$ and on $A(\theta)$, $L_{i(k)}(n, \delta_{(k)i}) < \frac{k-1}{1-P^*}$ for all $1 \leq n \leq M$. Then by Lemma 4.2, $L_{i(k)}(n, \delta^*) < \frac{k-1}{1-P^*}$ which means that π_i cannot be

labelled as good. Hence, π_i cannot be selected as a good population. Therefore, for all $\theta \in \Omega$, we have $P_\theta\{\pi_i \text{ is selected and } \delta_{(k)i} \leq \delta^* | \mathcal{P}'\} \geq P^*$.

Appendix

Proof of Lemma 3.1. The proof is analogous to that of Lemma 1 of Pollak and Siegmund (1975). Note that $\delta_{12} > 0$ since $\theta_1 > \theta_2$. For convenience, in the following, we let $W(m) = W_{12}(m)$, $Z(m) = Z_{12}(m)$, $\lambda(m, a, b) = \Psi(Z_{12}(m), a) - \Psi(Z_{12}(m), b)$, $\delta_0 = \delta_{12}$ and $\alpha = \frac{k-1}{1-P^*}$. Then,

$$L_{12}(n, 0) = \int_0^{\delta^*} \exp \left\{ y \sum_{m=1}^n W(m) - \sum_{m=1}^n \lambda(m, y, 0) \right\} dH(y).$$

For each n , define $T_{2n} = \min\{T_2, n\} - 1$, so that $T_{2n} + 1$ is a bounded stopping time.

Case 1. When $\delta_0 < c_0$ where c_0 is given in Condition A.

From the definition of T_2 , we have

$$\begin{aligned} \log \alpha &> \log L_{12}(T_{2n}, 0) \\ &= \delta_0 \sum_{m=1}^{T_{2n}} W(m) - \sum_{m=1}^{T_{2n}} \lambda(m, \delta_0, 0) \\ &\quad + \log \int_0^{\delta^*} \exp \left\{ (y - \delta_0) \sum_{m=1}^{T_{2n}} W(m) - \sum_{m=1}^{T_{2n}} \lambda(m, y, \delta_0) \right\} dH(y) \quad (\text{A.1}) \\ &\geq \delta_0 \sum_{m=1}^{T_{2n}} [W(m) - \Psi_\delta(Z(m), \delta_0)] - \sum_{m=1}^{T_{2n}} [\delta_0 \Psi(Z(m), \delta_0) - \lambda(m, \delta_0, 0)] \\ &\quad + \log \int_{|y - \delta_0| < \varepsilon} \exp \left\{ (y - \delta_0) \sum_{m=1}^{T_{2n}} W(m) - \sum_{m=1}^{T_{2n}} \lambda(m, y, \delta_0) \right\} dH(y). \end{aligned}$$

By using Taylor's expansion,

$$\lambda(m, y, \delta_0) = (y - \delta_0) \Psi_\delta(z, \delta_0) + \frac{1}{2} (y - \delta_0)^2 \Psi_{\delta\delta}(z, \xi) \quad (\text{A.2})$$

where $\xi = \xi(z, y, \delta_0)$ is a value between y and δ_0 . By Condition B, we can find an ε so small that $\frac{1}{2} |y - \delta_0| \Psi_{\delta\delta}(Z(m), \xi) \leq 1$ a.e. for all $y \in (\delta_0 - \varepsilon, \delta_0 + \varepsilon)$. We then obtain

$$\log \int_{|y - \delta_0| < \varepsilon} \exp \left\{ (y - \delta_0) \sum_{m=1}^{T_{2n}} W(m) - \sum_{m=1}^{T_{2n}} \lambda(m, y, \delta_0) \right\} dH(y)$$

$$\begin{aligned}
&\geq \log \int_{|y-\delta_0|<\varepsilon} \exp \left\{ (y - \delta_0) \sum_{m=1}^{T_{2n}} [W(m) - \Psi_\delta(Z(m), \delta_0)] - \varepsilon T_{2n} \right\} dH(y) \quad (\text{A.3}) \\
&\geq \log H(\delta_0 - \varepsilon, \delta_0 + \varepsilon) + (d - \delta_0) \sum_{m=1}^{T_{2n}} [W(m) - \Psi_\delta(Z(m), \delta_0)] - \varepsilon T_{2n},
\end{aligned}$$

where $H(\delta_0 - \varepsilon, \delta_0 + \varepsilon) = \int_{|y-\delta_0|<\varepsilon} dH(y)$, $d = \int_{|y-\delta_0|<\varepsilon} y dH(y) / H(\delta_0 - \varepsilon, \delta_0 + \varepsilon)$, and the second inequality is obtained by Jensen's inequality. From (2.6), (A.1) and (A.3), we have

$$\begin{aligned}
&\sum_{m=1}^{T_{2n}} [I(\delta_0|Z(m)) - \varepsilon] + d \sum_{m=1}^{T_{2n}} [W(m) - \Psi_\delta(Z(m), \delta_0)] \\
&\leq \log \alpha - \log H(\delta_0 - \varepsilon, \delta_0 + \varepsilon). \quad (\text{A.4})
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_{m=1}^{T_{2n}+1} [I(\delta_0|Z(m)) - \varepsilon] + d \sum_{m=1}^{T_{2n}+1} [W(m) - \Psi_\delta(Z(m), \delta_0)] \\
&\leq \log \alpha - \log H(\delta_0 - \varepsilon, \delta_0 + \varepsilon) + I(\delta_0|Z(T_{2n} + 1)) - \varepsilon \\
&\quad + d[W(T_{2n} + 1) - \Psi_\delta(Z(T_{2n} + 1), \delta_0)]. \quad (\text{A.5})
\end{aligned}$$

Now consider the expectation E_θ of both sides of (A.5). Since

$$\begin{aligned}
E_\theta[|W(m) - \Psi_\delta(Z(m), \delta_0)|] &= E_\theta \left[E_{\delta_0}[|W(m) - \Psi_\delta(Z(m), \delta_0)| \mid Z(m)] \right] \\
&\leq \left\{ E_\theta \left[E_{\delta_0}[|W(m) - \Psi_\delta(Z(m), \delta_0)|^2 \mid Z(m)] \right] \right\}^{\frac{1}{2}} \\
&= \{E_\theta[\Psi_\delta(Z(m), \delta_0)]\}^{\frac{1}{2}} \\
&\leq M^{\frac{1}{2}}(\delta_0),
\end{aligned}$$

then,

$$E_\theta[W(m) - \Psi_\delta(Z(m), \delta_0)] = E_\theta[E_{\delta_0}[W(m) - \Psi_\delta(Z(m), \delta_0) \mid Z(m)]] = 0$$

Also,

$$0 < E_\theta[I(\delta_0|Z(m))] \leq E_\theta[\delta_0 \Psi_\delta(Z(m), \delta_0)] = \delta_0 E_\theta[W(m)] < \infty.$$

Thus, by Wald's lemma, $E_{\tilde{\theta}} \left[\sum_{m=1}^{T_{2n}+1} [W(m) - \Psi_{\delta}(Z(m), \delta_0)] \right] = 0$, and

$$E_{\tilde{\theta}} \left[\sum_{m=1}^{T_{2n}+1} [I(\delta_0|Z(m)) - \varepsilon] \right] = [E_{\tilde{\theta}}(T_{2n} + 1)] \times [E_{\tilde{\theta}}[I(\delta_0|Z(1))] - \varepsilon].$$

Also, by the Schwarz inequality,

$$\begin{aligned} & E_{\tilde{\theta}}[W(T_{2n} + 1) - \Psi_{\delta}(Z(T_{2n} + 1), \delta_0)] \\ & \leq E_{\tilde{\theta}}[|W(T_{2n} + 1) - \Psi_{\delta}(Z(T_{2n} + 1), \delta_0)|] \\ & \leq \{E_{\tilde{\theta}}[(W(1) - \Psi_{\delta}(Z(1), \delta_0))^2][E_{\tilde{\theta}}(T_{2n} + 1)]\}^{\frac{1}{2}} \\ & \leq M^{\frac{1}{2}}(\delta_0)[E_{\tilde{\theta}}(T_{2n} + 1)]^{\frac{1}{2}}. \end{aligned}$$

Similarly,

$$\begin{aligned} E_{\tilde{\theta}}[I(\delta_0|Z(T_{2n} + 1))] & \leq \delta_0 E_{\tilde{\theta}}[\Psi_{\delta}(Z(T_{2n} + 1), \delta_0)] \\ & = \delta_0 E_{\tilde{\theta}}[W(T_{2n} + 1)] \\ & \leq \delta_0 \{E_{\tilde{\theta}}[W^2(1)]E_{\tilde{\theta}}[T_{2n} + 1]\}^{\frac{1}{2}}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \{E_{\tilde{\theta}}[I(\delta_0|Z(1))] - \varepsilon\} \times \{E_{\tilde{\theta}}[T_{2n} + 1]\} \\ & \leq \{\delta_0[E_{\tilde{\theta}}(W^2(1))]^{\frac{1}{2}} + dM^{\frac{1}{2}}(\delta_0)\} \times \{E_{\tilde{\theta}}[T_{2n} + 1]\}^{\frac{1}{2}} \\ & \quad - \log H(\delta_0 - \varepsilon, \delta_0 + \varepsilon) + \log \alpha, \end{aligned}$$

and so,

$$\begin{aligned} & \{E_{\tilde{\theta}}[I(\delta_0|Z(1))] - \varepsilon\} \times \{E_{\tilde{\theta}}[T_{2n} + 1]\}^{\frac{1}{2}} \\ & \leq \{\delta_0[E_{\tilde{\theta}}(W^2(1))]^{\frac{1}{2}} + dM^{\frac{1}{2}}(\delta_0)\} - \log H(\delta_0 - \varepsilon, \delta_0 + \varepsilon) + \log \alpha. \end{aligned} \quad (\text{A.6})$$

Now, since $I(\delta_0|Z(1))$ is positive, ε can be chosen so small that $E_{\tilde{\theta}}[I(\delta_0|Z(1))] - \varepsilon > 0$. While the right-hand-side of (A.6) is independent of n , so, as $n \rightarrow \infty$, the left-hand-side of (A.6) is still bounded. Hence $E_{\tilde{\theta}}(T_2) < \infty$ which implies that $P_{\tilde{\theta}}\{T_2 < \infty\} = 1$.

Case 2. When $\delta_0 > c_0$. From (A.1), we also have

$$\begin{aligned} & \sum_{m=1}^{T_{2n}+1} I(c_0|Z(m)) + d \sum_{m=1}^{T_{2n}+1} [W(m) - \Psi_{\delta}(Z(m), \delta_0)] + \sum_{m=1}^{T_{2n}+1} \{d[\Psi_{\delta}(Z(m), \delta_0) - \Psi(Z(m), c_0)] - \varepsilon\} \\ & \leq \log \alpha - \log H(c_0 - \varepsilon, c_0 + \varepsilon) + d[W(T_{2n} + 1) - \Psi_{\delta}(Z(T_{2n} + 1), \delta_0)] \\ & \quad + d[\Psi_{\delta}(Z(T_{2n} + 1), \delta_0) - \Psi(Z(T_{2n} + 1), c_0)]. \end{aligned}$$

Following an argument analogous to the above, we have $E_{\theta}[T_2] < \infty$ and so $P_{\theta}\{T_2 < \infty\} = 1$.

Proof of Lemma 3.2. Note that $\delta_{12} = 0$ since $\theta_1 = \theta_2$. Now,

$$L_{12}(n, \delta^*) = \int_0^{\delta^*} \exp \left\{ (y + \delta^*) \sum_{m=1}^n W(m) + \sum_{m=1}^n \lambda(m, \delta^*, y) \right\} dH(y).$$

By Jensen's inequality, we have

$$n^{-1} \log L_{12}(n, \delta^*) \geq n^{-1} \sum_{m=1}^n V_m, \quad (\text{A.7})$$

where $V_m = \int_0^{\delta^*} \{(y + \delta^*)W(m) + \lambda(m, \delta^*, y)\}dH(y)$, $m = 1, 2, \dots, n$, are iid. By Lemma 2.1 and the fact that $\delta_{12} = 0$, then,

$$\begin{aligned} E_{\theta}[V_m] &= \int_0^{\delta^*} E_{\theta}[\Psi(Z(m), \delta^*) - \Psi(Z(m), y)]dH(y) \\ &\geq \int_0^{c_0} E_{\theta}[\Psi(Z(m), \delta^*) - \Psi(Z(m), y)]dH(y) \\ &\geq \int_0^{c_0} E_{\theta}[\Psi(Z(m), \delta^*) - \Psi(Z(m), c_0)]dH(y) \\ &> 0. \end{aligned} \quad (\text{A.8})$$

By strong law of large numbers and by (A.7) and (A.8), we obtain $\liminf_{n \rightarrow \infty} n^{-1} \log L_{12}(n, \delta^*) \geq \int_0^{c_0} E_{\theta}[\Psi(Z(m), \delta^*) - \Psi(Z(m), c_0)]dH(y)$ a.e. while $n^{-1} \log \alpha \rightarrow 0$ as $n \rightarrow \infty$. Therefore, $P_{\theta}\{T_{12} < \infty\} = 1$ when $\theta_1 = \theta_2$. Similarly $P_{\theta}\{T_{21} < \infty\} = 1$ and thus $P_{\theta}\{T_1 < \infty\} = 1$ when $\theta_1 = \theta_2$.

References

- Bechhofer, R. E., Kiefer, J. and Sobel, M. (1968). *Sequential Identification and Ranking Procedures*. University of Chicago Press, Chicago.
- Desu, M. N. (1970). A selection problem. *Ann. Math. Statist.* **41**, 1596–1603.
- Edwards, D. (1987). Extended–Paulson sequential selection. *Ann. Statist.* **15**, 449–455.

- Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics* **7**, 225–245.
- Gupta, S. S. and Santner, T. J. (1973). On selection and ranking procedures – a restricted subset selection rule. *Proceedings of the 39th Session of the International Statistical Institute*, Vienna, Austria, **1**, 409–417.
- Hoel, D. G. and Mazumdar, M. (1968). An extension of Paulson's selection procedure. *Ann. Math. Statist.* **9**, 1034–1039.
- Hsu, J. C. and Edwards, D. G. (1983). Sequential multiple comparisons with the best. *J. Amer. Statist. Assoc.* **78**, 958–964.
- Kao, S. C. and Lai, T. L. (1980). Sequential selection procedures based on confidence sequences for normal populations. *Commun. Statist. –Theor. Meth.* **A9(16)** 1657–1676.
- Lam, K. (1986). A new procedure for selecting good populations. *Biometrika* **73**, 201–206.
- Perng, S. K. and Groves, B. D. (1977). Two sequential procedures for partitioning a set of one-parameter exponential populations with respect to a control. *Austral. J. Statist.* **19**, 56–60.
- Pollak, M. and Siegmund, D. O. (1975). Approximations to the expected sample size of certain sequential tests. *Ann. Statist.* **3**, 1267–1282.
- Robbins, H. and Siegmund, D. O. (1973). A class of stopping rules for testing parametric hypotheses. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **4**, 37–41, University of California Press.
- Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problems. *Ann. Statist.* **3**, 334–349.
- Santner, T. J. (1976). A two-stage procedure for selecting δ^* -optimal means in the normal model. *Commun. Statist. –Theor. Meth.* **A5(3)**, 283–292.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS			
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release, distribution unlimited.			
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE						
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report #88-20C			5. MONITORING ORGANIZATION REPORT NUMBER(S)			
6a. NAME OF PERFORMING ORGANIZATION Purdue University		6b. OFFICE SYMBOL (if applicable)		7a. NAME OF MONITORING ORGANIZATION		
6c. ADDRESS (City, State, and ZIP Code) Department of Statistics West Lafayette, IN 47907			7b. ADDRESS (City, State, and ZIP Code)			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research		8b. OFFICE SYMBOL (if applicable)		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-88-K-0170, DMS-8606964		
8c. ADDRESS (City, State, and ZIP Code) Arlington, VA 22217-5000			10. SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) On A Sequential Subset Selection Procedure For Exponential Family Distributions (Unclassified)						
12. PERSONAL AUTHOR(S) TaChen Liang						
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) May 1988		15. PAGE COUNT 18
16. SUPPLEMENTARY NOTATION						
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)			
FIELD	GROUP	SUB-GROUP	Bayes rule; Empirical Bayes; Asymptotically optimal; Rate of Convergence; Monotone decision problem; Least-concave majorant			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This paper deals with the problem of selecting the best population among k populations belonging to the same class of exponential family distributions through sequential subset selection approach. We desire that the best population should be selected and each selected population should be good. Based on the modified likelihood ratio of the conditional frequency function of some statistics, an elimination-type sequential subset selection procedure is proposed. This sequential procedure achieves the selection goal with guaranteed probability at least P* for some prespecified value P*. At each stage, this procedure also provides some statistical inference about an upper bound on the measure of separation from the unknown best population to each remaining contending population. Finally, a modified sequential procedure to select a good population is also studied. This modified sequential procedure achieves the goal of selecting a good population with guaranteed probability at least P*.						
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS				21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL TaChen Liang			22b. TELEPHONE (Include Area Code) (317) 494-6030		22c. OFFICE SYMBOL	