<Research Paper>

Journal of the Korean Statistical Society
Vol. 19, No. 1, 1990

A Two-Stage Elimination Type Selection Procedure for Stochastically Increasing Distributions: with an Application to Scale Parameters Problem+

Seung-Ho Lee*

ABSTRACT

The purpose of this paper is to extend the idea of Tamhane and Bechhofer(1977, 1979) concerning the normal means problem to some general class of distirbutions. The key idea in Tamhane and Bechhofer is the derivation of the computable lower bounds on the probability of a correct selection. To derive such lower bounds, they used the specific covariance structure of a multivariate normal distribution. It is shown that such lower bounds can be obtained for a class of stochastically increasing distributions under certain conditions, which is sufficiently general so as to include the normal means problem as a special application. As an application of the general theory to the scale parameters problem, a two-stage elimination type procedure for selecting the population associated with the smallest variance from among several normal populations is proposed. The design constants are tabulated and the relative efficiencies are computed.

1. Introduction

If k populations $\pi_1, \pi_2, \cdots, \pi_k$ are given and we wish to decide on the basis of a properly chosen sampling scheme which one of these populations is the best one, various approaches and methods have been studied up to now. A more detailed overview is provided by Gupta and Panchapakesan(1979). Among those, two-stage procedures with screening in the first

---

+ Research supported in part by the Korea Science & Engineering Foundation.
This research was done while the author was visiting the Department of Statistics at Purdue University.

* Department of Mathematics, Ajou University, Suwon 441-749, Korea.

stage seem to be quite appropriate, since they are more economical than single stage procedures but still technically not so complicated as sequential ones.

Cohen(1959) was the first to combine Gupta's(1956) maximum mean procedure in the first stage and Bechhofer's(1954) natural decision procedure in the second stage. Later, Alam(1970) proposed a minimax criterion in determining design constants of such two-stage procedures. But these results were mostly confined to the special case of $k=2$ normal populations with a common known variance.

Tamhane and Bechhofer(1977, 1979) extended Alam's(1970) work to the general case of $k \geq 2$ populations with some optimization criteria. They studied in detail a two-stage elimination type procedure using a u-minimax design criterion, and the two-stage procedure was found to be more efficient than the single-stage procedure of Bechhofer(1954). It should be noted that their work was also restricted to the normal means problem. For the normal means problem with a common unknown variance, it is well known that there does not exist any single-stage procedure satisfying the required minimum probability of correct selection. Bechhofer, Dunnet and Sobel(1954) were the first to use Stein's(1945) idea in devising two-stage selection procedures. Unlike the case with a known variance, the unknown variance is estimated at the first stage and the sample best in selected as the true best in the second stage. Later, Tamhane(1976), and Hochberg and Marcus(1981) considered three-stage procedures with the second stage set for elimination. Gupta and Kim(1984) proposed a two-stage procedure, in which the unknown variance is estimated and the bad ones are eliminated at the first stage.

The purpose of this paper is to extend the idea of Tamhane and Bechhofer(1977, 1979) concerning the normal means problem to some general class of distributions and to illustrate the extended theory by using some specific examples. The key idea in Tamhane and Bechhofer is the derivation of computable lower bounds on the probability of correct selection over the preference zone. To derive such lower bounds, they used the specific covariance structure of a multivariate normal distribution which heavily depends on the normality assumption. However, it is found that such lower bounds can be obtained for a class of stochastically increasing distributions under certain conditions, which is sufficiently general so as to include the normal means problem as a special application.

In Section 2, the formulation of the problem is given. A two-stage elimination type procedure for selecting the largest parameter value and a design criterion following the lines of Tamhane and Bechhofer(1977, 1979) are described. The main analytical results are contained in Sections 3 which deals with the probability of a correct selection and the expected total sample size. As an application of the general theory to the scale parameters problem, the problem of selecting the population associated with the smallest variance from among several normal populations is treated in Section 4. The design constants are tabulated and the reletive efficiencies of the two-stage procedures with respect to the corresponding single-stage procedure are computed.

2. A Two-Stage Procedure and Its Design Criterion

Let $\pi_i(1\le i\le k)$ be k populations, where the probability distribution of π_i depends only on an unknown parameter θ_i in an interval Θ of the real line$(1\le i\le k)$. Let $\theta_{[1]}\le\theta_{[2]}\le\cdots\le\theta_{[k]}$ denote the ordered values of the unknown parameters $\theta_1, \theta_2,\cdots,\theta_k$. We assume that the correct pairing between θ_i and $\theta_{[i]}$ is unknown. Any population associated with the largest parameter value $\theta_{[k]}$ is called the "best" population.

Following Santner(1975), an indifference-zone will be defined in the entire parameter space $\Omega=\{\underline{\theta}=(\theta_1, \theta_2,\cdots,\theta_k) \mid \theta_i\in\Theta, 1\le i\le k\}$ by means of a real valued function δ on Θ having the following properties:

(i) $\delta(\theta)<\theta$ for all $\theta\in\Theta$

(ii) δ, restricted on Θ', is a function onto Θ where $\Theta'=\{\theta\in\Theta \mid \delta(\theta)\in\Theta\}$.

Define the so-called preference-zone by

$$\Omega(\delta)=\{\underline{\theta}\in\Omega \mid \theta_{[k-1]}\le\delta(\theta_{[k]})\}, \tag{2.1}$$

where the best and the second best are sufficiently far apart so that the experimenter desires to insure the detection of the best with high probability. The complement of $\Omega(\delta)$ is called the indifference-zone. The following preference-zones have been used in the literatures of selection and ranking.

Example 2.1(a) A location type preference-zone defined by $\delta_1(\theta)=\theta-\delta^*(\delta^*>0)$ is given by

$$\Omega(\delta^*)=\{\underline{\theta} \mid \theta_{[k]}-\theta_{[k-1]}\ge\delta^*\}. \tag{2.2}$$

(b) A scale type preference-zone defined by $\delta_2(\theta)=\theta/\delta^*(\delta^*>1)$ is given by

$$\Omega(\delta^*)=\{\underline{\theta} \mid \theta_{[k]}\ge\delta^*\theta_{[k-1]}\}. \tag{2.3}$$

The goal of the experimenter is to select the best population. The event of correctly selecting the best population is denoted by CS. Following the indifference-zone approach, the attention is restricted to selection procedures R which guarantee the basic probability requirement;

$$P_{\underline{\theta}}\{CS \mid R\} \geq P^* \text{ for all } \underline{\theta} \in \Omega(\delta), \tag{2.4}$$

where $P^*(1/k < P^* < 1)$ is specified prior to the experiment.

A selection procedure with its design criterion

We now describe a two-stage elimination type selection procedure and its design criterion. At the first stage, the noncontending populations will be screened out using the statistics $T_i^{(1)} = T^{(1)}(X_{i,1}, \cdots, X_{i,n_1})(1 \leq i \leq k)$ based on n_1 independent observations $X_{i,1}, \cdots, X_{i,n_1}$ from each of $\pi_i(1 \leq i \leq k)$. At the second stage, we compute the statistics $T_i^{(2)} = T^{(2)}(X_{i,n_1+1}, \cdots, X_{i,n_1+n_2})$ based on n_2 additional independent observations from each of the retained populations, and selection is made using the statistics $T_i = u(T_i^{(1)}, T_i^{(2)})$ based on the overall sample with an appropriate function u.

Here, the screening process will be done using the following Gupta-type procedure:

Retain π_i if and only if $h(T_i^{(1)}) \geq \max_{1 \leq j \leq k} T_j^{(1)}$ where $h(\cdot)$ is a real valued function such that $h(x) > x$ for each x and h(x) is continuous and strictly increasing in x.

Typical examples of $h(\cdot)$ are given by $h(x) = x + d(d > 0)$ and $h(x) = cx(c > 1)$ for location type and scale type procedures, respectively.

Now the precise definition of a two-stage elimination type procedure R is given as follows.

<u>Stage 1.</u> Take n_1 independent observations $X_{i,1}, \cdots, X_{i,n_1}$ from each $\pi_i(1 \leq i \leq k)$ and compute $T_i^{(1)} = T^{(1)}(X_{i,1}, \cdots, X_{i,n_1})$. Define an index set I by

$$I = \{i \mid h(T_i^{(1)}) \geq \max_{1 \leq j \leq k} T_j^{(1)}, 1 \leq i \leq k\} \tag{2.5}$$

and let $|I|$ denote the number of elements in I.

(a) If $|I| = 1$, assert that the population associated with $\max_{1 \leq j \leq k} T_j^{(1)}$ is the best.

(b) If $|I| \geq 2$, proceed to the second stage.

Stage 2. Take n_2 additional observations $X_{i,n_2+1},\cdots,X_{i,n_1+n_2}$ from each population π_i, $i\in I$, and compute $T_i=u(T_i^{(1)}, T_i^{(2)})$ where $T_i^{(2)}=T^{(2)}(X_{i,n_1+1},\cdots,X_{i,n_1+n_2})$. We then claim that the population associated with $\max_{i\in I} T_i$ is the best.

In the definition of the above two-stage procedure, the sample sizes n_1, n_2 and the functin $h(\cdot)$ will be chosen so that the procedure guarantees the basic probability requirement (2.4) and different design criteria lead to different choices. We adopt the following unrestricted minimax criterion:

$$\begin{aligned} &\text{Minimize} && \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(TSS \mid R) \\ &\text{subject to} && \inf_{\underline{\theta}\in\Omega(\delta)} P_{\underline{\theta}}(CS \mid R)\geq P^*, \end{aligned} \tag{2.6}$$

where TSS is the total sample size needed in the experiment.

3. Lower Bounds on the Probability of a Correct Selection and Expected Total Sample Size

A main problem concerned with the construction of selection procedures using the indifference-zone approach is to find the infimum of the probability of a correct selection over the preference-zone $\Omega(\delta)$. Any parameter configuration achieving such an infimum is called a least favorable configuration (LFC) for the procedure under study.

However, as can be seen from Alam(1970), Tamhane and Bechhofer(1977, 1979), Miescke and Sehr(1980) and Gupta and Miescke(1982), there has been a conjecture that the LFC for the elimination type two-stage procedure would be the slippage one. Thus, instead of trying to find the LFC, some lower bounds will be derived here as in Tamhane and Bechhofer(1977, 1979). To do so, the following assumptions are made regarding the statistics $T_i^{(1)}$ and $T_i=u(T_i^{(1)}, T_i^{(2)})$ used in the procedure R.

Assumption (A1). The distributions of $T_i^{(1)}$ and $T_i^{(2)}$ are stochastically increasing in $\theta_i\in\Theta$ for $i=1, 2,\cdots, k$.

Assumption (A2). The function $u(t_1, t_2)$, used to define the statistic $T_i=u(T_i^{(1)}, T_i^{(2)})$, is strictly increasing in each variable.

In the sequel, let $F(\cdot \mid \theta_i)$ and $G(\cdot \mid \theta_i)$ denote the cdf's of $T_i^{(1)}$ and T_i, respectively and let $H(t_1, t_2 \mid \theta_i)$ denote the joint cdf of $T_i^{(1)}$ and T_i. It follows from the assumptions (A1) and (A2) that $F(\cdot \mid \theta_i)$, $G(\cdot \mid \theta_i)$ and $H(\cdot, \cdot \mid \theta_i)$ are non-increasing in $\theta_i (1 \leq i \leq k)$. From this fact the following results can be obtained.

Lemma 3.1 Under the assumptions (A1) and (A2), the following inequality holds.

$$\inf_{\underline{\theta} \in \Omega(\delta)} P_{\underline{\theta}}(CS \mid R) \geq \inf_{\theta \in \Theta'} A(\theta), \tag{3.1}$$

where

$$A(\theta) = E_\theta [H^{k-1}(h(T_k^{(1)}), T_k \mid \delta(\theta))]. \tag{3.2}$$

Proof. Without loss of generality, we may assume that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$. Then, for all $\underline{\theta} \in \Omega(\delta)$,

$$\begin{aligned} P_{\underline{\theta}}(CS \mid R) &= P_{\underline{\theta}}\{h(T_k^{(1)}) \geq \max_{1 \leq j \leq k} T_j^{(1)},\ T_k = \max_{i \in I} T_i\} \\ &\geq P_{\underline{\theta}}\{h(T_k^{(1)}) \geq T_i^{(1)},\ T_k \geq T_i \text{ for all } i=1,\cdots,k-1\} \\ &= \int \prod_{i=1}^{k-1} H(h(x), y \mid \theta_i)\, dH(x, y \mid \theta_k) \end{aligned} \tag{3.3}$$

Thus, the result (3.1) follows from (3.3) and the facts that $H(\cdot, \cdot \mid \theta_i)$ is non-increasing in θ_i and $\theta_i \leq \delta(\theta_k)$ for all $i=1,\cdots,k-1$, whenever $\underline{\theta} \in \Omega(\delta)$.

However, it would be very difficult to compute $A(\theta)$ in (3.1) in practice due to the dependence between $T_k^{(1)}$ and T_k. Thus it seems reasonable to find a lower bound for $A(\theta)$, which can be easily computed even though it is slightly less sharp. Such a lower bound can be obtained by the following result.

Lemma 3.2 Suppose that assumptions (A1) and (A2) hold. Then, for all $\theta \in \Theta'$,

$$A(\theta) \geq E_\theta \left[F^{k-1}(h(T_k^{(1)}) \mid \delta(\theta))\right] E_\theta \left[G^{k-1}(T_k \mid \delta(\theta))\right]. \tag{3.4}$$

Proof. The assumption (A2) insures that, for each fixed b, there exists a function $v(\cdot, b)$ such that

$$u(T_i^{(1)}, T_i^{(2)}) \leq b \text{ if and only if } T_i^{(1)} \leq v(T_i^{(2)}, b).$$

Thus, for each a and b,

$$
\begin{aligned}
&P_\theta\{T_i^{(1)} \le a,\ T_i = u(T_i^{(1)}, T_i^{(2)}) \le b\} \\
&= E_\theta[P_\theta\{T_i^{(1)} \le a,\ T_i^{(1)} \le v(T_i^{(2)}, b) \mid T_i^{(2)}\}] \\
&\ge E_\theta[P_\theta\{T_i^{(1)} \le a \mid T_i^{(2)}\}\ P_\theta\{T_i^{(1)} \le v(T_i^{(2)}, b) \mid T_i^{(2)}\}] \\
&= P_\theta\{T_i^{(1)} \le a\}\ P_\theta\{T_i \le b\}
\end{aligned}
$$

which in turn implies that

$$
\begin{aligned}
&E_\theta[H^{k-1}(h(T_k^{(1)}), T_k \mid \delta(\theta))] \\
&\ge E_\theta[F^{k-1}(h(T_k^{(1)}) \mid \delta(\theta))\ G^{k-1}(T_k \mid \delta(\theta))].
\end{aligned}
$$

Since $F(h(T_k^{(1)}) \mid \delta(\theta))$ and $G(T_k \mid \delta(\theta)) = G(u(T_k^{(1)}, T_k^{(2)}) \mid \delta(\theta))$ are non-decreasing in $T_k^{(1)}$,

$$
\begin{aligned}
&E_\theta[F^{k-1}(h(T_k^{(1)}) \mid \delta(\theta))\ G^{k-1}(T_k \mid \delta(\theta))] \\
&\ge E_\theta[F^{k-1}(h(T_k^{(1)}) \mid \delta(\theta))]\ E_\theta[G^{k-1}(T_k \mid \delta(\theta))]
\end{aligned}
$$

by the Chebychev's inequality. This completes the proof.

We summarize Lemmas 3.1 and 3.2 into the following theorem.

Theorem 3.1 Under the assumptions (A1) and (A2), the follwoing inequalities hold.

$$
\inf_{\underline{\theta} \in \Omega(\delta)} P_\theta(CS \mid R) \ge \inf_{\theta \in \Theta'} A(\theta) \ge \inf_{\theta \in \Theta'} B(\theta)\,, \tag{3.5}
$$

where $A(\theta)$ is given by (3.2) and $B(\theta)$ denotes the right hand side of (3.4).

Finally, it should be pointed out that any further simplification of the lower bounds $A(\theta)$ and $B(\theta)$ can not be done without further assumptions on the structure of the statistical model under study. The situation becomes quite simpler as can be seen in the following examples.

Example 3.1 (Location parameters problem). Suppose that θ_i is a location parameter of the population $\pi_i (1 \le i \le k)$ with the preference-zone given by (2.2) in Example 2.1 (a). Suppose further that θ_i is also a location parameter of the distributions of $T_i^{(1)}$ and T_i. Then, for the location type screening procedure with $h(x) = x + d$, $A(\theta)$ and $B(\theta)$ do not depend on the parameter θ.

In fact,

$$
A(\theta) = A(\delta^*) = E_{\theta=0}[H^{k-1}(T_k^{(1)} + d + \delta^*,\ T_k + \delta^*)] \tag{3.6}
$$

and

$$
B(\theta) = B(\delta^*) = E_{\theta=0}[F^{k-1}(T_k^{(1)} + d + \delta^*)]\ E_{\theta=0}[G^{k-1}(T_k + \delta^*)]\,, \tag{3.7}
$$

where F, G and H denote the cdf's of $T_i^{(1)}$, T_i and the joint cdf of $(T_i^{(1)}, T_i)$ when $\theta_i=0$, respectively.

Remark. As a typical application to the location parameters problem consider the normal populations π_i's with unknown means θ_i's and a common known variance $\sigma^2 (1\leq i\leq k)$. Define the two-stage procedure by setting

$$T_i^{(1)}=\bar{X}_i^{(1)}=\sum_{j=1}^{n_1} X_{i,j}/n_1, \qquad T_i^{(2)}=\bar{X}_i^{(2)}=\sum_{j=n_1+1}^{n_1+n_2} X_{i,j}/n_2,$$
$$T_2=u(T_i^{(1)}, T_i^{(2)})=(n_1T_i^{(1)}+n_2T_i^{(2)})/(n_1+n_2)=\bar{X}_i$$

and $h(t)=t+h(h>0)$. This is exactly the procedure of Tamhane and Bechhofer(1977, 1979). Clearly, the assumptions (A1) and (A2) hold in this case.

Also, from the definition of the statistics $T_i^{(1)}$, $T_i^{(2)}$ and T_i, the corresponding cdf's are given as follows:

$$F(t_1 \mid \delta(\theta)=P_{\delta(\theta)}\{\bar{X}_i^{(1)}\leq t_1\}=\Phi\{\sqrt{n_1}(t_1-\theta+\delta^*)/\sigma\}$$
$$G(t_2 \mid \delta(\theta))=P_{\delta(\theta)}\{\bar{X}_i\leq t_2\}=\Phi\{\sqrt{n_1+n_2}(t_2-\theta+\delta^*)/\sigma\}$$

and

$$H(t_1, t_2 \mid \delta(\theta))=P_{\delta(\theta)}\{\bar{X}_i^{(1)}\leq t_1, \bar{X}_i\leq t_2\}$$
$$=\Phi_2\{\sqrt{n_1}(t_1-\theta+\delta^*)/\sigma, \sqrt{n_1+n_2}(t_2-\theta+\delta^*)/\sigma \mid \sqrt{n_1/(n_1+n_2)}\},$$

where Φ is the cdf of the standard normal distribution and $\Phi_2\{\cdot, \cdot \mid \rho\}$ denotes the cdf of the bivariate normal distribution with means 0, variances 1 and correlation ρ.

Therefore the lower bounds in Theorem 3.1 are given as follows:

$$A(\theta)=E[\Phi_2^{k-1}\{\sqrt{n_1}(\bar{X}_k^{(1)}-\theta+\delta^*+h)/\sigma, \sqrt{n_1+n_2}[\bar{X}_k-\theta+\delta^*)/\sigma \mid p\}]$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Phi_2^{k-1}\{\sqrt{n_1}(\delta^*+h)/\sigma+x, \sqrt{n_1+n_2}\ \delta^*/\sigma+y \mid p\}\ d\Phi_2(x, y \mid p)$$

and

$$B(\theta)=\int_{-\infty}^{\infty}\Phi^{k-1}\{x+\sqrt{n_1}(\delta^*+h)/\sigma\}\ d\Phi(x)\int_{-\infty}^{\infty}\Phi^{k-1}(y+\sqrt{n_1+n_2}\ \delta^*/\sigma\}\ d\Phi(y)$$

with $p=\sqrt{n_1/(n_1+n_2)}$. Two lower bounds $A(\theta)$ and $B(\theta)$ do not depend on the unknown θ, and they are exactly the bounds of Tamhane and Bechhofer(1979), in which the performance of the procedure based on the lower bounds was investigated. The results indicate that the procedures improve upon the single stage procedure of Bechhofer(1954), with the one based on $A(\theta)$ being slightly better than that based on $B(\theta)$. It may be noted that the lower bound $A(\theta)$ in this case can be handled without much difficulty because the integration involves only a bivariate normal distribution.

Example 3.2 (Scale Parameters Problem). Suppose that θ_i is a scale parameter of the population $\pi_i(1 \le i \le k)$. In this case, the preforence-zone can be given as that in Example 2.1 (b). Suppose further that θ_i is a scale parameter of the distribution of $T_i^{(1)}$ and T_i. Then, for the scale-type screening procedure with $h(x)=cx(c>1)$, $A(\theta)$ and $B(\theta)$ are given by

$$A(\theta)=A(\delta^*)=E_{\theta=1}[H^{k-1}(c\delta^* T_k^{(1)}, \delta^* T_k)]$$

and

$$B(\theta)=B(\delta^*)=E_{\theta=1}[F^{k-1}(c\delta^* T_k^{(1)})]\ E_{\theta=1}[G^{k-1}(\delta^* T_k)],$$

where F, G and H denote the cdf's of $T_i^{(1)}$, T_i and $(T_i^{(1)}, T_i)$, respectively when $\theta_i=1$.

In order to employ the u-minimax criterion in Section 2, it is necessary to know the set of parameter points in Ω at which the supremum of $E_{\underline{\theta}}(TSS \mid R)$ occurs. It is shown that the supremum is attained when $\theta_1=\theta_2=\cdots=\theta_k$, the equal parameter configuration(EPC).

First, we derive a general expression of the expected total sample size. Note that the total sample size(TSS) can be written as

$$TSS=kn_1+n_2S \tag{3.8}$$

where S is the number of populations to be sampled at the second stage, i.e., $S=0$ if $|I|=1$ and $S=|I|$ otherwise.

Since

$$E_{\underline{\theta}}(S \mid R)=E_{\underline{\theta}}(|I| \mid R)-P_{\underline{\theta}}\{|I|=1 \mid R\}$$

we have

$$E_{\underline{\theta}}(S \mid R)=\sum_{i=1}^{k}[P_{\underline{\theta}}\{h(T_i^{(1)}) \ge \max_{1\le j\le k} T_j^{(1)}\}-P_{\underline{\theta}}\{T_i^{(1)} \ge \max_{j\ne i} h(T_j^{(1)})\}]$$

$$=\sum_{i=1}^{k}[\int \prod_{j\ne i} F(h(x) \mid \theta_j)dF(x \mid \theta_i)-\int \prod_{j\ne i} F(h^{-1}(x) \mid \theta_j)dF(x \mid \theta_i)].$$

Thus, a general expression of $E_{\underline{\theta}}(TSS \mid R)$ is given by

$$E_{\underline{\theta}}(TSS \mid R)=kn_1+n_2\sum_{i=1}^{k}[\int \prod_{j\ne i} F(h(x) \mid \theta_j)dF(x \mid \theta_i) -\int \prod_{j\ne i} F(h^{-1}(x) \mid \theta_j)dF(x \mid \theta_i)] \tag{3.9}$$

In order to know the behavior of $E_{\underline{\theta}}(TSS \mid R)$ as a function of $\underline{\theta}$, we need the following regularity condition.

Regularity Condition (C1). For the cdf $F(x \mid \theta_i)$ of $T_i^{(1)}$, the partial derivatives $f(x \mid \theta) = \frac{\partial}{\partial x} F(x \mid \theta)$, $\dot{F}(x \mid \theta) = \frac{\partial}{\partial \theta} F(x \mid \theta)$ and $\frac{\partial}{\partial \theta} f(x \mid \theta)$ exist for all $\theta \in \Theta$, and for the function $h(\cdot)$, the derivatives $\dot{h}(x) = \frac{d}{dx} h(x)$ and $\dot{h}^{-1}(x) = \frac{d}{dx} h^{-1}(x)$ exist for all x.

We now state the main result concerning the suprmum of the expected total sample size in the following theorem.

Theorem 3.2 Suppose that the regularity conditon (C1) holds. Then, the supremum of $E_{\underline{\theta}}(TSS \mid R)$ is attained whenever $\theta_1 = \theta_2 = \cdots = \theta_k$, provided that

$$\dot{F}(h(x) \mid \theta_1)\, f(x \mid \theta_2) - \dot{h}(x)\, f(h(x) \mid \theta_2)\, \dot{F}(x \mid \theta_1) \geq 0 \tag{3.10}$$

and

$$\dot{F}(h^{-1}(x) \mid \theta_1)\, f(x \mid \theta_2) - \dot{h}^{-1}(x)\, f(h^{-1}(x) \mid \theta_2)\, \dot{F}(x \mid \theta_1) \leq 0 \tag{3.11}$$

for all $\theta_1 \leq \theta_2$ and all x. Thus,

$$\sup_{\underline{\theta} \in \Omega} E_{\underline{\theta}}(TSS \mid R) = kn_1 + kn_2 \sup_{\theta \in \Theta'} \left\{ \int F^{k-1}(h(x) \mid \theta)\, dF(x \mid \theta) - \int F^{k-1}(h^{-1}(x) \mid \theta)\, dF(x \mid \theta) \right\}. \tag{3.12}$$

Proof. Consider, along the lines of Gupta(1965), a parameter configuration $\theta_1 = \cdots = \theta_q (= \theta) \leq \theta_{q+1} \leq \cdots \leq \theta_k$. Then, under the regularity condition, it is easily seen that

$$\begin{aligned}
&\frac{\partial}{\partial \theta} E_{\underline{\theta}}(S \mid R) \\
&= \int q(q-1)F^{q-2}(h(x) \mid \theta) \{\dot{F}(h(x) \mid \theta) f(x \mid \theta) - \dot{h}(x) f(h(x) \mid \theta) \dot{F}(x \mid \theta)\} \prod_{j=q+1}^{k} F(h(x) \mid \theta_j)dx \\
&\quad + \sum_{i=q+1}^{k} \int qF^{q-1}(h(x) \mid \theta) \{\dot{F}(h(x) \mid \theta) f(x \mid \theta_i) - \dot{h}(x) f(h(x) \mid \theta_i) \dot{F}(x \mid \theta)\} \prod_{\substack{j=q+1 \\ j \neq i}}^{k} F(h(x) \mid \theta_j)\, dx \\
&\quad - \int q(q-1)F^{q-2}(h^{-1}(x) \mid \theta) \{\dot{F}(h^{-1}(x) \mid \theta) f(x \mid \theta) - \dot{h}^{-1}(x) f(h^{-1}(x) \mid \theta) \dot{F}(x \mid \theta)\} \\
&\quad \times \prod_{j=q+1}^{k} F(h^{-1}(x) \mid \theta_j)dx \\
&\quad - \sum_{i=q+1}^{k} \int qF^{q-1}(h^{-1}(x) \mid \theta) \{\dot{F}(h^{-1}(x) \mid \theta) f(x \mid \theta_i) - \dot{h}^{-1}(x) f(h^{-1}(x) \mid \theta_i)\, \dot{F}(x \mid \theta)\} \\
&\quad \times \prod_{\substack{j=q+1 \\ j \neq i}}^{k} F(h^{-1}(x) \mid \theta_j)dx.
\end{aligned}$$

Thus, (3.10) and (3.11) imply $\frac{\partial}{\partial \theta} E_{\underline{\theta}}(S \mid R) \geq 0$ and hence the supremum of $E_{\underline{\theta}}(S \mid R)$ is attained whenever $\theta_1 = \theta_2 = \cdots = \theta_k$. Therefore the results follows from (3.8) and (3.9).

It should be remarked that the conditions (3.10) and (3.11) are reduced to the monotone likelihood ratio property of the density $f(x \mid \theta)$ of $F(x \mid \theta)$ in the location or scale parameters problem in the framework of Examples 3.1 and 3.2.

4. An Application to Normal Variances Problem

For a dual problem of selecting the population associated with the smallest parameter value $\theta_{[1]}$, the preference-zone is given by

$$\Omega(\delta)=\{\underline{\theta}\in\Omega \mid \theta_{[2]}\geq\delta(\theta_{[1]})\},$$

where the real-valued function $\delta(\cdot)$ satisfies $\delta(\theta)>\theta$ for all $\theta\in\Theta$.

Then, with some modification for final decision rules, a two-stage procedure R′ for this problem can be constructed in exactly the same manner as that in Section 2 except that the screening procedure is replaced by;

include π_i in the retained populations if and only if

$$h(T_i^{(1)})\leq\min_{1\leq j\leq k} T_j^{(1)}$$

where $h(\cdot)$ satisfies $h(x)<x$ for all x.

For the procedure R′, the following results can be obtained with some slight modifications of the arguments for the procedure R in Section 3.

Theorem 4.1 Under the assumptions (A1) and (A2) in Section 3, the following inequalities hold.

$$\inf_{\underline{\theta}\in\Omega(\delta)} P_{\underline{\theta}}(CS \mid R')\geq\inf_{\theta\in\Theta'} A'(\theta)\geq\inf_{\theta\in\Theta'} B'(\theta) \tag{4.1}$$

where $A'(\theta)$ and $B'(\theta)$ are defined by

$$A'(\theta)=E_\theta[M^{k-1}(h(T_1^{(1)}), T_1 \mid \delta(\theta))] \tag{4.2}$$

and

$$B'(\theta)=E_\theta[\{1-F(h(T_1^{(1)}) \mid \delta(\theta))\}^{k-1}]\ E_\theta[\{1-G(T_1 \mid \delta(\theta))\}^{k-1}] \tag{4.3}$$

with

$$M(x, y \mid \theta_i)=P_{\theta_i}[T_i^{(1)}>x,\ T_i>y],\ 1\leq i\leq k.$$

Theorem 4.2 Suppose that the regularity condition (C1) in Section 3 holds. Then the supremum of $E_{\underline{\theta}}(TSS \mid R')$ is attained whenever $\theta_1 = \theta_2 = \cdots = \theta_k$ provided (3.10) and (3.11) are satisfied. Thus,

$$\sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(TSS \mid R') = kn_1 + kn_2 \sup_{\theta\in\Theta'} \Big[\int \{1-F(h(x) \mid \theta)\}^{k-1}\, dF(x \mid \theta) - \int \{1-F(h^{-1}(x) \mid \theta)\}^{k-1}\, dF(x \mid \theta)\Big] \tag{4.4}$$

Remark. It is easily seen that the characterization of the procedures R and R′ depend on the parameters only through ststistics $T_i^{(1)}$ and $T_i(1 \le i \le k)$. Thus, the results obtained so far remain valid as long as the distributions of $T_i^{(1)}$ and T_i do not depend on the nuisance parameters.

For the problem of selecting the population associated with smallest variance from among several normal populations, Bechhofer and Sobel(1954) proposed a single-stage procedure R_0, in the framework of indifference-zone approach.

Gupta and Sobel(1962a, 1962b) investigated the same problem under the framework of subset selection. The values of the sample sizes needed in the single-stage procedure R_0 of Bechhofer and Sobel(1954) can also be obtained from the tables of Gupta and Sobel(1962a). Extended tables are also available from Gibbons, Olkin and Sobel(1977). Later Tamhane(1975) formulated this problem in the two-stage sampling scheme with screening in the first stage and proposed a lower bound on the probability of correct selection. However, due to the computational difficulties involved, no tables were given.

Let $\pi_i(1 \le i \le k)$ denote k normal populations with unknown means $\mu_i(-\infty < \mu_i < \infty, 1 \le i \le k)$ and unknown variances. The ordered variances are denoted by $\sigma^2_{[1]} \le \sigma^2_{[2]} \le \cdots \le \sigma^2_{[k]}$. It is assumed that there is no prior information available about the correct pairing between π_i and $\sigma^2_{[i]}(1 \le i \le k)$. The goal is to select a population associated with $\sigma^2_{[1]}$.

It can be easily shown that this problem falls into the general framework with $\delta(\sigma^2) = \sigma^2/\delta^*(0 < \delta^* < 1)$, while $\mu_1, \cdots, \mu_k$ are the nuisance parameters.

Let

$$T_i^{(1)} = \sum_{j=1}^{n_1} (X_{i,j} - \bar{X}_i^{(1)})^2$$

$$T_i^{(2)} = \sum_{j=n_1+1}^{n_1+n_2} (X_{i,j} - \bar{X}_i^{(2)})^2$$

$$T_i = u(T_i^{(1)}, T_i^{(2)}) = T_i^{(1)} + T_i^{(2)}$$

and

$$h(t) = ct \; (0 < c < 1)$$

where

$$\bar{X}_i^{(1)}=\sum_{j=1}^{n_1} X_{i,j}/n_1, \text{ and } \bar{X}_i^{(2)}=\sum_{j=n_1+1}^{n_1+n_2} X_{i,j}/n_2.$$

An elimination type two-stage selection procedure R_1 is proposed as follows:

Stage 1. Take n_1 independent observations $X_{i,1},\cdots,X_{i,n_1}$ from each $\pi_i(1\le i\le k)$, compute $T_i^{(1)}$ and determine an index set I of $\{1, 2,\cdots,k\}$ where

$$I=\{i \mid cT_i^{(1)}\le \min_{1\le j\le k} T_j^{(1)}\},\ 0<c<1 \tag{4.5}$$

(a) If $|I|=1$, assert that the population associated with $\min_{1\le j\le k} T_j^{(1)}$ is the best.

(b) If $|I|\ge 2$, proceed to the second stage.

Stage 2. Take n_2 additional independent observations $X_{i,n_1+1},\cdots,X_{i,n_1+n_2}$ from each populations π_i, $i\in I$, compute $T_i=T_i^{(1)}+T_i^{(2)}$ and assert that the population associated with $\min_{j\in I} T_j$ is the best.

Note that $T_i^{(1)}/\sigma_i^2$, $T_i^{(2)}/\sigma_i^2$ and T_i/σ_i^2 have the chi-squared distributions with $\nu_1=n_1-1$, $\nu_2=n_2-1$ and $\nu=\nu_1+\nu_2=n_1+n_2-2$ degrees of freedom, respectively. However the joint distribution of $T_i^{(1)}$ and T_i is rather complicated and inconvenient to compute in this case. Thus we use the lower bound $B'(\theta)$ in Theorem 4.1 to determine the design constants (n_1, n_2, c) for the two-stage procedure R_1. By straightforward computation,

$$B'(\delta^*)=\int_0^\infty \{1-F_{\nu_1}(c\delta^* x)\}^{k-1}\, dF_{\nu_1}(x) \int_0^\infty \{1-F_\nu(\delta^* y)\}^{k-1}\, dF_\nu(y), \tag{4.6}$$

where $F_\nu(\cdot)$ denotes the cdf of chi-squared distribution having ν degrees of freedom.

Remark. Tamhane(1975) proposed almost the same procedure as the procedure R_1. The only difference is the statistic T_i used in Stage 2. His T_i is defined by

$$T_i=\sum_{j=1}^{n_1}(X_{i,j}-\bar{X}_i)^2+\sum_{j=n_1+1}^{n_1+n_2}(X_{i,j}-\bar{X}_i)^2$$

where $\bar{X}_i=\sum_{j=1}^{n_1+n_2} X_{i,j}/(n_1+n_2)$. Hence the degrees of freedom of T_i is $\nu=n_1+n_2-1$. When the population mean $\mu_i(1\le i\le k)$ are all known, with the obvious definitions of the statistics, the two procedures are exactly the same. He also derived a lower bound $C(\delta^*)$, say, on the probability of a correct selection, of the form

$$C(\delta^*)=\int_0^\infty \{1-F_{\nu_1}(c\delta^* x)\}^{k-1}\, dF_{\nu_1}(x)+\int_0^\infty \{1-F_\nu(\delta^* x)\}^{k-1}\, dF_\nu(x)-1. \tag{4.7}$$

For the same ν_1 and ν(this is the case when all μ_i's are known), $B'(\delta^*)\geq C(\delta^*)$ since $ab\geq a+b-1$ for $a, b\in(0, 1)$, and hence $B'(\delta^*)$ is a less conservative lower bound.

The supremum of the expected total sample size can be obtained from Theorem 4.2 and is given as follows.

$$\sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(TSS \mid R_1) = kn_1 + kn_2\left[\int_0^\infty \{1-F_{\nu_1}(cx)\}^{k-1}\, dF_{\nu_1}(x) - \int_0^\infty \{1-F_{\nu_1}(x/c)\}^{k-1}\, dF_{\nu_1}(x)\right]. \quad (4.8)$$

Therefore the corresponding optimization problem to determine the design constants (n_1, n_2, c) is to minimize (4.8) subject to $B'(\delta^*)\geq P^*$. This is an extremely complicated integer programming problem with a non-linear objective function.

To solve the optimization problem, we have treated n_1 and n_2 as continuous variables, and used the SUMT(Sequential Unconstrained Minimization Technique) algorithm of Fiacco and McCormick(1968). We denote by ($\hat{n}_1$, $\hat{n}_2$, $\hat{c}$) a solution to this continuous version of the optimization problem. The problem has been solved numerically for $k=2(1)\,10$, $P^*=0.90$, 0.95 and $\sqrt{\delta^*}=0.50(0.05)\ 0.70$. The results are given in Table 1.

In supplying the objective function (4.8) and the constraint function (4.6) to SUMT algorithm, we used the 32-point Laguerre numerical quadrature formula to evaluate the integrals, and the values of the chi-squared cdf's were evaluated using the 32-points Legendre numerical quadrature formula. All computations were carried out in double precision arithmatic on VAX-11/780 at the Department of Statistics of Purdue University, indiana, U.S.A. Throughout the computations the convergence criterion was fixed to be 1×10^{-8}. The tabulated values are rounded off in the fourth decimal places for $\hat{n}_1$, and $\hat{n}_2$, and in the sixth decimal places for $\hat{c}$.

The Performance of R_1 Relative to R_0

In order to get some insight into the performance of the two-stage procedure R_1, we consider the ratio(termed relative efficiency RE),

$$RE=\{E_{\underline{\theta}}(TSS \mid R_1)/kn_0\}\times100(\%), \quad (4.9)$$

where n_0 is the sample size needed for the single-stage procedure of Bechhofer and Sobel(1954) to satisfy the same probability requirement. Clearly RE depends on $\underline{\theta}$, (δ^*, P^*) and k.

Table 1. Design Constants for the Two-Stage Procedure R_1

k=2

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$	P*=0.95 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$
0.50	4.672	6.029	0.99999	6.544	6.354	0.92158
0.55	5.365	3.894	0.76742	6.178	3.825	0.46474
0.60	5.039	2.211	0.17160	7.771	4.964	0.45385
0.65	6.435	3.426	0.21494	10.181	6.975	0.47896
0.70	8.621	5.420	0.27360	13.960	10.354	0.52255

k=3

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$	P*=0.95 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$
0.50	6.208	6.542	0.90887	6.901	5.003	0.63690
0.55	5.746	3.936	0.43711	8.328	5.527	0.56113
0.60	7.204	4.982	0.42234	10.606	7.296	0.56670
0.65	9.417	6.829	0.44267	14.037	10.107	0.59111
0.70	12.201	10.822	0.47227	19.426	14.711	0.62732

k=4

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$	P*=0.95 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$
0.50	5.551	4.649	0.54605	7.459	5.377	0.57673
0.55	6.721	5.609	0.51311	9.313	6.936	0.58219
0.60	8..479	7.249	0.51621	11.950	9.270	0.60041
0.65	11.119	9.822	0.53815	15.897	12.840	0.62928
0.70	15.293	14.016	0.57488	21.727	18.912	0.66103

k=5

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$	P*=0.95 $\hat{n}_1$	$\hat{n}_2$	$\hat{c}$
0.50	5.899	5.547	0.53513	7.850	6.194	0.56399
0.55	7.287	6.799	0.52954	9.880	8.087	0.58431
0.60	9.264	8.923	0.54652	12.733	10.866	0.60955
0.65	12.202	12.165	0.57446	16.992	15.071	0.64215
0.70	16.754	17.388	0.60969	23.766	21.753	0.68249

Table 1. Design Constants for the Two-Stage Procedure R_1(continued)

k=6

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	P*=0.90 $\hat{n}_2$	P*=0.90 $\hat{c}$	P*=0.95 $\hat{n}_1$	P*=0.95 $\hat{n}_2$	P*=0.95 $\hat{c}$
0.50	6.138	6.034	0.51152	8.128	6.906	0.55542
0.55	7.663	7.763	0.53191	10.260	9.061	0.58159
0.60	9.787	10.273	0.55662	13.257	12.206	0.61114
0.65	12.943	14.057	0.58952	17.737	16.930	0.64665
0.70	17.775	20.145	0.62632	24.894	24.289	0.68800

k=7

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	P*=0.90 $\hat{n}_2$	P*=0.90 $\hat{c}$	P*=0.95 $\hat{n}_1$	P*=0.95 $\hat{n}_2$	P*=0.95 $\hat{c}$
0.50	6.238	6.640	0.49599	8.345	7.534	0.54877
0.55	7.940	8.589	0.53035	10.537	9.900	0.57717
0.60	10.174	11.410	0.55998	13.643	13.344	0.60980
0.65	13.488	15.661	0.59628	18.563	18.491	0.65756
0.70	18.569	22.392	0.63444	25.741	26.644	0.69189

k=8

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	P*=0.90 $\hat{n}_2$	P*=0.90 $\hat{c}$	P*=0.95 $\hat{n}_1$	P*=0.95 $\hat{n}_2$	P*=0.95 $\hat{c}$
0.50	6.440	7.118	0.48408	8.528	8.075	0.54346
0.55	8.157	9.308	0.52725	10.754	10.640	0.57246
0.60	10.477	12.403	0.56057	13.946	14.345	0.60751
0.65	13.923	17.032	0.59933	19.344	20.491	0.66845
0.70	19.214	24.432	0.64026	26.431	28.944	0.69637

k= 9

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	P*=0.90 $\hat{n}_2$	P*=0.90 $\hat{c}$	P*=0.95 $\hat{n}_1$	P*=0.95 $\hat{n}_2$	P*=0.95 $\hat{c}$
0.50	6.504	7.693	0.47554	8.685	8.558	0.53907
0.55	8.335	9.949	0.52377	10.931	11.297	0.56789
0.60	10.724	13.274	0.55959	14.192	15.237	0.60478
0.65	14.269	18.254	0.60020	19.834	21.536	0.66514
0.70	19.741	26.317	0.64419	26.924	30.716	0.69566

Table 1. Design Constants for the Two-Stage Procedure R_1(continued)

k=10

$\sqrt{\delta^*}$	P*=0.90 $\hat{n}_1$	P8=0.90 $\hat{n}_2$	$\hat{c}$	$\hat{n}_1$	P*=0.95 $\hat{n}_2$	$\hat{c}$
0.50	6.609	8.126	0.46947	8.822	8.991	0.53317
0.55	8.484	10.524	0.52004	11.085	11.888	0.56385
0.60	10.932	14.058	0.55798	14.399	16.037	0.60188
0.65	14.560	19.347	0.60004	20.013	22.682	0.66819
0.70	20.125	27.832	0.64840	27.292	31.970	0.69114

Table 2. Relative Efficiences RE of the Procedure R_1

P*=0.90

	$\sqrt{\delta^*}$									
	0.50		0.55		0.60		0.65		0.70	
k	EPC	LFC	EPC	LFC	EPC	LFC	EPC	LFC	EPC	LFC
2	99.9	94.5	99.9	97.6	99.9	94.9	96.3	91.2	97.5	92.0
3	98.1	92.7	98.3	85.2	95.6	83.6	95.9	84.5	97.5	86.0
4	99.5	85.7	92.7	79.5	94.7	81.3	94.0	80.8	96.4	83.0
5	95.8	82.3	91.9	78.8	95.7	81.9	88.4	75.6	94.3	80.5
6	91.0	78.4	89.4	76.7	89.3	76.5	93.1	79.5	90.4	77.0
7	93.4	80.7	86.2	74.2	87.7	75.2	88.8	76.0	88.5	75.5
8	88.4	76.5	88.5	76.3	85.5	73.5	84.6	72.4	86.2	73.5
9	89.9	77.9	79.6	68.9	87.3	75.2	86.4	74.2	85.8	73.4
10	84.3	73.2	80.9	70.1	84.6	73.0	81.9	70.4	82.5	70.7

P*=0.95

	$\sqrt{\delta^*}$									
	0.50		0.55		0.60		0.65		0.70	
k	EPC	LFC	EPC	LFC	EPC	LFC	EPC	LFC	EPC	LFC
2	99.9	96.0	99.9	85.8	99.9	84.3	99.9	85.2	98.4	83.4
3	99.9	86.8	99.7	82.7	98.1	81.2	96.9	79.9	96.9	80.0
4	94.9	79.1	95.5	79.1	97.5	80.3	91.9	75.4	93.6	76.5
5	92.6	77.1	94.8	78.6	93.7	77.2	91.0	74.7	92.6	75.7
6	89.1	74.4	92.7	76.9	89.0	73.4	88.7	72.9	89.8	73.6
7	91.5	76.5	90.0	74.8	87.7	72.5	85.8	70.8	86.5	71.0
8	87.1	73.0	87.0	72.5	86.0	71.3	83.6	69.0	86.6	71.2
9	88.5	74.4	84.0	70.2	84.2	69.9	85.4	70.7	84.5	69.6
10	84.2	70.9	81.1	67.9	85.2	71.0	81.0	67.4	82.1	68.0

Since R_0 is a special case of R_1(with $c=1$ or ∞), it immediately follows that $1 \geq RE(EPC) \geq RE(LFC)$ and R_1 is at least as good as R_0. The values of RE are given in Table 2.

From Table 2 one can see that the relative savings by applying the two-stage procedure R_1 are not dramatic. However, even a small relative saving means a lot in terms of total sample size when k and/or n_1 is moderately large. Also it can be observed that the relative saving increases as k becomes large. This is in accordance with one's intuition that the screening process would be helpful when k is large.

To illustrate the use of Table 1, suppose that $k=6$ and that the experimenter specifies $\delta^*=(0.7)^2$, $P^*=0.90$. Then the design constants necessary to implement the two-stage procedure R_1 are given by $\hat{n}_1=17.775$, $\hat{n}_2=20.145$ and $\hat{c}=0.62632$. Thus we take $n_1=18$ observations from each populations and compute the sample variances $S_i^2(1 \leq i \leq 6)$. If the number of S_i^2's smaller than min $S_i^2/0.62632$ is one, stop sampling and assert that the population associated with min S_i^2 is the best. If more than one S_i^2's are smaller than min $S_i^2/0.62632$, take $n_2=21$ additional observations from each of the contending populations and assert that the population associated with the smallest sample variance based on the pooled sample of size $n_1+n_2=39$ is the best. In using this two-stage procedure R_1, the average value of the total number of observations is 90.4% at EPC and 77.0% at LFC compared with that of the single-stage procedure R_0 of Bechhofer and Sobel(1954).

Large Sample Approximation

For solving the optimization problem involving(4.8) and (4.6), it is extremely tedious to compute the integrals when n_1 and/or n_2 are large. Hence an approximate solution for large sample size is useful. We shall give an approximate solution to the problem based on normal theory.

It is well known that if S^2 is the sample variance associated with the variance σ^2, then $\sqrt{(\nu-1)/2}\,\log(S^2/\sigma^2)$ is asymptotically normally distributed with mean zero and variance unity as the number of degrees of freedom ν, associated with S^2, tends to infinity. From this fact, it can be shown that, when ν is large

$$\int_0^\infty \{1-F_\nu(ax)\}^{k-1}\, dF_\nu(x) \cong \int_{-\infty}^{\infty} \Phi^{k-1}(x+d)\,d\Phi(x) \tag{4.10}$$

where $d=\sqrt{(\nu-1)/2}\,\log(a^{-1})$, and $F_\nu(\cdot)$ is the cdf of chi-squared distribution with ν degrees of freedom.

Replacing the integrals in (4.6) and (4.8) involving the chi-squared cdf's by the corresponding integrals of the right hand side of (4.10) involving the normal cdf, we can obtain after slight modifications the following asymptotic version of the optimization problem:

Minimize

$$kc_1^2+kc_2^2\int_{-\infty}^{\infty}\{\Phi^{k-1}(x+d)-\Phi^{k-1}(x-d)\}d\Phi(x) \qquad (4.11)$$

subject to

$$\int_{-\infty}^{\infty}\Phi^{k-1}(x+d+c_1)d\Phi(x)\int_{-\infty}^{\infty}\Phi^{k-1}(x+\sqrt{c_1^2+c_2^2})\,d\Phi(x)\geq P^* \qquad (4.12)$$

If we denote the solutions to (4.11) and (4.12) by $(\hat{c}_1, \hat{c}_2, \hat{d})$, then the approximate values of the design constants$(\hat{n}_1, \hat{n}_2, \hat{c})$ of the procedure R_1 are computed using the following formulas:

$$\hat{n}_1=2\left\{\frac{\hat{c}_1}{\log(\delta^{*-1})}\right\}^2+2\log(k-1),$$

$$\hat{n}_2=2\left\{\frac{\hat{c}_2}{\log(\delta^{*-1})}\right\}^2+2\log(k-1)$$

and

$$\hat{c}=\exp\{-\hat{d}\sqrt{2/(n_1-2)}\}$$

The second term in the formula for $\hat{n}_1$(or $\hat{n}_2$) is a slight correction term based on empirical results cited from Gibbons, Olkin and Sobel(1977). The correction term is added since the first term drifts below the true value of $\hat{n}_1$(or $\hat{n}_2$) as k increases.

The values of $(\hat{c}_2, \hat{c}_2, \hat{d})$ can be found in the tables of Tamhane and Bechhofer(1979). To illustrate numerically the closeness of the normal approximation we take the values of $(\hat{c}_1, \hat{c}_2, \hat{d})$ out of Table 2 of Tamhane and Bechhofer(1979) corresponding to $P^*=0.95$, $k=10$, namely, $\hat{c}_1=2.452$, $\hat{c}_2=2.744$ and $\hat{d}=1.322$. Then the approximate values of $(\hat{n}_1, \hat{n}_2, \hat{c})$ for $\delta^*=(0.7)^2$ are (28.025, 33.988, 0.69317). These approximate values are slightly larger than the corresponding exact values (27.292, 31.970, 0.69114) given in Table 1.

Acknowledgement

The author wishes to thank to Professor Shanti S. Gupta, Professor Woo-Chul Kim and Dr. Tachen Liang for their extensive comments for this work. Also thanks are to the Department of Statistics, Purdue University for giving me an opportunity to visit during 1986-87 and for providing facilities for this research.

References

1. Alam, K.(1970). A Two-Sample Procedure for Selecting the Population with the Largest Mean from k Normal Populations. *Annals of the Institute of Statistical Mathematics,* Vol. 22, 127-136.
2. Bechhofer, R.E.(1954). A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances. *Annals of Mathematical Statistics,* Vol. 25, 16-39.
3. Bechhofer, R.E., Dunnett, C.W. and Sobel, M.(1954). A Two-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with a Common Unknown Variance. *Biometrika,* Vol. 41, 170-176.
4. Bechhofer, R.E. and Sobel, M.(1954). A Single-Sample Multiple Decision Procedure for Ranking Variances of Normal Populations. *Annals of Mathematical Statistics,* Vol. 25, 273-289.
5. Cohen, D.S.(1959). A Two-Sample Decision Procedure for Ranking Means of Normal Populations with a Common Known Variance. M.S. Thesis, Department of Operatiosn Research, Cornell Univ., Ithaca, New York.
6. Fiacco, A.V. and McCormick, G.P.(1968). *Nonlinear Sequential Unconstrained Minimization Techniques,* New York.: John Wiley and Sons, Inc.
7. Gibbsons, J.D., Olkin, I. and Sobel, M.(1977). *Selecting and Ordering Populations: A New Statistical Methodology*, New York.: John Wiley and Sons, Inc.
8. Gupta, S.S.(1956). On a Decision Rule for a Problem in Ranking Means. Ph. D. Thesis(Mimeo. Ser. No. 150). Inst. of Statist., Univ., of North Carolina, Chapel Hill.
9. Gupta, S.S.(1965). On Some Multiple Decision(Selection and Ranking) Rules. *Technometrics,* Vol. 7, 225-245.
10. Gupta, S.S. and Kim, W.C.(1984). A Two-Stage Elimination Type Procedure for Selecting the Largest of Several Normal Means with a Common Unknown Variance. *Design of Experiments: Ranking and Selections.* (T.J. Santner and A.O. Tamhane, eds.), New York.: Marcel Dekker, 77-94.
11. Gupta, S.S. and Miescke, K.J.(1982). On the Least Favorable Configurations in Certain Two-Stage Selection Procedures. *Statistics and Probability: Essays in Honor of C.R. Rao.*(G. Kallianpur et al., eds.), Amsterdam-New York-Oxford: North-Holland, 295-305.
12. Gupta, S.S. and Panchapakesan, S.(1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations,* New York.: John Wiley and Sons, Inc.
13. Gupta, S.S. and Sobel, M.(1962a). On Selecting a Subset Containing the Populations with the Smallest Variance. *Biometrika,* Vol. 49, 495-507.

14. Gupta, S.S. and Sobel, M.(1962b). On the Smallest of Several Correlated F-stiatistics. *Biometrika,* Vol. 49, 509-523.

15. Hochberg, Y. and Marcus, R.(1981). Three Stage Elimination Type Procedures for Selecting the Best Normal Population when Variances are Unknown. *Communication in Statistics, Theory and Method,* A 10, 597-612.

16. Miescke, K.J. and Sehr, J.(1980). On a Conjecture Concerning Least Favorable Configurations in Certain Two-Stage Selection Procedures, *Communication in Statistics, Theory and Method,* A 9, 1609-1617.

17. Santner, T.J.(1975). A Restricted Subset Selection Approach to Ranking and Selection Problems. *The Annals of Statistics,* Vol. 3, 334-349.

18. Stein, C.(1945). A Two-Sample Test for a Linear Hypothesis whose Power is Independent of the Variance. *Annals of Mathematical Statistics,* Vol. 16, 243-258.

19. Tamhane, A.C.(1975). A Minimax Two-Stage Permanent Elimination Type Procedure for Selecting the Smallest Normal Variance. Tech. Report 260, Department of Operations Research, Cornell Univ., Ithaca, New York.

20. Tamhane, A.C.(1976). A Three-Stage Elimination Type Procedure for Selecting the Largest Normal Mean(Common Unknown Variance). *Sankhyá,* B 38, 339-349.

21. Tamhane, A.C. and Bechhofer, R.E.(1977). A Two-Stage Minimax Procedure with Screening for Selecting the Largest Normal Mean. *Communication in Statistics, Theory and Method,* A 6, 1003-1033.

22. Tamhane, A.C. and Bechhofer, R.E.(1979). A Two-Stage Minimax Procedure with Screening for Selecting the Largest Normal Mean(II): An Improved PCS Lower Bound and Associated Tables. *Communication in Statistics, Theory and Method,* A 8, 337-358.