ON A MAXIMIN STRATEGY FOR SAMPLING BASED ON SELECTION PROCEDURES FROM SEVERAL POPULATIONS

Shanti S. Gupta and Wen-Tao Huang

Purdue University

ABSTRACT

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k populations such that π_i has cdf $F(x;\theta_i)$. It is permitted to draw n samples from each of these k populations and the sum of the n observations in the sample is considered to be the reward. It is desired to make the expected reward as large as possible. This problem is formulated as a game based on a selection procedure. The maximin strategy of the game formulated is used for the sampling scheme. Some properties of selection rules advantageous for the sampling scheme are studied. Sufficient conditions for the existence of the value of the game are given. The least favorable configuration for the game is found for the case where the populations are normal with common known variance. An asymptotic optimal property for the maximin strategy is shown to hold.

0. INTRODUCTION AND SUMMARY

In a pioneering paper Robbins (1952), proposed a problem of sequential sampling of two populations. In this problem two populations are given such that their cumulative distribut ons belong to some class. Suppose their respective means exist; h e draw n samples from the two populations if one's object is to achieve the greatest possible expected value of the sum of the n observations? In the special case of Bernoulli trials i.e. the two-armed-bandit problem, much work has been done since then by several authors, for example, Smith and Pyke (1965). But in the general formulation of this problem, no good sampling scheme has been proposed so far. In this paper, we propose a two-stage scheme of sampling based on a selection procedure. We formulate the problem in a game theoretic set-up and obtain a maximin strategy.

In Section 1 we give the formulation of the problem. A maximin strategy is derived in Section 2. Asymptotic optimal properties are studied in Section 3. Some numerical computations related to the maximin strategy are given in Tables 1-6 at the end of the paper.

1. NOTATION AND FORMULATION OF THE PROBLEM

Let $\pi_1,\pi_2,\ldots,\pi_k$ be k populations such that π_i has the cumulative distribution, c.d.f., $F(x;\theta_i)$, $i=1,2,\ldots,k$. Let Ω denote the parameter space of $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$. For $d > 0$, define

$\Omega(d) = \{\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k) | \theta_{[k]}-\theta_{[k-1]} \geq d\}$, where $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ are the ordered values of $\theta_1,\theta_2,\ldots,\theta_k$. Let X_{ij} denote the jth independent observation from π_i and define $S_{in} = \sum_{j=1}^{n} X_{ij}$, $i=1,2,\ldots,k$, $n=1,2,\ldots$. Let $W(X_1,X_2,\ldots,X_n)$ denote a statistic of random variables $X_1,X_2,\ldots,X_n$.

Suppose we are allowed to draw a total of n observations from the k populations. We are perfectly free to choose a population from which we draw our αth observation, $\alpha=1,2,\ldots,n$. Let $x_1,x_2,\ldots,x_n$ be the samples drawn, then a reward $W(x_1,x_2,\ldots,x_n)$ is received. The problem is how to design a sampling strategy so that the expected reward $E_{\underline{\theta}}W(X_1,X_2,\ldots,X_n)$ is as large as possible. In this paper we confine ourselves to the reward function $W(X_1,X_2,\ldots,X_n) = X_1 + X_2 +\ldots+ X_n$. We note that when π_i is a coin with the probability of a head p_i, this becomes the k-armed-bandit problem with a finite number of tosses.

2. A MAXIMIN STRATEGY

By a test block $U(m)$ we mean a sequence of random outcomes $\{X_{11}, X_{12}, \ldots, X_{1m}, X_{21}, X_{22}, \ldots, X_{2m}, \ldots, X_{k1}, X_{k2}, \ldots, X_{km}\}$ where X_{ij} is the jth outcome of π_i. By a trial block $V(i,m)$ we mean a sequence of m random outcomes of π_i, i.e. $\{X_{i1}, X_{i2}, \ldots, X_{im}\}$. Let R be a selection procedure for selecting the unique population associated with $\theta_{[k]}$, the largest parameter, when $\Omega = \Omega(d)$. Let $\pi(R;m)$ be a random variable taking value in $\{1,2,\ldots,k\}$ such that $\pi(R;m) = i$ means π_i is selected based on m independent observations from each population using the rule R. It should be pointed out that here we are using the indifference zone type selection (see Bechhofer (1954) and not the subset type selection (see Gupta (1965)). Define $\gamma_i(m;\underline{\theta}) = P_{\underline{\theta}}\{\pi(R;m)=i\}$, $i=1,2,\ldots,k$. For convenience, when there is no possible confusion, we write $\gamma_i(m)$ instead of $\gamma_i(m;\underline{\theta})$. Let $E_{\underline{\theta}}X_{i1} = t(\theta_i)$, $i=1,2,\ldots,k$, where $\underline{\theta}=(\theta_1,\theta_2,\ldots,\theta_k)$. Without loss of generality, we assume $t(\theta_k)= \max_{1<i<k} t(\theta_i)$. By a correct selection, CS, we mean that π_k is selected. Define $\overline{\gamma(m)} = \inf_{\Omega(d)} P\{CS|R\}$. For a given integer m ($0\le m\le[\frac{n}{k}]$), a $UV(R;m)$ scheme is a sampling strategy which follows a test block $U(m)$ first and then follows a trial block $V(\pi(R;m),n-km)$. Let $W(R;m)$ denote the reward using $UV(R;m)$ scheme. Then, it is obvious that $W(R,m) = \sum_{i=1}^{k} S_{im} + S_{\pi(R;m),n-km}$. Let $A = \{0,1,2,\ldots,[\frac{n}{k}]\}$ and $B = \Omega(d)$. For given R, define $G(a,b) = E_b W(R;a)$ for $a \in A$ and $b \in B$. Now if we consider a zero-sum two person game such that a statistician (player I) plays a game against nature (player II) with reward G, then, following the UV scheme we have a game (A,B,G). In this game, player II tries to minimize the reward G while player I wants to maximize the reward. Hence, a good strategy for player I is a maximin strategy, i.e. player I needs to choose $m^*=m(R,k,n) \in A$ so that

$$G(m^*,\underline{\theta}) = \max_{0\le m\le[\frac{n}{k}]} \inf_{\underline{\theta}\in\Omega(d)} E_{\underline{\theta}} W(R;m). \tag{2.1}$$

When player I chooses m* as his maximin strategy, he follows the sampling scheme UV(R;m*). It is clear that a maximin strategy always exists for our problem.

We note that

(2.2) $$G(m,\underline{\theta}) = m \sum_{i=1}^{k} t(\theta_i) + (n-km) \sum_{i=1}^{k} \gamma_i(m;\underline{\theta})\, t(\theta_i)$$

Let

(2.3) $$L(R;m) = md + (n-km)\gamma(m)d$$

$$U(R;m) = md + (n-km)\gamma_k(m;\underline{\theta}_0) \quad \text{where } \underline{\theta}_0 = (0,0,\ldots,0,d).$$

Then, if $\theta_i \geq 0$ and $t(\theta)=\theta$, we have

(2.4) $$\max_{m\epsilon A} L(R;m) \leq \max_{m\epsilon A} \inf_{\underline{\theta}\epsilon\Omega(d)} G(m,\underline{\theta}) \leq \max_{m\epsilon A} U(R;m)$$

Let A* denote the convex hull of A. Let $[\frac{n}{k}]=\ell$, then A* is a closed convex subset in R^ℓ, the ℓ-dimensional Euclidean space. Let $\underline{\gamma}(m;\underline{\theta}) = (\gamma_1(m;\theta_1), \gamma_2(m;\theta_2),\ldots, \gamma_k(m;\theta_k))$ and $\underline{t}(\underline{\theta})=(t(\theta_1),t(\theta_2),\ldots,t(\theta_k))$. Let $\gamma(\alpha)$ denote the polygonal interpolation of $\{\gamma(m);\ m=0,1,2,\ldots,\ell\}$ for any real α, $0 \leq \alpha \leq \ell$. Also, let $\gamma_i(\alpha)$ be extended by the polygonal interpolation, $i=1,2,\ldots,k$. Then, $\gamma(\alpha)$ and $\gamma_i(\alpha)$ are continuous in α and we have the following theorem.

Theorem 2.1. If $\Omega(d)$ is a closed convex subset of R^k and $t(\theta)$ is continuous and convex in $\underline{\theta}$ and $\gamma_i(m;\underline{\theta})$ is continuous in $\underline{\theta}$ and $\underline{t}(\underline{\theta})$. $\underline{\gamma}(a;\underline{\theta})$ is convex in $\underline{\theta}$ for every $a\epsilon A^*$,then, $(A^*,\Omega(d),G)$ has a value and player II has a good pure strategy and player I has a good strategy which is a mixture of at most $\min(\ell+1,k)$ pure strategies.

Proof: For $a\epsilon A^*$ and $\underline{\theta}\epsilon\Omega(d)$ there exists $\underline{\alpha} = (\alpha_0,\alpha_1,\ldots,\alpha_\ell)$ such that $\sum_{i=0}^{\ell} \alpha_i = 1$ and $a = \sum_{i=1}^{\ell} i\alpha_i$. We note that $G(a,\underline{\theta}) = (\sum_{i=1}^{\ell} i\alpha_i)(\sum_{j=1}^{k} t(\theta_j)) + \sum_{i=0}^{\ell} (n-ki\alpha_i) \sum_{j=1}^{k} \gamma_j(i\alpha_i;\underline{\theta})t(\theta_j)$. By our assumption we see that $G(a,\underline{\theta})$ is continuous and convex in $\underline{\theta}$ for every $a\epsilon A^*$ and $G(a,\underline{\theta})$ is continuous in a for every $\underline{\theta}$ in $\Omega(d)$. Then, using some results in game theory (see, for example,

p. 53 of Blackwell and Girshick (1954), it follows that the game has a value and player II has a good strategy which is a mixture of at most $\min(\ell+1,k)$ pure strategies.

Corollary 2.1. If $\Omega(d)$ is closed and convex in R^k and $t(\theta)$ is linear in θ and $\gamma_i(a;\underline{\theta})$ is convex in $\underline{\theta}$ and monotome increasing in θ_i for each $i=1,2,\ldots,k$, then, the result of theorem 2.1 holds.

Proof: Since $t(\theta)$ is linear it is convex and continuous in θ. To see $\underline{t}(\theta)\cdot\underline{\gamma}(a;\underline{\theta})$ is convex in $\underline{\theta}$, it suffices to consider the case $t(\theta)=\theta$. Let $f(\underline{\theta})=\underline{\theta}\cdot\underline{\gamma}(a;\underline{\theta})$, the inner product of two vectors. Suppose $\underline{\theta}_1$ and $\underline{\theta}_2$ are in $\Omega(d)$ and $0<\alpha<1$, $\beta=1-\alpha$. Let $\underline{\theta}_i=(\theta_{i1},\theta_{i2},\ldots,\theta_{ik})$, $i=1,2$. Then,

$$(2.5)\quad f(\alpha\underline{\theta}_1+\beta\underline{\theta}_2)=\alpha\sum_{j=1}^{k}\theta_{1j}\gamma_j(a;\alpha\underline{\theta}_1+\beta\underline{\theta}_2)+\beta\sum_{j=1}^{k}\theta_{2j}\gamma_j(a;\alpha\underline{\theta}_1+\beta\underline{\theta}_2)$$

$$\le\alpha^2\sum_{j=1}^{k}\theta_{1j}\gamma_j(\underline{\theta}_1)+\beta^2\sum_{j=1}^{k}\theta_{2j}\gamma_j(\underline{\theta}_2)+\alpha\beta\sum_{j=1}^{k}[\theta_{1j}\gamma_j(\underline{\theta}_2)+\theta_{2j}\gamma_j(\underline{\theta}_1)]$$

$$(2.6)\quad \alpha f(\underline{\theta}_1)+\beta f(\underline{\theta}_2)=\alpha\sum_{j=1}^{k}\theta_{1j}\gamma_j(\underline{\theta}_1)+\beta\sum_{j=1}^{k}\theta_{2j}\gamma_j(\underline{\theta}_2).$$

It follows from (2.5) and (2.6) that $f(\underline{\theta})$ is convex, if, and only if,

$$(2.7)\quad \alpha(1-\alpha)\sum_{j=1}^{k}\theta_{1j}\gamma_j(\underline{\theta}_1)+\beta(1-\beta)\sum_{j=1}^{k}\theta_{2j}\gamma_j(\underline{\theta}_2)\ge\alpha\beta\sum_{j=1}^{k}[\theta_{1j}\gamma_j(\underline{\theta}_2)+\theta_{2j}\gamma_j(\underline{\theta}_1)]$$

Or, equivalently,

$$(2.8)\quad (\underline{\gamma}(\underline{\theta}_1)-\underline{\gamma}(\underline{\theta}_2))\cdot(\underline{\theta}_1-\underline{\theta}_2)\ge 0,\ \text{since}\ \alpha(1-\alpha)=\beta(1-\beta)=\alpha\beta\ne 0.$$

Note that (2.8) is satisfied if $\gamma_i(\underline{\theta})$ is monotone increasing in θ_i, $i=1,2,\ldots,k$.

This completes the proof.

Remark 2.1: When $t(\theta)$ is non-decreasing in θ, $\Omega(d)$ need only be bounded below and be convex in R^k.

Definition. A selection rule R is most economical if, for any other selection rule R', we have

(i) $\gamma(R;m)\le\gamma(R';n)$ implies $n\ge m$ and

(ii) $\lim_{m\to\infty} \gamma(R;m) = 1$

It has been shown by Hall (1959) that Bechhofer's procedure (1954) R_B, and Sobel-Huyett procedure (1957), R_{SH}, both satisfy condition (i) of the above definition. By the strong law of large numbers, R_B and R_{SH} also satisfy (ii). For the problem of selecting the largest $t(\theta_i)$, let C denote the set of the most economical selection rules. Then, the rules in C have the following property.

Corollary 2.2. If $t(\theta)=\theta$ and $R\varepsilon C$, then, for any rule R' other than R, we have

$$\max_{m\varepsilon A} L(R';m) \le \max_{m\varepsilon A} L(R;m) + \min\{n\delta(R;\ell),\ k\gamma(R;\ell+1)\}.d$$

where $L(R;m)$ is defined by (2.3) and $\ell = [\frac{n}{k}]$, $\delta(R;\ell) = \max_{m\varepsilon A}\{\gamma(R;m+1)-\gamma(R;m)\}$

Proof: Since $R\varepsilon C$, for any R', we have $\gamma(R';m) < \gamma(R;m+1)$ for $m\varepsilon A$. Hence, we have

$$(2.9)\quad \max_{s\varepsilon A} L(R';s) \le md + (n-km)\gamma(R;m+1)d \text{ for some } m\varepsilon A$$

$$= [(m+1)d+(n-k(m+1))\gamma(R;m+1)d]-d+k\gamma(R;m+1)d \text{ for some } m\varepsilon A.$$

$$\le \max_{s\varepsilon A} L(R;s)+k\gamma(R;\ell+1)d \text{ since } \gamma(R;m) \text{ is monotone increasing in } m.$$

On the other hand,

$$(2.10)\quad \max_{s\varepsilon A} L(R';s) \le md+(n-km)[\gamma(R;m)+\delta(R;\ell)]d \text{ for some } m\varepsilon A$$

$$\le \max_{m'\varepsilon A} [m'd+(n-km')\gamma(R;m')d]+(n-km)\delta(R;\ell)d$$

$$\le \max_{s\varepsilon A} L(R;s) + n\delta(R;\ell)d$$

It follows from (2.9) and (2.10) that the result holds.

Henceforth, we confine ourselves to C whenever it is non-empty.

To find a maximin strategy we need to find a value $\underline{\theta}$ in $\overline{\Omega(d)}$, the closure of $\Omega(d)$, at which $G(m,\underline{\theta})$ attains its minimum. There is no general rule to do this. However, for $k = 2$, we have the following sufficient conditions.

Corollary 2.3. Suppose $t(\theta)$ is twice differentiable and $\gamma_1(m;\underline{\theta})$ is also twice partially differentiable with respect to θ_1 and θ_2. If $\underline{\theta}^* = (\theta_1,\theta_2)$ satisfies $|\theta_1 - \theta_2| \geq d$ such that $A(\theta_1,\theta_2)=B(\theta_1,\theta_2) = 0$ and $C(\theta_1,\theta_2)\cdot E(\theta_1,\theta_2) > D^2(\theta_1,\theta_2)$, then $G(m,\theta^*)$ attains its minimum on $\Omega(d)$, where

$$A(\theta_1,\theta_2) = m\, t'(\theta_1) + (n-2m)\left[\frac{\partial\gamma_1}{\partial\theta_1}(t(\theta_1)-t(\theta_2))+\gamma_1 t'(\theta_1)\right]$$

$$B(\theta_1,\theta_2) = m\, t'(\theta_2) + (n-2m)\left[\frac{\partial\gamma_1}{\partial\theta_2}(t(\theta_1)-t(\theta_2))+(1-\gamma_1) t'(\theta_2)\right]$$

$$C(\theta_1,\theta_2) = m\, t''(\theta_1) + (n-2m)\left[\frac{\partial^2\gamma_1}{\partial\theta_1^2}(t(\theta_1)-t(\theta_2))+2\frac{\partial\gamma_1}{\partial\theta_1} t'(\theta_1)+\gamma_1 t''(\theta_1)\right]$$

$$D(\theta_1,\theta_2) = (n-2m)\left[\frac{\partial^2\gamma_1}{\partial\theta_1\partial\theta_2}(t(\theta_1)-t(\theta_2))+\frac{\partial\gamma_1}{\partial\theta_1}(t'(\theta_1)-t'(\theta_2)\right]$$

$$E(\theta_1,\theta_2) = m\, t''(\theta_2) + (n-2m)\left[\frac{\partial^2\gamma_1}{\partial\theta_2^2}(t(\theta_1)-t(\theta_2))+(1-\gamma_1)t''(\theta_2)-2\frac{\partial\gamma_1}{\partial\theta_2} t'(\theta_2)\right]$$

The proof is straightforward. We now call such θ^* the least favorable configuration for the reward function.

Theorem 2.2. If π_i is a normal population with mean θ_i ($\theta_i \geq 0$) and common variance 1, for $i=1,2,\ldots,k$, then, the least favorable configuration with respect to the reward function is $\theta^* = (0,0,\ldots,0,\ d)$.

Proof: Since $G(m,\underline{\theta}) = m \sum_{i=1}^{k} \theta_i + (n-km) \sum_{i=1}^{k} \gamma_i(m)\theta_i$, it is clear that $G(m,\underline{\theta})$ is invariant with respect to any permutation of $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$. Then, we may assume for given $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$ that $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k - d$. Let $\delta_{ij} = \theta_i - \theta_j$. Then, we obtain $\gamma_i(m;R_B) = \int_{-\infty}^{\infty} \prod_{\substack{j=1\\ j\neq i}}^{k} \Phi(x;\delta_{ij}\sqrt{m},1)\, d\Phi(x)$

for $i=1,2,\ldots,\ k$, where $(x;\alpha,\beta)$ denotes the normal cdf with mean α and variance β.

Let $h_{i+1} = \theta_{i+1} - \theta_i$ for $i=1,2,\ldots,k-1$ and let $h_1 = \theta_1$. Then, for given $\underline{\theta}$ there corresponds a unique $\underline{h} = (h_1,h_2,\ldots,h_k)$. We note that if $\underline{\theta}_1$ and $\underline{\theta}_2$

are associated with the same $(h_2,h_3,\ldots,h_k)$, $\gamma_i(m;R_B)$ remains same value under $\underline{\theta}_1$ and $\underline{\theta}_2$ respectively, for $i=1,2,\ldots,k$. It suffices to show the following facts.

(i) Let $\underline{\theta}_1$ and $\underline{\theta}_2$ be associated with same vector $(h_2,h_3,\ldots,h_k)$ and let h_1 and h_1' be respectively associated with $\underline{\theta}_1$ and $\underline{\theta}_2$. Then, $G(m,\underline{\theta}_1) \leq G(m,\theta_2)$ if $h_1 = 0$. The inequality holds if $h_1' > 0$. It is because $G(m,\underline{\theta}_2) = m[k\ h_1' + (k-1)h_2 + \ldots + k\ h_k] + (n-km)\sum_{i=1}^{k}\gamma_i(m)(h_1'+h_2+\ldots+h_i) \geq m\ [(k-1)h_2+(k-2)h_3+\ldots+k\ h_k] + (n-km)\sum_{i=1}^{k}\gamma_i(m)(h_2+h_3+\ldots+h_i) = G(m,\underline{\theta}_1)$.

(ii) Let $\theta^* = (\theta_1^*,\theta_2^*,\ldots,\theta_k^*)$ be the least favorable configuration for G. Let $\underline{h}^* = (h_1^*,h_2^*,\ldots,h_k^*)$ be associated with $\underline{\theta}^*$. Let θ_j^* be the first $j\neq k$ such that $\theta_j^* > 0$. Let $\underline{\tilde{\theta}} = (0,0,\ldots,0,\ \theta_{j+1}^*-\theta_j^*,\theta_{j+2}^*-\theta_j^*,\ldots,\theta_k^*-\theta_j^*)$. Then the associated $\underline{\tilde{h}}$-vector keeps the same as that of $\underline{\theta}^*$ except $\tilde{h}_j=0$ and $h_j^*>0$. By the same argument of (i), we see that for $m\in A$ $G(m,\underline{\tilde{\theta}}) < G(m,\underline{\theta}^*)$ and which contradicts our assumption that $\underline{\theta}^*$ is least favorable. Hence, we conclude that $\underline{\theta}^* = (0,0,\ldots,0,\theta_k^*)$ with $\theta_k^* \geq d$. Let $\underline{\theta}'^* = (0,0,\ldots,0,d)$, then, we see that $G(m,\underline{\theta}'^*) < G(m,\underline{\theta}^*)$ if $\theta_k^* > d$. Therefore, we conclude that $\underline{\theta}^* = (0,0,\ldots,0,d)$.

This completes the proof.

Some tables are tabulated in this paper. Table I and Table II give m*-values and $G(m^*,\underline{\theta}^*)$ for the normal distributions. Table III and IV give the m*-value and $G(m^*,\underline{\theta}^*)$ for 2 binomial populations. Maximum values of m for the lower bounds L defined in (2.3) for binomial populations are given in Table 5. Table 6 gives some values of the lower bounds of $G(m^*,\underline{\theta}^*)$ for the binomial populations.

3. AN ASYMPTOTIC OPTIMALITY PROPERTY

It is natural to ask how good the maximum strategy is when n increases to infinity. In the following we prove a result concerning this asymptotic behavior.

Lemma 3.1. Let $\{n_i; i=1,2,\ldots\}$ and $\{m_i; i=1,2,\ldots\}$ be increasing sequences of positive integers such that $m_i \to \infty$ and $m_i/n_i \to 0$ as $i \to \infty$. Then, $E_{\underline{\theta}}\, W(R;m_i)/n_i \to \max_{1\le i\le k} t(\theta_i)$ as $i \to \infty$ $\forall R\varepsilon C$, where for each n_i, we follow $UV(R;m_i)$ scheme.

Proof: We note that $E_{\underline{\theta}}\, W(R;m_i) = m_i \sum_{j=1}^{k} t(\theta_j) + (n_i - km_i) \sum_{j=1}^{k} \gamma_j(m_i) t(\theta_j)$.

Hence, $E_{\underline{\theta}}\, W(R;m_i)/n_i = r_i \sum_{j=1}^{k} t(\theta_j) + (1-kr_i) \sum_{j=1}^{k} \gamma_j(m_i;\theta) t(\theta_j)$ where $r_i = m_i/n_i$.

Since $r_i \to 0$ and $\gamma_k(m_i;\underline{\theta}) \to 1$ for every $\underline{\theta}\varepsilon\Omega(d)$ as $i \to \infty$. Hence, we conclude that $E_{\underline{\theta}}\, W(R;m_i)/n_i \quad \to \quad t(\theta_k) = \max_{1\le i\le k} t(\theta_i)$.

Corollary 3.1. Under the same assumptions of Lemma 3.1,

$W(R;m_i)/n_i \to \max_{1\le r\le k} t(\theta_r)$ a.s. $\forall R\varepsilon C$, where $\underline{\theta}=(\theta_1,\ldots,\theta_r,\ldots,\theta_k)$ is the true parameter.

Proof: According to the $UV(R;m_i)$ scheme we have

$$W(R;m_i)/n_i = (S_{1m_i} + S_{2m_i} + \ldots + S_{km_i})/n_i + S_{\pi(R;m_i),(n_i-km_i)}/n_i$$

$$= \sum_{j=1}^{k} \frac{S_{jm_i}}{m_i} \cdot \frac{m_i}{n_i} + \frac{S_{\pi(R;m_i),n_i-km_i}}{n_i - km_i} \cdot \frac{n_i - km_i}{n_i}$$

By the strong law of large numbers we see that $\frac{S_{jm_i}}{m_i} \to t(\theta_j)$ a.s. as $i \to \infty$, $j=1,2,\ldots,k$. Since $\frac{m_i}{n_i} \to 0$ and $\frac{n_i-km_i}{n_i} \to 1$ it follows then $W(R;m_i)/n_i \to t(\theta_j)$ a.s. for some j, as $i \to \infty$. If $t(\theta_j) \neq \max_{1\le i\le k} t(\theta_i)$, then Lemma 3.1 leads to a contradiction since $t(\theta)$ is not a constant and $\theta_{[k]} > \theta_{[i]}$ $i \neq k$.

This completes the proof.

Let $\{n_i; i=1,2,\ldots\}$ be a strictly increasing sequence of positive integers. For a fixed $R\varepsilon C$, let $\{m_i^*; i=1,2,\ldots\}$ and $\{\underline{\theta}_i^*; i=1,2,\ldots\}$ be respectively the associated maximin strategies and the least favorable configurations with respect to the reward function. Let $\theta_i^* = \{\theta_{i1}^*, \theta_{i2}^*, \ldots, \theta_{ik}^*\}$. Then, we have the following theorem.

Theorem 3.1. If $\{\theta_i^*; i=1,2,\ldots\}$ is bounded in R^k and $t(\theta_{ij}^*)$ is bounded for each $j=1,2,\ldots,k$ and $i=1,2,\ldots$, then, there exists a subsequence $\{n_{i_j}; j=1,2,\ldots\}$ of $\{n_i\}$ such that $G(m_{i_j}^*, \underline{\theta})/n_{i_j} \to \max_{1\le r\le k} t(\theta_r)$ as $j\to\infty$ $\forall \underline{\theta}\in\Omega(d)$.

Proof: (i) Let $\ell_i = [\frac{n_i}{k}]$ for $i=1,2,\ldots$. We are going to show that $\{m_i^*; i=1,2,\ldots\}$ is unbounded. Suppose there is some integer M such that $m_i^* < M$ for $i=1,2,\ldots$, then, we note that for $m_i^* < M \le \ell_i$ for sufficiently large i,

$$G(m_i^*, \underline{\theta}_i^*) < G(M, \underline{\theta}_i^*) \qquad \text{if } n_i > \frac{k \sum_{j=1}^{k} [M\gamma_j(M) - m_i^*\gamma_j(m_i^*)]t(\theta_{ij}^*)}{\sum_{j=1}^{k} [\gamma_j(M) - \gamma_j(m_i^*)]\ t(\theta_{ij}^*)}.$$

Since $t(\theta_{ij}^*)$ is bounded for all i and j, it is easy to see that there exists some i_0 such that $G(m_{i_0}^*, \underline{\theta}^*_{i_0}) < G(M, \underline{\theta}^*_{i_0})$. This contradicts the assumption that $m_{i_0}^*$ is a maximum strategy.

(ii) Suppose $\frac{m_i^*}{n_i} \to \alpha$ $(0 \le \alpha \le 1)$. If $\{\underline{\theta}_i^*\}$ does not tend to limit, we can choose a subsequence $\{\underline{\theta}_{i_j}^*; j=1,2,\ldots\}$ of $\{\underline{\theta}_i^*\}$ such that $\underline{\theta}_{i_j}^* \to \underline{\theta}^*$, say, since $\{\underline{\theta}_i^*\}$ is bounded. Then, $t(\theta_j^*)$ is finite where $\underline{\theta}^* = (\theta_1^*, \theta_2^*, \ldots, \theta_k^*)$. Then the associated subsequences $\{m_{i_j}^*\}$ and $\{n_{i_j}\}$ satisfy $m_{i_j}^* \to \infty$ and $m_{i_j}^*/n_{i_j} \to \alpha$ $(0\le\alpha\le 1)$. We note that

$$G(m_{i_j}^*, \underline{\theta}^*)/n_{i_j} \to \alpha \sum_{r=1}^{k} t(\theta_r^*) + (1-k\alpha)t(\theta_k^*), \text{ assuming}$$

$$t(\theta_k^*) = \max_{1\le i\le k} t(\theta_j^*), \text{ as } j\to\infty, \text{ since } m_{i_j}^* \to \infty \text{ and } R\varepsilon C.$$ We note that

$\alpha \sum_{r=1}^{k} t(\theta_r^*) + (1-k\alpha)t(\theta_k^*)$ is a decreasing function of α, hence, we conclude that $\alpha = 0$.

(iii) If m_i^*/n_i does not tend to a limit, we can choose subsequences $\{m_{i_j}^{\prime *}; j=1,2,\ldots\}$ and $\{n_{i_j}^{\prime}; j=1,2,\ldots\}$ of $\{m_i^*\}$ and $\{n_i\}$, respectively, so that $m_{i_j}^{\prime *}/n_{i_j} \to \beta$ $(0\le\beta\le 1)$ since $0 \le m_i/n_i \le 1$ for each i. Again, we can choose subsequences $\{m_{i_j}^*; j=1,2,\ldots\}$ and $\{n_{i_j}; j=1,2,\ldots\}$ of $\{m_{i_j}^{\prime *}\}$ and $\{n_{i_j}^{\prime}\}$ respectively so that the associated sequence $\{\theta_{i_j}^*\}$ has a limit, say, θ^* and

TABLE II

The values of G(m*,θ*) and Its Percentages of Maximum Reward for k Normal Populations.
The upper entry is the value of G(m*,θ*) associated with the maximin strategy and the lower entry is the percentage of the ratio of G(m*,θ*) over n d.

k = 2

d \ n	20	40	60	80	100
0.01	0.1007 50.3420	0.2019 50.485	0.3036 50.595	0.4055 50.687	0.5077 50.768
0.05	0.5171 51.709	1.0484 52.422	1.5890 52.967	2.1369 53.424	2.6912 53.825
0.10	1.0682 53.412	2.1929 54.823	3.3539 55.898	4.5434 56.792	5.7572 57.572
0.50	6.6090 66.090	14.2947 71.473	22.5096 75.032	31.0467 77.617	39.8072 79.615
1.00	15.4615 77.308	33.4832 83.708	52.1538 86.923	71.1690 88.961	90.3611 90.361
3.00	56.0848 93.475	115.0680 95.890	174.0511 96.695	233.6922 97.372	293.6112 97.870
5.00	94.9817 94.982	194.9613 97.481	294.9410 98.314	394.9206 98.730	495.9003 98.980

k = 3

d \ n	20	40	60	80	100
0.01	0.0672 33.613	0.1349 33.730	0.2029 33.821	0.2712 33.897	0.3396 35.964
0.05	0.3475 34.745	0.7068 35.339	1.0742 35.808	1.4480 36.200	1.8272 36.545
0.10	0.7237 36.185	1.4961 37.402	2.3018 38.364	3.1335 39.169	3.9878 39.878
0.50	4.8237 48.237	10.8718 54.359	17.6258 58.753	24.8657 62.164	32.4543 64.909
1.00	12.4329 62.164	28.3141 70.785	45.5568 75.928	63.5507 79.438	81.9454 81.945
3.00	52.4086 87.348	110.5363 92.114	168.6640 93.702	227.4188 94.758	287.2617 95.754
5.00	89.9659 89.9659	189.9259 94.963	289.8858 96.629	389.8457 97.461	489.8056 97.961

TABLE II (cont'd)

k = 5

d \ n	20	40	60	80	100
0.01	0.0404 20.175	0.0810 20.253	0.1219 20.312	0.1629 20.360	0.2040 20.404
0.05	0.2089 20.887	0.4259 21.296	0.6481 21.603	0.8742 21.856	1.1045 22.089
0.10	0.4361 21.804	0.9066 22.664	1.3985 23.309	1.9080 23.850	2.4350 24.350
0.50	3.0090 30.090	7.1242 35.621	11.8740 39.580	17.1711 42.928	22.9399 45.880
1.00	8.4055 42.027	20.9914 52.478	35.5136 59.189	51.2703 64.088	67.9090 67.909
3.00	45.5390 75.898	102.2577 85.215	158.9764 88.320	215.6950 89.873	274.6597 91.553
5.00	79.9415 79.942	179.8635 89.932	279.786 93.262	379.7075 94.927	479.6295 95.926

k = 10

d \ n	20	40	60	80	100
0.01	0.0202 10.078	0.0405 10.118	0.0609 10.148	0.0814 10.170	0.1019 10.190
0.05	0.1040 10.397	0.2119 10.595	0.3227 10.756	0.4350 10.875	0.5490 10.980
0.10	0.2163 10.814	0.4488 11.220	0.6941 11.568	0.9464 11.830	1.2057 12.057
0.50	1.4950 14.960	3.5333 17.666	6.0665 20.222	8.9862 22.465	12.2284 24.457
1.00	4.4095 22.048	11.5837 28.959	21.1674 35.279	32.3708 40.464	44.4183 44.418
3.00	30.0746 50.124	84.2237 70.186	138.3728 76.874	192.5219 80.218	246.6711 82.224
5.00	54.9171 54.917	154.7513 77.376	254.5855 84.862	354.4197 88.605	454.2540 90.851

TABLE III

For given n and d, this table gives the maximin strategy m* for k binomial populations

k = 2

d \ n	5	10	15	20	25	30
0.01	1	2	4	5	6	7
0.05	1	2	4	5	6	7
0.10	1	2	3	4	5	6
0.30	1	2	3	3	4	4
0.50	1	2	2	3	3	3

TABLE IV

$G(m^*,\underline{\theta}^*)$ and its Percentage of Maximum Reward For Two Binomial Populations

The upper entry is the value of $G(m^*,\underline{\theta}^*)$ associated with the maximin strategy and the lower entry is the percentage of ratio

d \ n	5	10	15	20	25	30
0.01	0.0751 75.150	0.1506 75.296	0.2263 75.447	0.3024 75.590	0.3786 75.725	0.4551 75.856
0.05	0.1787 59.583	0.3645 60.747	0.5566 61.843	0.7546 62.883	0.9575 63.833	1.1648 64.713
0.10	0.3150 57.273	0.6565 59.679	1.0198 61.807	1.4012 63.698	1.7977 65.370	2.2065 66.863
0.30	0.9350 60.323	2.0588 66.299	3.2753 70.437	4.5616 73.575	5.9072 76.223	7.2682 78.153
0.50	1.6750 65.686	3.7176 72.895	5.9490 77.765	8.2370 80.755	10.6217 83.307	13.0063 85.009

TABLE V

Maximum value of m for L of (2.3) for binomial populations using $R=R_{SH}$

For k = 2, d = 0.20, m* = 6 the entry 36-41 in the table, shows that for n from 36 to 41, the m*-value of L is 6.

k = 2

d \ m	1	2	3	4	5	6	8	9	10
0.05	2-8	9-14	15-20	21-26	27-32	33-38	39-51	52-57	58-60
0.10	2-8	9-14	15-20	21-26	27-32	33-39	40-52	53-59	60
0.20	2-8	9-14	15-21	22-28	29-35	36-41	42-60		
0.30	2-8	9-15	16-22	23-31	32-40	41-47	48-60		
0.50	2-9	10-19	20-32	33-49	59-60				
0.80	2-15	16-54	55-60						

k = 3

d \ m	1	2	3	4	5	6	7	8	9	10
0.05	3-8	9-17	18-29	30-38	39-47	48-55	56	57-74	75-83	84-90
0.10	3-8	9-19	20-29	30-38	39-46	47-55		56-75	76-84	85-90
0.20	3-9	10-20	21-29	30-39	40-48	49-57		58-81	82-90	
0.30	3-9	10-21	22-31	32-43	44-53	54-62		63-90		
0.50	3-10	11-24	25-40	41-61	62-89	90				
0.80	3-16	17-63	64-90							

k = 4

d \ m	1	2	3	4	5	6	8	9	10	11
0.05	4-11	12-19	20-31	32-49	50-61	62-72	73-98	99-109	110-120	
0.10	4-11	12-19	20-37	38-49	50-60	61-71	72-96	97-108	109-119	120
0.20	4-11	12-20	21-37	38-49	50-61	62-71	72-101	102-114	115-120	
0.30	4-11	12-23	24-38	39-52	53-66	67-76	77-120			
0.50	4-12	13-29	30-48	49-72	73-103	104-120				
0.80	4-18	19-70	71-120							

k = 5

d \ m	1	2	3	4	5	6	8	9	10	11
0.05	5-14	15-24	25-33	34-52	53-75	76-88	89-120	121-134	135-148	149-150
0.10	5-14	15-23	24-34	35-59	60-73	74-86	87-117	118-131	132-145	146-150
0.20	5-14	15-22	23-41	42-59	60-73	74-85	86-120	121-136	137-150	
0.30	5-13	14-24	25-45	46-61	62-77	78-90	91-140	141-150		
0.05	5-14	15-31	32-55	56-82	83-116	117-141	142-150			
0.80	5-19	20-76	77-150							

TABLE VI

The Maximum Value of L of (2.3) and Percentage of L over Maximum Reward for k Binomial Populations

The upper entry is the maximum value of L and the lower entry is the percentage of ratio of maximum of L over n d.

k = 2

d \ n	20	30	40	50	60
0.05	0.5328 53.280	0.8113 54.087	1.0952 54.760	1.3849 55.396	1.6772 55.907
0.10	1.1304 56.520	1.7428 58.093	2.3748 59.370	3.0310 60.620	3.6938 61.563
0.30	4.1123 68.538	6.4810 72.001	8.9545 74.621	11.5371 76.914	14.2245 79.025
0.50	7.7754 77.754	12.2578 81.719	16.8711 84.356	21.5215 86.086	26.2768 87.589
0.80	14.0416 87.760	21.8176 90.907	29.5936 92.480	37.3696 93.424	45.2302 94.230

k = 3

d \ n	10	30	50	70	90
0.05	0.1738 34.760	0.5505 36.700	0.9453 37.812	1.3576 38.789	1.7794 39.542
0.10	0.3635 36.350	1.2069 40.230	2.1253 42.506	3.1170 44.529	4.1461 46.068
0.30	1.2947 43.157	4.9300 54.778	9.1303 60.869	13.7980 65.705	18.7536 69.458
0.50	2.5417 50.834	10.1880 67.920	18.7015 74.806	27.6408 78.974	36.7937 81.764
0.80	5.3547 66.934	19.8416 82.673	35.0430 87.608	50.4247 90.040	66.1706 91.904

k = 5

d \ n	30	60	90	120	150
0.05	0.3249 21.660	0.6941 23.137	1.0800 24.000	1.4881 24.802	1.9067 25.423
0.10	0.7109 23.697	1.6012 26.687	2.5769 28.632	3.6452 30.377	4.7664 31.776
0.30	3.1340 34.822	7.8433 43.574	13.3900 49.593	19.9763 55.490	26.6509 59.224
0.50	7.1016 47.344	18.1168 60.389	30.4022 67.560	43.3452 72.242	47.0755 76.100
0.80	16.2175 67.573	38.1438 79.466	60.6831 84.282	83.9963 87.496	107.3095 89.425

Selection of a Restricted Subset of Normal Populations containing the One with Largest Mean*

by

Shanti S. Gupta and Thomas J. Santner

Purdue University

Division of Mathematical Sciences
Department of Statistics

Mimeograph Series #299

August 1972

*This research was supported in part by the Office of Naval Research Contract N00014-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Selection of a Restricted Subset of Normal Populations Containing the One with Largest Mean*

by

Shanti S. Gupta and Thomas J. Santner

Purdue University

1. Introduction and Terminology

One of the prime motivations for the use of subset selection procedures is to enable the experimenter to screen a group of populations selecting a subset of the best ones which will be further studied in a more intensive fashion. However, in practice, the experimenter has only limited resources to use for secondary exploration. Hence the goal in this paper is to give more flexibility to the experimenter than does the usual subset selection procedure by allowing him to specify an upper bound, m, on the number of populations included in the selected subset. Should the data clearly indicate a single population is best, this procedure still retains that advantage of the subset selection approach which would allow selection of fewer than the maximum number of populations, m. On the other hand, if the data make the choice of the best population less obvious this procedure still selects a subset for further study but guarantees that no more than m populations are selected.

---

*This research was supported in part by the Office of Naval Research Contract N00014-A-0226-00014 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Two special cases of this goal are the following. When $m = 1$ we select exactly one population and claim it is best. Such rules have been widely studied in the literature and in particular Bechhofer (1954) solves the normal means problem when the common variance is known under this formulation. When $m = k$ we select a subset whose size is a random variable $(1 \leq S \leq k)$ and claim the best population is a member of the subset. Such rules have also been widely studied in the literature and in particular Gupta (1956,65) solves the above normal means problem using such a procedure.

Formally our method of viewing the selection problem relates the subset selection formulation and the indifference zone formulation by showing both are special cases of a general theory. In practice our method allows us to blend some of the advantages of each method in the solution of the selection problem.

To fix ideas we introduce the following terminology which will distinguish the various types of rules used.

Let S be the number of populations selected by the procedure R. The goal is to select the "best" population. Ω is the set of all possible parameter configurations.

<u>Definition 1.1</u>: R is a fixed size subset rule means $\exists$ s $(1 \leq s < k)$ such that

$$P_\theta[S=s] = 1 \;\forall \theta \epsilon \Omega$$

Rules for which $s=1$ are also known as indifference zone rules and were introduced by Bechhofer (1954). In the more general case these rules were introduced by Mahamunulu (1966,67).

Definition 1.2: R is a restricted subset selection procedure means $\exists\ 1 < s < k$ such that $P_\theta[1 \leq S \leq s] = 1\ \forall\theta\epsilon\Omega$ and R is not a fixed size subset rule.

Definition 1.3: R is a subset selection procedure means $P_\theta[1 \leq S \leq k] = 1$ $\forall\theta\epsilon\Omega$ and R is neither a restricted subset selection procedure nor a fixed size subset selection procedure.

2. Statement of the Problem

Let $\pi_i \sim N(\mu_i,\sigma^2)$ for $i=1,\ldots, k$ where the common σ^2 is known. Also let $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ be the ordered means and $\pi_{(i)}$ the population with mean $\mu_{[i]}$, the best population being $\pi_{(k)}$. We assume there is no a priori knowledge concerning the pairing of the $\{\pi_{(i)}\}$ and $\{\pi_i\}$. Let $\delta \geq 0$ and

$$\Omega = \{\mu = (\mu_1,\ldots,\mu_k)\,|\,\mu_i \in (-\infty,\infty)\forall i\}$$

$$\Omega(\delta) = \{\mu \in \Omega\,|\,\mu_{[k]} - \mu_{[k-1]} \geq \delta\}$$

$$\Omega^o(\delta) = \{\mu \in \Omega(\delta)\,|\,\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta\}.$$

Goal G: Given P*, m and also possibly n and $\delta \geq 0$ define a procedure R based on a common sample size n from each population which selects a subset of the populations not exceeding m in size such that the subset contains the population $\pi_{(k)}$ and satisfies the basic probability requirement

$$P_\mu[CS|R] \geq P^*,\ \forall\mu \in \Omega(\delta) \qquad (2.1)$$

As we shall see later, by fixing δ, n, and $m < k$, the admissable range of P* values becomes

$$(1/k < P^* < (k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}[1-\Phi(t-\tfrac{\sqrt{n}}{\sigma}\delta)][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1}\,d\Phi(t)\ .$$

The event $[CS|R]$ is the selection of any subset containing $\pi_{(k)}$.

We propose the following rule based on a sample of common size n from each of the k populations. As usual let $\bar{X}_i$ be the sample mean from π_i and $\bar{X}_{[1]} \leq \cdots \leq \bar{X}_{[k]}$.

<u>Rule R</u>: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\}$ (2.2)

The following are special cases of the goal G and rule R.

A) $m = k, \ \delta = 0$

$\Omega(0) = \Omega$

$\underline{G}$: Choose a subset of $\{\pi_1, \ldots, \pi_k\}$ containing the best population such that $P_\mu[CS|R] \geq P^* \ \forall \mu \in \Omega$

$\underline{R}$: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \bar{X}_{[k]} - d\sigma/\sqrt{n}$

These are the goal and procedure studied by Gupta (1956,65).

B) $m = 1, \ \delta > 0$

$\underline{G}$: Choose a single population such that $P_\mu[CS|R] \geq P^* \ \forall \mu \in \Omega(\delta)$

$\underline{R}$: Select that population π_i corresponding to $\bar{X}_{[k]}$.

Bechhofer (1954) studied this goal and procedure.

C) $m = s(1 < s < k), \ d = +\infty$

$\underline{R}$: Select the populations corresponding to $\bar{X}_{[k-s+1]}, \ldots, \bar{X}_{[k]}$

This procedure was studied by Mahamunulu (1966,67) and Desu and Sobel (1968). The procedure is a fixed size subset type and must satisfy (2.1).

3. Probability of a Correct Selection

We introduce the following notation. For every $\ell = 1, \ldots, k$ and for every $i = k-m, \ldots, k-1$ let $\{S_j^i(\ell) : j = 1, \ldots, \binom{k-1}{i}\}$ be the collection of all subsets of size i from $\{1, \ldots, k\} - \{\ell\}$. Also let $\bar{S}_j^i(\ell) = \{1, \ldots, k\} - \{\ell\} - S_j^i(\ell)$.

<u>Theorem 3.1</u>. For any $\mu \in \Omega$, $P_\mu[CS|R] =$

$$\sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} \Phi\left(t + \frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right) \prod_{\ell \in \bar{S}_j^i(k)} \left\{\Phi\left(t + d + \frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right) - \Phi\left(t + \frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})\right)\right\} d\Phi(t)$$

Proof:

Let $\bar{X}_{(i)}$ denote the mean from population $\pi_{(i)}$, then

$$P[CS|R] = P[\bar{X}_{(k)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\}]$$

$$= P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n} \text{ for } \ell < k \text{ and } \bar{X}_{(k)} \geq \text{at least } (k-m)\ \bar{X}_{(\ell)}\text{'s with } \ell \neq k].$$

Now for every $i = k-m, \ldots, k-1$ and $j = 1, \ldots, \binom{k-1}{i}$ let

$$A_j^i = [\bar{X}_{(k)} \geq \bar{X}_{(\ell)}\ \forall \ell \in S_j^i(k) \text{ and } \bar{X}_{(k)} < \bar{X}_{(\ell)}\ \forall \ell \in \bar{S}_j^i(k)]$$

$$\Rightarrow P[CS|R] = P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n}\ \forall \ell < k \text{ and } \bigcup_{i=k-m}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i]$$

$$= \sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} P[\bar{X}_{(k)} \geq \bar{X}_{(\ell)} - d\sigma/\sqrt{n}\ \forall \ell < k \text{ and } A_j^i]$$

Now fix i and j

$$P[\overline{X}_{(k)} \geq \overline{X}_{(\ell)} - d\sigma/\sqrt{n} \ \ \forall \ell < k \text{ and } A_j^i]$$

$$= P[\overline{X}_{(k)} \geq \overline{X}_{(\ell)} \ \forall \ell \in S_j^i(k) \text{ and } \overline{X}_{(k)} < \overline{X}_{(\ell)} < \overline{X}_{(k)} + d\sigma/\sqrt{n} \ \ \forall \ell \in \overline{S}_j^i(k)]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})) \prod_{\ell \in \overline{S}_j^i(k)} \{\Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}) + d)$$

$$- \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}))\} \, d\Phi(t)$$

QED

Remark 3.1: As special cases we immediately obtain the results of Bechhofer (1954) and Gupta (1965)

A) Bechhofer ($m = 1$, $0 < d < \infty$)

$$P[CS|R] = P[\overline{X}_{(k)} \geq \overline{X}_{[k]} - d\sigma/\sqrt{n} \text{ and } A_1^{k-1}]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell=1}^{k-1} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]})) \, d\Phi(t)$$

since $A_1^{k-1} = [\overline{X}_{(k)} \geq \overline{X}_{(\ell)} \ \forall \ell < k] \subset [\overline{X}_{(k)} \geq \overline{X}_{(S)} - d\sigma/\sqrt{n} \ \forall \ell < k]$

B) Gupta ($m = k$, $0 < d < \infty$)

$$P[CS|R] = P[\overline{X}_{(k)} \geq \overline{X}_{(k)} - d\sigma/\sqrt{n} \text{ and } \bigcup_{i=0}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i]$$

$$= \int_{-\infty}^{\infty} \prod_{\ell=1}^{k-1} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[k]} - \mu_{[\ell]}) + d) \, d\Phi(t)$$

since $\bigcup_{i=0}^{k-1} \bigcup_{j=1}^{\binom{k-1}{i}} A_j^i \supset [\overline{X}_{(k)} \geq \overline{X}_{[k]} - d\sigma/\sqrt{n}]$

Remark 3.2: An application of the dominated convergence theorem shows

$$P[CS|R] \to 1 \text{ as } \mu_{[k]} - \mu_{[k-1]} \to \infty .$$

Next we determine the infimum over $\Omega(\delta)$ of the probability of a correct selection.

<u>Theorem 3.2</u>. $\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^o(\delta)} P[CS|R]$

$$\sum_{i=k-m}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} \Phi^i(t+\frac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\frac{\sqrt{n}}{\sigma}\delta)-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i} d\Phi(t)$$

Proof:

We use the following lemma due to Alam and Rizvi (1965) and also to D. Mahamunulu (1966).

Lemma

Let $X = (X_1, \ldots, X_k)$ have $k \geq 1$ independent components such that for every i, X_i has cdf $H(x_i|\theta_i)$. Suppose $\{H(x|\theta)\}$ form a stochastically increasing family. If $\phi(X)$ is a monotone function of x_i when all other components of X are held fixed, then $E_\theta[\phi(X)]$ is monotone in θ_i in the same direction.

$$\text{Let } \phi(X) = \begin{cases} 1, & \bar{X}_{(k)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise} \end{cases}$$

We claim $\phi(X)$ is non increasing in $\bar{X}_{(i)}$ for $i=1,\ldots,k-1$. Let

$$\bar{X}_{(i)} < \bar{X}'_{(i)},\ X = (\bar{X}_{(1)}, \ldots, \bar{X}_{(k)}),\ X' = (\bar{X}_{(1)}, \ldots, \bar{X}_{(i-1)}, \bar{X}'_{(i)}, \bar{X}_{(i+1)}, \ldots, \bar{X}_{(k)})$$

$$\Rightarrow \max\left\{\bar{X}_{[k]} - d\sigma/\sqrt{n},\ \bar{X}_{[k-m+1]}\right\} \leq \max\left\{\bar{X}'_{[k]} - d\sigma/\sqrt{n},\ \bar{X}'_{[k-m+1]}\right\}$$

where the primes denote the order statistics from X'. So if $\phi(X) = 0$ $\Rightarrow \phi(X') = 0$. Hence

$P_{\mu}[CS|R] = E_{\mu}[\phi(X)]$ is nonincreasing in each of $\mu_{[1]}, \ldots, \mu_{[k-1]}$ when all other means are fixed. So

$\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^{o}(\delta)} P[CS|R]$ and substituting the vector of means $(\mu_{[1]}, \ldots, \mu_{[1]}, \mu_{[1]} + \delta)$ gives the result.

QED

Remark 3.3: As special cases we get the results obtained by Gupta (1965), Bechhofer (1954) and Desu and Sobel (1968).

A) Bechhofer $(m = 1, \delta > 0)$

$$\inf_{\Omega(\delta)} P[CS|R] = \int_{-\infty}^{\infty} \Phi^{k-1}(t + \frac{\sqrt{n}}{\sigma}\delta)\, d\Phi(t)$$

B) Gupta $(m = k, 0 < d < \infty, \delta = 0)$

$$\inf_{\Omega} P[CS|R] = \sum_{i=0}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} \Phi^{i}(t)\{\Phi(t+d) - \Phi(t)\}^{k-i-1} d\Phi(t)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(t+d) \sum_{i=0}^{k-1} \binom{k-1}{i} \left[\frac{\Phi(t)}{\Phi(t+d)}\right]^{i} \left[1 - \frac{\Phi(t)}{\Phi(t+d)}\right]^{k-1-i} d\Phi(t)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(t+d)\, d\Phi(t)$$

C) Desu and Sobel $(1 \le m < k, d = +\infty, \delta > 0)$

$$\inf P[CS|R] = \int_{-\infty}^{\infty} \sum_{i=k-m}^{k-1} \binom{k-1}{i} \Phi^{i}(t + \frac{\sqrt{n}}{\sigma}\delta)\{1 - \Phi(t + \frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i} d\Phi(t)$$

$$= (k-m) \binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)][1-\Phi(t)]^{m-1}[\Phi(t)]^{k-m-1}\, d\Phi(t)$$

as we will later show.

4. Properties of R

Next we study the properties of the procedure R. To facilitate this study we let $p_\mu(i) = P_\mu[R \text{ selects } \pi_{(i)}]$ and recall the following two definitions.

Definition 4.1: R is a monotone procedure means $\forall \mu \in \Omega$ and $i < j$ $p_\mu(i) \le p_\mu(j)$.

Definition 4.2: R is an unbiased procedure means $\forall \mu \in \Omega$ and $j < k$

$$P_\mu[R \text{ does not select } \pi_{(j)}] \ge P_\mu[R \text{ does not select } \pi_{(k)}].$$

Of course R monotone $\Rightarrow$ R unbiased. Other optimal properties are

Definition 4.3: R is consistent wrt Ω' means $\lim_{n\to\infty} \inf_{\Omega'} P[CS|R] = 1$

Definition 4.4: R is strongly monotone in $\pi_{(i)}$ means

$$p_\mu(i)\begin{cases} \uparrow \text{ in } \mu_{[i]} \text{ when all other components of } \mu \text{ are fixed} \\ \downarrow \text{ in } \mu_{[j]} \text{ when all other components of } \mu \text{ are fixed } (j \neq i) \end{cases}$$

Remard 4.1: If R is non decreasing for $\pi_{(k)}$ then

$$\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^0(\delta)} P[CS|R]$$

Theorem 4.1. For every $i = 1,\dots,k$ and for all procedures R of the form (2.2), R is strongly monotone in $\pi_{(i)}$

Proof:

1) We have already shown this result for $i = k$. Since for $i < k$ we have $p_\mu(i) = E_\mu[\eta(x)]$ where $\eta(x) = \begin{cases} 1, & \bar{X}_{(i)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise,} \end{cases}$

the same argument applies to give the desired conclusion.

QED

Corollary 4.1. All rules of the form (2.2) are monotone and unbiased.

The proof follows from the definition of monotonicity and the property of being strongly monotone in $\pi_{(i)} \forall i$.

Theorem 4.2. For every rule R of form (2.2) and every $\delta > 0$, R is consistent wrt $\Omega(\delta)$.

Proof:

We must show

$$\sum_{i=k-m}^{k-1} \binom{k-1}{i} \lim_{n\to\infty} \int_{-\infty}^{\infty} \Phi^i(t+\frac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\frac{\sqrt{n}}{\sigma}\delta) - \Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\, d\Phi(t) = 1$$

We note each integrand is bounded wrt Φ measure and so dominated convergence applies.

For every $i < k-1$ we have

$$\lim_{n\to\infty} \Phi^i(t+\frac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+\frac{\sqrt{n}}{\sigma}\delta+d) - \Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i} = 0$$

and for $i = k-1$

$\lim_{n\to\infty} \Phi^{k-1}(t+\frac{\sqrt{n}}{\sigma}\delta) = 1$. Hence the result follows.

QED

This theorem says that no matter what probability requirement $(\delta > 0, P^*)$ is made and which rule is used, (2.1) can be made to hold by choosing a sufficiently large sample.

Theorem 4.3. For every n and rule R of form (2.2), $\lim_{n\to\infty} \inf_{\Omega(\delta)} P[CS|R] = 1$.

For every n, $m < k$, and $\delta > 0$, $\lim_{n\to\infty} \inf_{\Omega(\delta)} P[CS|R] =$

$$(k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}[1-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)]\ \Phi^{k-m-1}(t)\ [1-\Phi(t)]^{m-1}\ d\Phi(t)$$

Proof:

The first result follows from dominated convergence. The second result follows from the same theorem and

$$\lim_{n\to\infty}\ \inf_{\Omega(\delta)} P[CS|R] = \sum_{i=k-m}^{k-1}\binom{k-1}{i}\int_{-\infty}^{\infty}\Phi^{i}(t+\frac{\sqrt{n}}{\sigma}\delta)\ \{1-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\ d\Phi(t)$$

$$= \int_{-\infty}^{\infty}(k-m)\binom{k-1}{k-m}\int_{1-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)}^{1} y^{m-1}(1-y)^{k-m-1}dy\ d\Phi(t)$$

Letting $w = \Phi^{-1}(1-y)$ and changing the order of integration yields

$$(k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}\int_{w-\frac{\sqrt{n}}{\sigma}\delta}^{\infty} d\Phi(t)[1-\Phi(w)]^{m-1}[\Phi(w)]^{k-m-1}\ d\Phi(w)$$

QED

Remark 4.2: The first part states that by taking δ sufficiently large we can attain any P* probability requirement for any rule d based on any number of observations. The second result says that given an indifferent zone $\delta \geq 0$ and common sample size n we can not achieve all P* values. We can only attain

$$P^* < (k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}[1-\Phi(t-\tfrac{\sqrt{n}}{\sigma}\delta)][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1}d\Phi(t) < 1$$

This interpretation follows since $\inf_{\Omega(\delta)} P[CS|R]$ is a monotone non decreasing function of d.

Remark 4.3: Using the monotonicity of $\inf_{\Omega(\delta)} P[CS|R]$ we can obtain the following bounds: For $m < k$ and $d \geq 0$

$$\int_{-\infty}^{\infty}\Phi^{k-1}(t+\tfrac{\sqrt{n}}{\sigma}\delta)\, d\Phi(t) \leq \inf_{\Omega(\delta)} P[CS|R]$$

$$\leq (k-m)\binom{k-1}{k-m}\int_{-\infty}^{\infty}[1-\Phi(t-\tfrac{\sqrt{n}}{\sigma}\delta][\Phi(t)]^{k-m-1}[1-\Phi(t)]^{m-1}\, d\Phi(t).$$

For the purpose of implementing the procedure R we have prepared Table I found at the end of the paper. The body of the table contains the values of $\frac{\sqrt{n}}{\sigma}\delta$ necessary to obtain $P^* = .75, .90, .975$ using rules $d = .4, .7, 1.3$ and 1.6 for $k = 3(1)5$ with $m = 2(1)k-1$ and also for $k = 6(1)10, 15, 20$ with $m = 2(1)5$. In general given P^*, d, k and m the corresponding $\frac{\sqrt{n}}{\sigma}\delta$ is the solution of the following equation:

$$P^* = \sum_{i=k-m}^{k-1}\binom{k-1}{i}\int_{-\infty}^{\infty}\Phi^{i}(t+\tfrac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\tfrac{\sqrt{n}}{\sigma}\delta) - \Phi(t+\tfrac{\sqrt{n}}{\sigma}\delta)\}^{k-1-i}\, d\Phi(t).$$

To compare this rule to the fixed size subset rule, we have calculated

$$e(P^*,k,m,d) = \frac{n(d)}{n(\infty)}$$

where $n(d')$ is the sample size necessary to achieve probability requirement (2.1) using rule R with k,m and d'. The ratio shows the relative samples sizes of the restricted subset selection rule to the fixed size subset rule when both attain the same probability requirements. For larger d values this ratio is close to one indicating that in many cases a slight additional cost will allow use of a restricted subset selection procedure and still meet the same probability requirement. The exact savings in terms of $(m - E_\mu[S|R])$ depends of course on the underlying μ. Some exact comparisons for the equispaced means and slippage configurations will be described in the next section.

5. Expected Number of Selected Populations

As usual define

$$Y_i = \begin{cases} 1, & \bar{X}_{(i)} \geq \max\{\bar{X}_{[k-m+1]}, \bar{X}_{[k]} - d\sigma/\sqrt{n}\} \\ 0, & \text{otherwise} \end{cases}$$

which gives $S = \sum_{i=1}^{k} Y_i$ = number of populations selected.

Then $E_\mu[S] = \sum_{1}^{k} p_\mu(i)$

<u>Theorem 5.1</u>. For every $\mu \in \Omega$, $E_\mu[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^P(i)} \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})\right) \prod_{\ell \in \bar{S}_j^P(i)} \left\{\Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})+d\right)\right.$$

$$\left. - \Phi\left(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})\right)\right\} d\Phi(t)$$

Proof:

From the above discussion we see that it suffices to calculate $p_\mu(i)$ for $i = 1,\ldots,k$. Using arguments similar to those above we get

$$p_{\mu}(i) = \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} P[\bar{X}_{(i)} > \bar{X}_{(\ell)} \ \forall \ell \in S_j^P(i) \text{ and } \bar{X}_{(i)} + d\sigma/\sqrt{n} > \bar{X}_{(\ell)} > \bar{X}_{(i)} \forall \ell \in \bar{S}_j^P(i)]$$

$$= \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^P(i)} \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})) \prod_{\ell \in \bar{S}_j^P(i)} \{\Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]})+d)$$

$$- \Phi(t+\frac{\sqrt{n}}{\sigma}(\mu_{[i]} - \mu_{[\ell]}))\} \, d\Phi(t)$$

QED

Remark 5.1: $E_{\mu}[S|R] \leq m \quad \forall \mu \in \Omega$

Remark 5.2: If $m = k \Rightarrow \sup_{\Omega(\delta)} E_{\mu}[S|R] = \sup_{\Omega^{0}(\delta)} E_{\mu}[S|R]$

This was proved by Gupta (1965).

Since $E_{\mu}[S|R]$ is increasing in d the experimenter should seek to use rules with small d. On the other hand for fixed δ and P* the smaller d is the larger n must be to achieve (2.1). Hence, the experimenter must decide what trade off between n, d, and δ he is willing to accept.

To investigate his interdependence in more detail we have tabulated $E[S|R]$ under the following configurations.

A) Equispaced Means $\mu = (\alpha, \alpha+\delta, \alpha+2\delta, \ldots, \alpha+(k-1)\delta)$

Given P*, d, $\frac{\sqrt{n}}{\sigma}\delta$, k and m, Table III displays $E_{\mu}[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^P(i)} \Phi(t+\frac{\sqrt{n}}{\sigma}(i-\ell)\delta) \prod_{\ell \in \bar{S}_j^P(i)} \{\Phi(t+d+\frac{\sqrt{n}}{\sigma}(i-\ell)\delta) - \Phi(t+\frac{\sqrt{n}}{\sigma}(i-\ell)\delta)\} \, d\Phi(t)$$

B) Slippage $\mu = (\alpha, \ldots, \alpha, \alpha + \delta)$

Again given P*, d, $\frac{\sqrt{n}}{\sigma}\delta$, k and m, Table IV gives

$$E_\mu[S|R] = \sum_{p=k-m}^{k-1} \binom{k-1}{p} \int_{-\infty}^{\infty} \Phi^P(t+\frac{\sqrt{n}}{\sigma}\delta)\{\Phi(t+d+\frac{\sqrt{n}}{\sigma}\delta)-\Phi(t+\frac{\sqrt{n}}{\sigma}\delta)\}^{k-1-p} d\Phi(t)$$

$$+ (k-1)\sum_{p=k-m}^{k-1} [\binom{k-2}{p-1} \int_{-\infty}^{\infty} \Phi(t-\frac{\sqrt{n}}{\sigma}\delta)\Phi^{P-1}(t)\{\Phi(t+d)-\Phi(t)\}^{k-1-p} d\Phi(t)$$

$$+ \binom{k-2}{p} \int_{-\infty}^{\infty} \Phi^P(t)\{\Phi(t+d-\frac{\sqrt{n}}{\sigma}\delta)-\Phi(t-\frac{\sqrt{n}}{\sigma}\delta)\}\{\Phi(t+d)-\Phi^P(t)\}^{k-2-p} d\Phi(t)].$$

The same two tables also list

A) $\sum_{i=1}^{k} (k-i+1)p_\mu(i)$, the expected sum of ranks of the selected populations ($\pi_{(k)}$ is assigned rank 1 etc.) and

B) $E_\mu[S|R]/m$, the expected proportion of selected populations.

As an application of the theory we give the following example. An experimenter is sampling from nine normal populations with $\pi_i \sim N(\mu_i,1)$. He wishes to select a subset of size at most four which contains the population with largest mean. For his screening process he wishes to have a probability of correct selection at least .975 whenever $\mu_{[9]} - \mu_{[8]} \geq .8$. As $E_\mu[S|R]$ is increasing in d and n, he wishes both to be small.

Examining the four rules specified in Table I he finds his choices

are

d	.4	.7	1.3	1.6
n	18	15	11	10

He decides to base his preliminary research on a sample of size 10 from each population and uses the rule:

R: Select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[6]}, \bar{X}_{[9]} - 1.6/\sqrt{10}\}$.

6. Extension to Location and Scale Parameter Family

We assume we are given independent random variables $X_1, \ldots, X_k$ from k populations $\pi_1, \ldots, \pi_k$ with cdf's $F_{\theta_i}(x)$ where

A) $F_\theta(x) = F(x-\theta)$ $(\theta\varepsilon(-\infty,\infty))$ in the location parameter case and

B) $F_\theta(x) = F(x/\theta)$ $(F(0)=0$ and $\theta > 0)$ in the scale parameter case.

Here the θ_i's are unknown but F is known. Our goal is to select a subset of the populations not exceeding m in size such that

$$P_\theta[CS] \geq P^* \quad \forall\theta\varepsilon\Omega(\delta) \tag{6.1}$$

The event [CS] is the selection of any subset containing population $\pi_{(k)}$ and

A) $\Omega(\delta) = \{\theta | \theta_{[k]} - \theta_{[k-1]} \geq \delta\}$ in the location parameter case and

B) $\Omega(\delta) = \{\theta | \theta_{[k]}/\theta_{[k-1]} \geq \delta\}$ in the scale parameter case.

As usual there is no knowledge of the correct pairing of the $\{\pi_j\}$ and $\{\pi_{(j)}\}$. Our rules are the following:

A) R: Select $\pi_i \Leftrightarrow X_i \geq \max\{X_{[k]} - d\ ,\ X_{[k-m+1]}\}$

where $d > 0$ in the location parameter case and (6.2)

B) R': Select $\pi_i \Leftrightarrow X_i \geq \max\{c\, X_{[k]}, X_{[k-m+1]}\}$ where $0 < c < 1$

in the scale parameter case (6.3)

As the results for the present cases are completely analogous to those in the normal case we present them only for the location parameter family and then without proofs.

<u>Theorem 6.1</u>. $P_\theta[CS|R] =$

$$\sum_{i=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{i}} \int_{-\infty}^{\infty} \prod_{\ell \in S_j^i(k)} F(t+\theta_{[k]} - \theta_{[\ell]}) \prod_{\ell \in \bar{S}_j^i(k)} \{F(t+d+\theta_{[k]} - \theta_{[\ell]}) - F(t+\theta_{[k]}-\theta_{[\ell]})\}\, dF(t)$$

and further

$$\inf_{\Omega(\delta)} P[CS|R] = \inf_{\Omega^0(\delta)} P[CS|R]$$

$$= \sum_{i=k-m}^{k-1} \binom{k-1}{i} \int_{-\infty}^{\infty} F^i(t+\delta)\{F(t+d+\delta) - F(t+\delta)\}^{k-1-i}\, dF(t)$$

The optimality properties of R and R' parallel those in the normal case.

<u>Theorem 6.2</u>. For every $i = 1,\ldots,k$ and every R of the form (6.2), R is non decreasing in $\pi_{(i)}$. Hence R is monotone and unbiased.

<u>Theorem 6.3</u>. For every indifference zone $\delta \geq 0$ and $m < k$.

$$\lim_{d \to \infty} \inf_{\Omega(\delta)} P[CS|R] = \sup_{d \geq 0} \inf_{\Omega(\delta)} P[CS|R]$$

$$= (k-m)\binom{k-1}{k-m} \int_{-\infty}^{\infty} [1-F(w-\delta)][F(w)]^{k-m-1}[1-F(w)]^{m-1}\, dF(w) \qquad (6.4)$$

As before we are not able to attain all P* values for a given δ merely by choosing sufficiently large d. The right hand side of (6.4) is the upper bound on the attainable P* values.

Theorem 6.4. For every rule of form (6.2),

$$\lim_{\delta\to\infty} \inf_{\Omega(\delta)} P[CS|R] = 1.$$

Again analogous to the previous results, for any rule R and P* we can always choose an indifference zone large enough so that (6.2) holds. Finally we obtain the general expression for the expected number of populations selected.

Theorem 6.4. $E_\theta[S|R] =$

$$\sum_{i=1}^{k} \sum_{p=k-m}^{k-1} \sum_{j=1}^{\binom{k-1}{p}} \int_{-\infty}^{\infty} \prod_{\ell\in S_j^p(i)} F(t+\theta_{[i]} - \theta_{[\ell]}) \prod_{\ell\in \bar{S}_j^p(i)} \{F(t+d+\theta_{[i]} - \theta_{[\ell]})-F(t+\theta_{[i]} - \theta_{[\ell]})\} dF(t).$$

Here as before the expected savings using a restricted subset procedure over the fixed size subset procedure depends upon the underlying F(x) and parameter point in question.

<u>Table I</u>

Tables $\frac{\sqrt{n}}{\sigma}\delta$ needed to attain P* levels .75, .90, .975 for the rules given by d = .4, .7, 1.3 and 1.6

k	m \ d	.4			.7		
		P*			P*		
		.75	.90	.975	.75	.90	.975
3	2	1.078	1.891	2.766	0.863	1.645	2.520
4	2	1.342	2.123	2.998	1.156	1.906	2.781
	3	1.286	2.067	2.942	1.017	1.767	2.642
5	2	1.516	2.266	3.141	1.348	2.098	2.973
	3	1.456	2.225	3.125	1.192	1.942	2.817
	4	1.448	2.216	3.104	1.151	1.932	2.807
6	2	1.647	2.397	3.272	1.490	2.209	3.084
	3	1.591	2.341	3.216	1.335	2.054	2.929
	4	1.570	2.321	3.196	1.283	2.033	2.908
	5	1.569	2.320	3.194	1.281	2.030	2.907
7	2	1.747	2.465	3.340	1.601	2.319	3.194
	3	1.690	2.440	3.315	1.438	2.157	3.032
	4	1.680	2.430	3,305	1.389	2.139	3.014
	5	1.667	2.417	3.292	1.370	2.120	2.995
8	2	1.830	2.549	3.424	1.684	2.403	3.278
	3	1.772	2.491	3.366	1.529	2.248	3.123
	4	1.758	2.475	3.350	1.468	2.212	3.087
	5	1.755	2.474	3.347	1.462	2.181	3.056
9	2	1.906	2.625	3.500	1.766	2.453	3.328
	3	1.841	2.560	3.435	1.609	2.296	3.171
	4	1.829	2.549	3.423	1.541	2.260	3.135
	5	1.822	2.541	3.416	1.526	2.245	3.120
10	2	1.956	2.675	3.550	1.837	2.525	3.337
	3	1.884	2.603	3.478	1.666	2.385	3.198
	4	1.871	2.590	3.470	1.601	2.319	3.194
	5	1.869	2.587	3.464	1.585	2.304	3.179
15	2	2.175	2.862	3.737	2.056	2.744	3.494
	3	2.101	2.820	3.695	1.894	2.582	3.394
	4	2.086	2.798	3.673	1.825	2.542	3.391
	5	2.080	2.791	3.660	1.794	2.513	3.388
20	2	2.321	3.008	3.821	2.207	2.895	3.645
	3	2.245	2.933	3.808	2.045	2.732	3.482
	4	2.218	2.905	3.780	1.968	2.656	3.468
	5	2.213	2.900	3.775	1.935	2.622	3.449

Table I (cont.)

k	m \ d	1.3			1.6		
		p*			p*		
		.75	.90	.975	.75	.90	.975
3	2	0.559	1.340	2.215	0.464	1.246	2.121
4	2	0.943	1.662	2.537	0.884	1.634	2.509
	3	0.547	1.297	2.172	0.365	1.115	1.990
5	2	1.178	1.897	2.772	1.130	1.849	2.724
	3	0.805	1.524	2.399	0.678	1.397	2.272
	4	0.609	1.359	2.234	0.389	1.107	1.982
6	2	1.326	2.044	2.906	1.308	1.995	2.870
	3	0.992	1.679	2.554	0.889	1.576	2.451
	4	0.783	1.502	2.377	0.610	1.329	2.204
	5	0.697	1.447	2.322	0.431	1.181	2.056
7	2	1.457	2.145	3.020	1.442	2.130	2.942
	3	1.127	1.814	2.627	1.047	1.735	2.485
	4	0.930	1.649	2.524	0.772	1.459	2.334
	5	0.828	1.547	2.422	0.600	1.318	2.193
8	2	1.556	2.244	3.056	1.544	2.231	2.981
	3	1.234	1.922	2.734	1.168	1.855	2.605
	4	1.048	1.736	2.611	0.918	1.606	2.356
	5	0.925	1.644	2.519	0.724	1.442	2.255
9	2	1.645	2.332	3.082	1.619	2.307	3.059
	3	1.327	2.015	2.765	1.269	1.957	2.707
	4	1.141	1.829	2.641	1.020	1.707	2.457
	5	1.023	1.711	2.586	0.850	1.537	2.350
10	2	1.725	2.412	3.162	1.706	2.394	3.144
	3	1.407	2.049	2.844	1.367	2.024	2.774
	4	1.219	1.907	2.657	1.117	1.805	2.555
	5	1.101	1.789	2.603	0.947	1.635	2.385
15	2	1.967	2.655	3.405	1.952	2.640	3.390
	3	1.692	2.349	3.099	1.659	2.315	3.065
	4	1.516	2.173	2.923	1.449	2.106	2.856
	5	1.388	2.075	2.825	1.284	1.941	2.691
20	2	2.138	2.794	3.544	2.116	2.772	3.522
	3	1.871	2.527	3.277	1.842	2.498	3.248
	4	1.695	2.351	3.101	1.644	2.301	3.051
	5	1.578	2.234	2.984	1.492	2.149	2.898

Table II

This table gives $n(d)/n(+\infty)$ where $n(a)$ is the sample size necessary for the rule R: select $\pi_i \Leftrightarrow \bar{X}_i \geq \max\{\bar{X}_{[k]} - a\sigma/\sqrt{n},\ \bar{X}_{[k-m+1]}\}$ to satisfy $P[CS] \geq P^* \ \forall \mu \in \Omega(\delta)$.

k	m (d →)	.4		.7		1.3		1.6	
		P*		P*		P*		P*	
		.90	.975	.90	.975	.90	.975	.90	.975
3	2	2.999	1.999	2.270	1.659	1.506	1.282	1.302	1.175
4	2	1.895	1.615	1.527	1.389	1.161	1.156	1.123	1.131
	3	7.660	3.488	5.598	2.813	3.016	1.901	2.229	1.596
5	2	1.598	1.473	1.370	1.319	1.120	1.147	1.064	1.108
	3	3.459	2.480	2.634	2.015	1.623	1.402	1.363	1.311
	4	17.615	5.392	13.389	4.091	6.625	2.793	4.396	2.198
6	2	1.490	1.419	1.265	1.261	1.083	1.119	1.039	1.092
	3	2.595	2.093	1.997	1.735	1.334	1.319	1.176	1.215
	4	5.706	3.358	4.384	2.783	2.394	1.860	1.873	1.598
7	2	1.388	1.358	1.228	1.242	1.051	1.110	1.036	1.054
	3	2.234	1.921	1.746	1.607	1.235	1.207	1.129	1.080
	4	3.912	2.773	3.030	2.306	1.802	1.618	1.411	1.383
8	2	1.348	1.336	1.198	1.224	1.046	1.064	1.033	1.013
	3	1.983	1.784	1.615	1.535	1.181	1.177	1.101	1.069
	4	3.081	2.413	2.462	2.050	1.515	1.466	1.297	1.193
9	2	1.325	1.323	1.157	1.196	1.045	1.026	1.023	1.011
	3	1.859	1.714	1.495	1.461	1.151	1.110	1.086	1.064
	4	2.691	2.227	2.129	1.876	1.394	1.331	1.214	1.152

Table III

Using the rule R and under the configuration $(\alpha, \alpha+\delta, \ldots, \alpha+(k-1)\delta)$ this table gives in order the triple a) the expected number of selected populations, b) the expected sum of ranks of the selected populations and c) the expected proportion of selected populations ((a) divided by m)

Number of Populations Studied

k = 3

m	d \ $\sqrt{\frac{n}{\sigma}}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3111	1.2800	1.2237	1.1649	1.1156
		2.5300	2.1262	1.7606	1.4906	3.3121
		0.6555	0.6400	0.6118	0.5825	0.5578
	.7	1.5039	1.4588	1.3751	1.2839	1.2038
		2.9134	2.4731	2.0451	1.7073	1.4698
		0.7520	0.7294	0.6875	0.6420	0.6019

k = 4

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3619	1.3090	1.2316	1.1660	1.1157
		3.2056	2.3924	1.8184	1.4971	1.3124
		0.6810	0.6545	0.6158	0.5830	0.5578
	.7	1.5691	1.4972	1.3862	1.2855	1.2039
		3.7113	2.8056	2.1237	1.7172	1.4704
		0.7845	0.7486	0.6931	0.6427	0.6020
3	.4	1.4391	1.3629	1.2568	1.1750	1.1183
		3.3970	2.5213	1.8765	1.5173	1.3183
		0.4797	0.4543	0.4189	0.3917	0.3728
	.7	1.7789	1.6483	1.4611	1.3139	1.2126
		4.2343	3.1766	2.3037	1.7845	1.4910
		0.5930	0.5494	0.4870	0.4380	0.4042

k = 5

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3956	1.3208	1.2326	1.1660	1.1157
		3.8362	2.5299	1.8277	1.4973	1.3124
		0.6978	0.6604	0.6163	0.5830	0.5578
	.7	1.6097	1.5119	1.3875	1.2855	1.2039
		4.4502	2.9794	2.3170	1.7175	1.4704
		0.8048	0.7560	0.6938	0.6428	0.6020
3	.4	1.4995	1.3845	1.2588	1.1751	1.1183
		4.1402	2.6964	1.8893	1.5176	1.3183
		0.4998	0.4615	0.4196	0.3917	0.3728
	.7	1.8785	1.6862	1.4650	1.3140	1.2125
		5.2408	3.4475	2.3276	1.7851	1.4910
		0.6262	0.5621	0.4884	0.4380	0.4042
4	.4	1.5165	1.3920	1.2601	1.1752	1.1183
		4.1910	2.7184	1.8932	1.5180	1.3183
		0.3791	0.3480	0.3150	0.2938	0.2796
	.7	1.9571	1.7230	1.4724	1.3148	1.2126
		5.4774	3.5593	2.3499	1.7875	1.4912
		0.4893	0.4308	0.3681	0.3287	0.3031

Table IV

Using the rule R and under the configuration $(\alpha,\alpha,\ldots,\alpha+\delta)$ the table gives in order the triple a) the expected number of selected populations, b) the expected sum of ranks of the selected populations and c) the expected proportion of selected populations ((a) divided by m)

Number of Populations Studied

k = 3

m	d \ $\sqrt{\frac{n}{\sigma}}\delta$	.10	.50	.90	1.30	1.70
2	.4	1.3120	1.2996	1.2702	1.2270	1.1766
		2.5773	2.3611	2.1156	1.8627	1.6259
		0.6560	0.6498	0.6351	0.6135	0.5883
	.7	1.5052	1.4872	1.4437	1.3783	1.3003
		2.9629	2.7352	2.4657	2.1740	1.8861
		0.7526	0.7436	0.7219	0.6892	0.6502

k = 4

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3641	1.3529	1.3243	1.2792	1.2233
		3.3491	3.0598	2.7192	2.3554	2.0028
		0.6821	0.6765	0.6622	0.6396	0.6116
	.7	1.5720	1.5568	1.5169	1.4523	1.3696
		3,8654	3.5571	3.1877	2.7804	2.3685
		0.7860	0.7784	0.7585	0.7261	0.6848
3	.4	1.4423	1.4266	1.3877	1.3288	1.2583
		3.5441	3.2426	2.8768	2.4792	2.0908
		0.4808	0.4755	0.4626	0.4429	0.4194
	.7	1.7844	1.7578	1.6920	1.5915	1.4701
		4.3959	4.0606	4.6299	3.1363	2.6292
		0.5948	0.5859	0.5640	0.5305	0.4900

k = 5

m	d	.10	.50	.90	1.30	1.70
2	.4	1.3993	1.3893	1.3622	1.3172	1.2587
		4.1254	4.7752	3.3491	2.8800	2.4125
		0.6997	0.6947	0.6811	0.6586	0.6294
	.7	1.6145	1.6015	1.5653	1.5033	1.4198
		4.7653	4.3894	3.9297	3.4130	2.8799
		0.8072	0.8007	0.7827	0.7516	0.7099
3	.4	1.5055	1.4904	1.4512	1.3887	1.3108
		4.4422	4.0725	3.6089	3.0886	2.5649
		.5018	.4968	.4837	.4629	.4369
	.7	1.8882	1.8635	1.7988	1.6946	1.5627
		5.5835	5.1660	4.6218	3.9837	3.3107
		0.6294	0.6212	0.5996	0.5649	0.5209
4	.4	1.5230	1.5067	1.4649	1.3990	1.3177
		4.4949	4.1216	3.6502	3.1198	2.5859
		0.3808	0.3767	0.3662	0.3498	0.3294
	.7	1.9692	1.9392	1.8631	1.7437	1.5964
		5.8267	5.3950	4.8181	4.1356	3.4161
		0.4923	0.4848	0.4658	0.4359	0.3991

REFERENCES

[1] Alam, K. and Rizvi, M. H. (1966). Selection from Multivariate Normal Populations. Ann. Inst. Statist. Math. 18, 307-318.

[2] Bechhofer, R. E. (1954). A single sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25, 16-39.

[3] Desu, M. M. and Sobel, M. (1968). A fixed subset-size approach to the selection problem. Biometrika 55, 401-410.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill, N.C.

[5] Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 7, 225-245.

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

[7] Mahamunulu, D. M. (1966). On a generalized goal in fixed sample ranking and selection problems. Technical Report No. 72, Department of Statistics, University of Minnesota.

[8] Mahamunulu, D. M. (1967). Some fixed sample ranking and selection problems. Ann. Math. Stat. 38, 1079-91.

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