OPTIMALITY OF SUBSET SELECTION PROCEDURES FOR RANKING MEANS OF THREE NORMAL POPULATIONS*

by

Shanti S. Gupta
Purdue University

and

Klaus J. Miescke
University of Mainz

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #78-19

August 1978

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University.

OPTIMALITY OF SUBSET SELECTION PROCEDURES FOR RANKING MEANS OF THREE NORMAL POPULATIONS

by

Shanti S. Gupta
Purdue University

and

Klaus J. Miescke
University of Mainz

SUMMARY

This paper deals with the classical Gupta (1956,65) - approach ("Minimize the expected subset size under the P*-condition") in the case of three normal populations with a common known variance and equal sample sizes n.

By the method of Lagrangian (undetermined) multipliers a function (involving Φ- and φ-terms only) is derived which is a convenient tool to find optimal procedures within Seal's (1955,57) class. Numerical work together with asymptotical results lead to the conclusion that for every fixed P* and mean vector $\underline{\mu}$, Gupta's (1956) means procedure is optimal within Seal's class for sufficiently large sample size n.

Typographical Corrections

"Optimality of subset selection procedures for ranking means of three normal populations" by S. S. Gupta and K. J. Miescke, Mimeo Series #78-19.

3_8: "... where $\{i,j,k\} = \{1,2,3\}$ and h is appropriately ..."

$3_{7,6}$: insert between 3_7 and 3_6: "For a symmetric h this type of monotonicity can, equivalently, be described by"

6^1: "... of a standard normal ..."

7_6: replace "y" by "u".

12_8: replace "$P^*(b,c)$" by "$1 - P^*(b,c)$".

OPTIMALITY OF SUBSET SELECTION PROCEDURES FOR RANKING MEANS OF THREE NORMAL POPULATIONS*

by

Shanti S. Gupta
Purdue University

and

Klaus J. Miescke
University of Mainz

1. Introduction

Suppose there are k normal populations $\pi_1,\ldots,\pi_k$ with unknown means $\mu_1,\ldots,\mu_k$ and a common known variance which for convenience we assume to be unity. Further let $\bar{X}_1,\ldots,\bar{X}_k$ be the sample means of k independent samples each of size n from $\pi_1,\ldots,\pi_k$. If our goal is to select a (non-empty) subset S of $\{\pi_1,\ldots,\pi_k\}$ which contains the "best" population - i.e. the population associated with the largest mean - there are several reasonable requirements which we could impose upon such a subset selection procedure S. A classical approach due to Gupta (1956,1965) is the following:

"Minimize the expected subset size $E_{\underline{\mu}}(|S|)$ under

(1.1) $\inf_{\underline{\mu}} P_{\underline{\mu}}\{CS|S\} = P^*$, where $0 < P^* < 1$ is a predetermined constant and "CS" denotes a correct selection - i.e. the selection of any subset which includes the best population.

Now Seal (1955) proposed the following natural class $\mathcal{C}$ of procedures:

(1.2) Include π_i in the selected subset S_α i.e. $\pi_i \in S_\alpha$ iff

$$\bar{X}_i \geq \sum_{j=1}^{k-1} \alpha_j Y_j - n^{-1/2}\, c(\underline{\alpha},P^*,k),\ i = 1,\ldots,k, \text{ where}$$

$Y_1 \leq \ldots \leq Y_{k-1}$ are the ordered values of $\{\bar{X}_1,\ldots,\bar{X}_{i-1}, \bar{X}_{i+1},\ldots,\bar{X}_k\}$ and $\alpha_1,\ldots,\alpha_{k-1}$ are non-negative constants with $\alpha_1 + \ldots + \alpha_{k-1} = 1$ and $c(\underline{\alpha},P^*,k)$ is determined by (1.1), the least favorable configuration (LFC) being $\underline{\mu} = (\mu,\ldots,\mu)$, $\mu \in \mathbb{R}$.

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

If we require S to be non-empty, then c must be non-negative and therefore a lower bound (greater than k^{-1}) for P* (depending on $\underline{\alpha}$ and k) has to be observed. Let us denote this sub-class of procedures by $\mathcal{C}_+$. We shall return to this point in the next section. In the sequel let us denote $S_{\underline{\alpha}}$ by S_0, S_* and S_1 if $\underline{\alpha}$ equals to $(1,0,\ldots,0)$, $((k-1)^{-1},\ldots,(k-1)^{-1})$ and $(0,\ldots,0,1)$, respectively. If we fix P* in (0,1) $((k^{-1},1))$ and $\underline{\mu} \in \mathbb{R}^k$, then clearly there exist optimal procedures in $\mathcal{C}$ ($\mathcal{C}_+$), since $E_{\underline{\mu}}(|S_{\underline{\alpha}}|)$ is continuous in $\underline{\alpha}$ and the ranges of $\underline{\alpha}$ are compact in $\mathbb{R}^k$. But how to find them? It is conjectured by several authors that S_1 (Gupta's means-procedure) is optimal over much of the parameter space $\{\underline{\mu} | \underline{\mu} \in \mathbb{R}^k\}$, but an explicit proof of this conjecture has been missing up to now.

Seal (1957) heuristically reduced the problem in $\mathcal{C}$ to a comparison of S_0, S_* and S_1 only, and then he showed for k = 3 that in special parameter situations S_0 is inferior to S_* and S_* is inferior to S_1. Furthermore, superiority of S_1 w.r.t. other members of $\mathcal{C}_+$ was deduced by Berger (1977) heuristically, who proved that there are P*-values for which S_1 is the only procedure within $\mathcal{C}_+$ which is minimax w.r.t. the expected subset size.

Finally Deely and Gupta (1968) showed that in the special slippage configuration $\underline{\mu} = (\mu,\ldots,\mu+\delta)$, $\delta > 0$, $E_{\underline{\mu}}(|S_1|) < E_{\underline{\mu}}(|S_*|)$ if $n^{-1/2}\delta$ is greater than a constant depending on P* and k only. But as they pointed out: "Because of the difficult distribution problems involved, the general comparison of the rule R (S_1 here) in the class $\mathcal{C}_+$ is hard to make".

In this paper we extend the results of Deely and Gupta (1968) in the case of k = 3 to the following: Let ε_1 $\varepsilon(0,1)$, $\varepsilon_2 > 0$ be fixed. Then there exists a lower bound $N(\varepsilon_1,\varepsilon_2)$ for n, above which S_1 is optimal within both $\mathcal{C}_+$ and $\mathcal{C}$ for all $P^* \in [\varepsilon_1, 1-\varepsilon_1]$ and all $\underline{\mu} \in \mathbb{R}^3$ with $|\underline{\mu}| = \sqrt{\mu_1^2+\ldots+\mu_k^2} \geq \varepsilon_2$.

Especially, we can prove that S_* within $\mathcal{C}$ can never be optimal, a result which is to some extent contradictory to the results of Seal (1957).

2. A general class of procedures for k = 3.

From now on we restrict ourselves to the case of $k = 3$. Thus we start with independent variables $\bar{X}_i \sim N(\mu_i, n^{-1})$, $i = 1,2,3$. Since we restrict our considerations to procedures which are invariant under permutations as well as under common location shifts of the observables $\bar{X}_1, \bar{X}_2, \bar{X}_3$, we henceforth assume without loss of generality that $\underline{\mu} = (\mu_1, \mu_2, \mu_3) = (0, \Delta, \Delta+\delta)$, $\Delta, \delta \geq 0$, holds.

By the imposed location invariance we arrive at the maximal invariant $(\bar{X}_1-\bar{X}_2, \bar{X}_1-\bar{X}_3, \bar{X}_2-\bar{X}_3)$ as a suitable statistic, where of course one of the three differences is redundant. Let $h\colon \mathbb{R}^2 \to \mathbb{R}$ be a continuous symmetric function with the properties that $h(0,0) = 0$ and $h(s_1,s_2) \leq h(t_1,t_2)$ for all $s_1 \leq t_1$, and $s_2 \leq t_2$. Because of the permutation invariance, the only natural procedures that are appropriate are of the form:

(2.1) Include π_i in the selected subset $S_h(\bar{X}_1, \bar{X}_2, \bar{X}_3)$ iff $h(n^{1/2}(\bar{X}_i-\bar{X}_j), n^{1/2}(\bar{X}_i-\bar{X}_k)) \geq 0$ where $\{i,j,k\} = \{1,2,3\}$ where h is an appropriately chosen so that (1.1) is satisfied.

(2.2) a) $h(t_1,t_2) \leq h(t_1+u, t_2)$, $0 \leq u \leq t_2-t_1$,

b) $h(t_1,t_2) \leq h(t_1, t_2+u)$, $t_1 \leq t_2$, $u \geq 0$,

c) $h(t_1,t_2) \leq h(t_1+u, t_2+u)$, $t_1 \leq t_2$, $u \geq 0$.

Let H be an auxiliary function defined by

$H(t,v) = h(t,t+v)$, $t \in \mathbb{R}$, $v \geq 0$.

Then (2.2) rewritten in terms of the function H reads as follows:

(2.3) a) $H(t,v) \leq H(t+u,v-u)$, $0 \leq u \leq v$, $t \in \mathbb{R}$,

b) $H(t,v) \leq H(t,v+u)$, $0 \leq u$, v, $t \in \mathbb{R}$,

c) $H(t,v) \leq H(t+u,v)$, $0 \leq u$, v, $t \in \mathbb{R}$.

By c) we get the function $r(v) = \inf\{t \mid H(t,v) \geq 0\}$, $v \geq 0$, with $r(0) \leq 0$ which by b) is non-increasing and obviously has the following property:

(2.4) $H(t,v) \geq 0$ iff $t \geq r(v)$, $t \in \mathbb{R}$, $v \geq 0$.

Finally from a) and (2.4) we conclude that

(2.5) $r(v-u) \leq r(v)+u$, $0 \leq u \leq v$, holds.

Putting $c = -r(0)$ and $q(v) = r(v) + c$ we arrive at the following representation of the general class of procedures given by (2.1):

(2.6) $\pi_i \in S_q(\bar{X}_1,\bar{X}_2,\bar{X}_3)$ iff $\bar{X}_i \geq \max(\bar{X}_j,\bar{X}_k) + n^{-1/2}q(n^{1/2}|\bar{X}_j-\bar{X}_k|)-n^{-1/2}c(q,P^*)$ where $\{i,j,k\} = \{1,2,3\}$, q is a continuous non-increasing function with $q(0) = 0$ and $q(v-u) \leq q(v) + u$, $0 \leq u \leq v$ and $c(q,P^*)$ is a constant determined by (1.1).

There are of course many possible choices for q. The simplest, however, is a function q being linear, more explicitly $q(v) = av$, $v \geq 0$, with $a \in [-1,0]$. Besides we note that no other choice within the class of polynomials is possible, since there are no other polynomials satisfying $-1 \leq \frac{d}{dx} q(x) \leq 0$ for all $x \geq 0$. In this way we arrive at Seal's class of procedures, since by letting $b = a+1$, (2.6) reduces to $\mathcal{C}$ or $\mathcal{C}_+$:

(2.7) $\pi_i \in S_b(\bar{X}_1,\bar{X}_2,\bar{X}_3)$ iff $\bar{X}_i \geq b \max(\bar{X}_j,\bar{X}_k) + (1-b)\min(\bar{X}_j,\bar{X}_k)-n^{-1/2}c(b,P^*)$, where $\{i,j,k\} = \{1,2,3\}$, $b \in [0,1]$ and $c(b,P^*)$ is a constant determined by (1.1).

If b equals to 0, 1/2 or 1, S_b turns out to be S_0, S_* or S_1, respectively.

If we restrict our considerations to procedures which never select empty subsets (i.e. to $\mathcal{C}_+$), then $c(b,P^*) \geq 0$ must be observed. In this case for every $b \in [0,1]$ the possible P*-values for S_b are restricted from below by

(2.8) $$P_{(0,0,0)}\{\bar{X}_3 \geq b \max(\bar{X}_1,\bar{X}_2) + (1-b)\min(\bar{X}_1,\bar{X}_2)\}$$
$$= 2^{-1} - \pi^{-1} \text{ arc tg}(3^{-1/2}(2b-1)),\ b \in [0,1].$$

(This result also follows from (3.1) with c = 0, differentiation w.r.t. b, transformation as in (3.5) and finally by (3.7).) Besides we remark that for $P^* \geq 1/3$ (i.e. $c(1,P^*) \geq 0$) S_1 can be put in the following perhaps more familiar form:

(2.9) $\pi_i \in S_1(\bar{X}_1,\bar{X}_2,\bar{X}_3)$ iff $\bar{X}_i \geq \max(\bar{X}_1,\bar{X}_2,\bar{X}_3) - n^{-1/2}\, c(1,P^*)$.

To prove that for any fixed P* and $\underline{\mu}$, S_1 is optimal within both $\mathcal{C}$ and $\mathcal{C}_+$ for sufficiently large n, it suffices to do the following two steps:

1) To show with the method of Lagrangian (undetermined) multipliers, that for sufficiently large n no S_b with $0 < b < 1$ has an (not even local) extremal $E_{\underline{\mu}}(|S_b|)$, a fact which clearly implies monotonicity of $E_{\underline{\mu}}(|S_b|)$ in b.
2) To show that for sufficiently large n, S_1 is superior to S_0.

3. Optimality within Seal's class for k = 3.

We will now simplify the notation considerably by putting n equal to 1 in the sequel. Thereby we do not really lose any generality, since we can always transfer to cases n > 1 very quickly by only replacing $\underline{\mu}$ by $n^{1/2}\underline{\mu}$ throughout the following sections. Especially we point out that our results for "large $\Delta(\delta)$"- derived later on are to be interpreted as results for "large n and fixed (or bounded from below) $\Delta(\delta)$" in the general case.

Let φ and Φ denote the density and the c.d.f. of a standard distribution. Then we have for $\{i,j,k\} = \{1,2,3\}$ and for Z_1, Z_2, Z_3 i.i.d. $N(0,1)$:

$$P_{\underline{\mu}}\{\pi_i \in S_b(X_1,X_2,X_3)\} =$$

$$P_{\underline{\mu}}\{Z_i \geq b(Z_k + \mu_k) + (1-b)(Z_j + \mu_j) - c - \mu_i,\ Z_j+\mu_j \leq Z_k + \mu_k\} +$$

$$P_{\underline{\mu}}\{Z_i \geq b(Z_j + \mu_j) + (1-b)(Z_j + \mu_k) - c - \mu_i,\ Z_k+\mu_k \leq Z_j + \mu_j\}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\eta} [1-\Phi(b\eta + (1-b)\xi - c+\mu_k-\mu_i)]\varphi(\xi+\mu_k-\mu_j)d\xi\ \varphi(\eta)d\eta$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{\eta} [1-\Phi(b\eta+(1-b)\xi-c+\mu_j-\mu_i)]\varphi(\xi+\mu_j-\mu_k)d\xi\varphi(\eta)d\eta$$

and

$$(3.1) \qquad 2\int_{-\infty}^{\infty} \int_{-\infty}^{\eta} [1-\Phi(b\eta+(1-b)\xi-c)]\varphi(\xi)d\xi\varphi(\eta)d\eta = P^*.$$

Thus for $\underline{\mu} = (0,\Delta,\Delta+\delta)$, $\Delta,\delta \geq 0$, we have

$$(3.2) \qquad E_{\underline{\mu}}(|S_b|) =$$

$$\sum_{j=1}^{2} \sum_{i=1}^{3} \int_{-\infty}^{\infty} \int_{-\infty}^{\eta} [1-\Phi(b\eta+(1-b)\xi+A_{ij})]\varphi(\xi+B_{ij})d\xi\varphi(\eta)d\eta$$

where the A_{ij}'s and B_{ij}'s are given by Table 1 below.

Table 1	$i = 1$	$i = 2$	$i = 3$
A_{i1}	$-c+\Delta+\delta$	$-c+\delta$	$-c-\delta$
B_{i1}	δ	$\Delta+\delta$	Δ
A_{i2}	$-c+\Delta$	$-c-\Delta$	$-c-\Delta-\delta$
B_{i2}	$-\delta$	$-\Delta-\delta$	$-\Delta$

Now if $\underline{\mu} = (0,\Delta,\Delta+\delta)$ and $P^* \in (0,1)$ are fixed, b varies over [0,1] and $c(b,P^*)$ is determined by (3.1), then $E_{\underline{\mu}}(|S_b|)$ is a continuous function of $b \in [0,1]$ and therefore assumes at least one minimal value at $\tilde{b}$, say, which gives us an optimal procedure $S_{\tilde{b}}$ in $\mathbb{C}$. At this point it seems to be natural to use the method of undetermined (Lagrangian) multipliers to find such a $\tilde{b}$. (By Lebesgue's dominated convergence theorem it is easy to convince oneself that steps up to (3.4) are valid). The three equations are (3.1) and

(3.3) (a) $\frac{\partial}{\partial b} E_{\underline{\mu}}(|S_b|) = \lambda \frac{\partial}{\partial b} P_{(0,0,0)}\{CS|S_b\}$,

(b) $\frac{\partial}{\partial c} E_{\underline{\mu}}(|S_b|) = \lambda \frac{\partial}{\partial c} P_{(0,0,0)}\{CS|S_b\}$

where λ is the undetermined multiplier.

This can be put in the following form:

(3.4) (a) $$\sum_{j=1}^{2}\sum_{i=1}^{3}\int_{-\infty}^{\infty}\int_{-\infty}^{\eta}(\eta-\xi)\varphi(b\eta + (1-b)\xi + A_{ij})\varphi(\xi+B_{ij})d\xi\varphi(\eta)d\eta$$

$$= 2\lambda \int_{-\infty}^{\infty}\int_{-\infty}^{\eta}(\eta-\xi)\varphi(b\eta + (1-b)\xi-c)\varphi(\xi)d\xi\varphi(\eta)d\eta,$$

(b) $$\sum_{j=1}^{2}\sum_{i=1}^{3}\int_{-\infty}^{\infty}\int_{-\infty}^{\eta}\varphi(b\eta + (1-b)\xi + A_{ij})\varphi(\xi+B_{ij})d\xi\varphi(\eta)d\eta$$

$$= 2\lambda \int_{-\infty}^{\infty}\int_{-\infty}^{\eta}\varphi(b\eta + (1-b)\xi +c)\varphi(\xi)d\xi\varphi(\eta)d\eta.$$

By change of variables $u = \xi$ and $v = \eta-\xi$, we arrive at

(3.5) (a) $$\sum_{j=1}^{2}\sum_{i=1}^{3}\iint_{\{v>0\}} v\varphi(y+bv+A_{ij})\varphi(u+B_{ij})\varphi(u+v)d(u,v)$$

$$= 2\lambda \iint_{\{v>0\}} v\varphi(u+bv-c)\varphi(u)\varphi(u+v)d(u,v)$$

(b) $$\sum_{j=1}^{2}\sum_{i=1}^{3}\iint_{\{v>0\}} \varphi(u+bv+A_{ij})\varphi(u+B_{ij})\varphi(u+v)d(u,v)$$

$$= 2\lambda \iint_{\{v>0\}} \varphi(u+bv-c)\varphi(u)\varphi(u+v)d(u,v).$$

Clearly the next step is to eliminate λ by dividing equation (3.5) (a) by equation (3.5) (b). Moreover, we shall see that it is possible to reduce the double-integrals to relative simple terms. But we clearly point out that this is possible only in case of $k = 3$, since otherwise higher integrals are involved. We proceed now to carry out this reduction.

<u>Lemma 1</u>. <u>For all</u> $u,v,b,A,B \in \mathbb{R}$

$$\varphi(u+bv+A)\varphi(u+B)\varphi(u+v)$$

$$= \varphi(3^{1/2}u+\alpha)\varphi(\beta v+\gamma)\varphi(\varepsilon), \quad \underline{\text{where}}$$

(3.6) $$\alpha = 3^{-1/2}(A+B+(1+b)v)$$

$$\beta = (2/3)^{1/2}(b^2-b+1)^{1/2}$$

$$\gamma = 6^{-1/2}((2b-1)A-(1+b)B)(b^2-b+1)^{-1/2}$$

$$\varepsilon = 2^{-1/2}(A+(b-1)B)(b^2-b+1)^{-1/2}$$

The proof is straightforward and therefore omitted here.

<u>Lemma 2</u>. <u>For all</u> $b,A,B \in \mathbb{R}$

(3.7) $$\iint_{\{v>0\}} v\varphi(u+bv+A)\varphi(u+B)\varphi(u+v)d(u,v)$$

$$= 3^{-1/2}\beta^{-2}\,\varphi(\varepsilon)[\varphi(\gamma)-\gamma(1-\Phi(\gamma))]$$

<u>where</u> β,γ <u>and</u> ε <u>are given by</u> (3.6).

<u>Proof</u>: By lemma 1 the l.h.s. of (3.7) equals to

$$\iint_{\{v>0\}} v\varphi(3^{1/2}u+\alpha)\varphi(\beta v+\gamma)\varphi(\varepsilon)d(u,v)$$

$$= \varphi(\varepsilon)\int_{-\infty}^{\infty}\varphi(3^{1/2}u+\alpha)du\int_0^{\infty} v\varphi(\beta v+\gamma)dv$$

$$= 3^{-1/2}\varphi(\varepsilon)\int_0^{\infty} v\varphi(\beta v+\gamma)dv.$$

Substituting $w = \beta v+\gamma$ this equals

$$3^{-1/2}\varphi(\varepsilon)\beta^{-2}\int_{\gamma}^{\infty}(w-\gamma)\varphi(w)dw$$

which in turn equals the r.h.s. of (3.7).

Since the proof of the next result proceeds analogously, we omit it for brevity.

Lemma 3. For all $b,A,B \in \mathbb{R}$

$$(3.8)\qquad \iint_{\{v>0\}} \varphi(u+bv+A)\varphi(u+B)\varphi(u+v)d(u,v) = 3^{-1/2}\beta^{-1}\varphi(\varepsilon)[1-\Phi(\gamma)]$$

where β,γ and ε are given by (3.6).

To simplify the forthcoming formulas we introduce the following auxiliary function f:

$$f(x) = \varphi(x) - x(1-\Phi(x)),\ x \in \mathbb{R},$$

which is positive, strictly decreasing and convex, since $d/dx\, f(x) = \Phi(x)-1$, $x \in \mathbb{R}$, is negative and strictly increasing in x.

Now we are in a position to state our main result:

Theorem 1. Let $\underline{\mu} = (0,\Delta,\Delta+\delta)$, $\Delta,\delta \geq 0$, and $0 < P^* < 1$ be fixed. If S_b, $b \in (0,1)$, minimizes the expected subset size subject to the P^*-condition (3.1), then necessarily b and $c = c(b,P^*)$ must satisfy the equation

$$(3.9)\qquad \frac{\sum_{j=1}^{2}\sum_{i=1}^{3}\varphi(\varepsilon_{ij})f(\gamma_{ij})}{\sum_{j=1}^{2}\sum_{i=1}^{3}\varphi(\varepsilon_{ij})[1-\Phi(\gamma_{ij})]} = \frac{f(\gamma_*)}{1-\Phi(\gamma_*)}$$

where the ε_{ij}'s and γ_{ij}'s are given by Table 2 below, and $\gamma_* = 6^{-1}\sum_{j=1}^{2}\sum_{i=1}^{3}\gamma_{ij} = 6^{-1/2}(b^2-b+1)^{-1/2}(1-2b)c$.

Table 2	i = 1	i = 2	i = 3
$3^{1/2}\rho\gamma_{i1}+(2b-1)c$	$(2b-1)\Delta-(2-b)\delta$	$-(b+1)\Delta-(2-b)\delta$	$-(b+1)\Delta-(2b-1)\delta$
$\rho\varepsilon_{i1}+c$	$\Delta+b\delta$	$-(1-b)\Delta+b\delta$	$-(1-b)\Delta-\delta$
$3^{1/2}\rho\gamma_{i2}+(2b-1)c$	$(2b-1)\Delta+(b+1)\delta$	$(2-b)\Delta+(b+1)\delta$	$(2-b)\Delta-(2b-1)\delta$
$\rho\varepsilon_{i2}+c$	$\Delta+(1-b)\delta$	$-b\Delta+(1-b)\delta$	$-b\Delta-\delta$

In the above table, $\rho = 2^{1/2}\,(b^2 - b + 1)^{1/2}$

Corollary 1. S_* is not optimal in $\mathbb{C}$ except when $\underline{\mu} = (0,0,0)$.

Proof: Since $b = 1/2$, we have $\gamma_* = 0$ and $\gamma_{i1} = -\gamma_{i2}$, $\varepsilon_{i1} = \varepsilon_{i2}$, $i = 1,2,3$. Therefore the l.h.s. of (3.9) reduces to

$$\frac{\sum_{i=1}^{3}\varphi(\varepsilon_{i1})[f(\gamma_{i1})+f(\gamma_{i2})]}{\sum_{i=1}^{3}\varphi(\varepsilon_{i1})[1-\Phi(\gamma_{i1})+1-\Phi(\gamma_{i2})]}$$

$$= \sum_{i=1}^{3}\varphi(\varepsilon_{i1})[f(\gamma_{i1}) + f(-\gamma_{i1})]/\sum_{i=1}^{3}\varphi(\varepsilon_{i1}).$$

By the convexity of f this last expression is strictly greater than

$$\sum_{i=1}^{3}\varphi(\varepsilon_{i1})2f(0)/\sum_{i=1}^{3}\varphi(\varepsilon_{i1}) = 2f(0) \quad \text{if } \underline{\mu} \neq (0,0,0).$$

But the r.h.s. of (3.9) equals $2f(0)$ since $\gamma_* = 0$. Thus (3.9) holds only in the case $\underline{\mu} = (0,0,0)$, where, of course, all procedures in $\mathbb{C}$ are optimal (since $E_{(0,0,0)}(|S_b|) \equiv 3P^*$).

Though we cannot get explicit solutions by using (3.9), some more analysis is possible.

In our asymptotic considerations in Section 5 we restrict ourselves to the three types of parametric-configurations:

(I) $\Delta = 0$ i.e. $\underline{\mu} = (0,0,\delta)$, $\delta > 0$,

(II) $\Delta = \delta$ i.e. $\underline{\mu} = (0,\delta,2\delta)$, $\delta > 0$ and

(III) $\delta = 0$ i.e. $\underline{\mu} = (0,\Delta,\Delta)$, $\Delta > 0$,

It should be noted that our numerical studies in the next section are performed without the above restrictions.

4. Numerical results.

Let us denote the difference of the l.h.s. and the r.h.s. of (3.9) by $G(b,c,\Delta,\delta)$. We computed some interesting G-values as follows:

(i) $b = 0.1\ (0.1)\ 0.9$.

(ii) $c = 0\ (0.5)\ 4$.

In fact, we cover all values of $c = -4\ (0.5)\ 4$, since there is a symmetry in our problem. More precisely, we have for all $\underline{\mu}$, i, b, c, Δ and δ

(4.1) $P_{\underline{\mu}}\{\pi_i \in S_{b,c}\} = 1-P_{-\underline{\mu}}\{\pi_i \in S_{1-b,-c}\}$,

which implies

(4.2) $P^*(b,c) = 1 - P^*(1-b,-c)$

and

(4.3) $E_{(0,\Delta,\Delta+\delta)}(|S_{b,c}|) = 3-E_{(0,\delta,\delta+\Delta)}(|S_{1-b,-c}|)$.

It is easy to see (cf. Table 2) that

(4.4) $G(b,c,\Delta,\delta) = G(1-b,-c,\delta,\Delta)$ for all b,c,Δ,δ.

Remark: With our choice of c-values we cover for every b at least a P^*-interval of [0.005, 0.995], since $P^*(1,4) > 0.995$. As to the configuration of the means, we studied the cases $\underline{\mu} = (0,0,\delta)$ (i.e. $\Delta = 0$) and $\underline{\mu} = (0,\Delta,\Delta+t\Delta)$ (i.e. $\delta = t\Delta$) for

(iii) t = 0 (0.1) 0.5 (0.5) 2 and

(iv) Δ = 0.25 (0.25) 2 (1) 9.

It suffices to study the above configurations, since we are in fact interested in the general case, where $\underline{\mu}$ is fixed (or bounded away from 0) and $n \geq 1$ varies.

Our numerical findings are as follows: In most situations including $\Delta = 0$ we have positive G-values. Negative values occur only when $b > 0.5$, $c > 2$ and $t < 0.5$. But, and this should be emphasized, we also found that for every such (b,c,t) - point there is a lower bound for Δ (increasing in c), beyond which we merely have positive G-values.

Morever, the (large Δ resp. δ) ends of our tables clearly indicate the beginning of the asymptotic behavior of G, which we shall study in the next section. Thus there seems to be no gap between our numerical and our asymptotical results. And therefore we believe that our assertion stated toward the end of Section 1 is sufficiently well confirmed.

Finally we remark that the interesting question that remains open is what really happens in cases where G = 0 occurs. Because of (4.2) and (4.3) we know that if at a certain $\underline{\mu} = (0,\Delta,\Delta+\delta)$, $S_{b,c}$ is optimal for $P^*(b,c)$, then at $\underline{\mu} = (0,\delta,\delta+\Delta)$ $S_{1-b,-c}$ is worst for $P^*(b,c)$. But we do not know whether there is any extreme at all. Some first attempts to find an answer with the help of Monte-Carlo-experiments did not lead to a definitive conclusion, but it seems worthwhile to study this point in more detail at another occasion.

5. Asymptotic results in cases (I) - (III).

Theorem 2. Let P^* and b in (0,1) and therefore $c(b,P^*)$ be fixed.

Then in all cases (I) - (III) the l.h.s. of (3.9) tends to infinity if δ or Δ tends to infinity.

Proof: Since all terms $\varphi(\varepsilon_{ij})$, $f(\gamma_{ij})$, $1-\Phi(\gamma_{ij})$, $i = 1,2,3$, $j = 1,2$, are positive, it suffices to show that

$$(5.1) \qquad \frac{\varphi(\varepsilon_{ij})\,[1-\Phi(\gamma_{ij})]}{\varphi(\varepsilon_{21})f(\gamma_{21})} \xrightarrow[\delta \text{ or } \Delta \to \infty]{} 0 \text{ for all } (i,j).$$

Now this evidently holds for $(i,j) = (2,1)$, since in all cases (I)-(III) γ_{21} tends to $-\infty$ and therefore $f(\gamma_{21})$ tends to ∞. The remaining five terms are studied separately for cases (I) - (III) in the sequel.

Case (I) $\Delta = 0$, $\underline{\mu} = (0,0,\delta)$.

Here the six pairs $(\varepsilon_{ij}, \gamma_{ij})$ reduce to three distinct pairs (cf. Table 2). Thus we have to prove (5.1) for $(i,j) = (3,1)$, (1.2) only. By $\varphi(\varepsilon_{31})/\varphi(\varepsilon_{21}) = \varphi((-c-\delta)/\rho)/\varphi((-c+b\delta)/\rho)$

$$= \exp\{(-1/2\rho^2)[(1-b^2)\delta^2+2c(1+b)\delta]\} \xrightarrow[\delta \to \infty]{} 0 \text{ and}$$

$0 \leq [1-\Phi(\gamma_{31})]/f(\gamma_{21}) \leq f(\gamma_{21})^{-1} \xrightarrow[\delta \to \infty]{} 0$ we see that (5 .1) holds for $(i,j) = (3.1)$.

For $(i,j) = (1,2)$ we make use of the following inequalities (cf. Feller (1968): "Large Deviations")

$$(5.2) \qquad 1-\Phi(x) \leq \varphi(x)/x \quad \text{for all } x > 0.$$

$$(5.3) \qquad f(x) = \varphi(x) - x(1-\Phi(x)) \geq -x \text{ for all } x < 0.$$

Now if δ is sufficiently large we have $\gamma_{12} > 0$ and $\gamma_{21} < 0$ and therefore by (5.2) and (5.3)

$$0 \leq \varphi(\varepsilon_{12})[1-\Phi(\gamma_{12})]/\varphi(\varepsilon_{21})f(\gamma_{21})$$

$$\leq (-1/\gamma_{12}\gamma_{21})\varphi(\varepsilon_{12})\varphi(\gamma_{12})/\varphi(\varepsilon_{21}).$$

Clearly we have $-1/\gamma_{12}\gamma_{21} \xrightarrow[\delta \to \infty]{} 0$. And the asymptotic behavior of $\varphi(\varepsilon_{12})\varphi(\gamma_{12})/\varphi(\varepsilon_{21})$ can be found by looking at the corresponding term associated with δ^2:

$$\exp\{-(b-2)^2\delta^2/6\rho^2\} \xrightarrow[\delta \to \infty]{} 0.$$

Thus (5.1) also holds for $(i,j) = (1,2)$.

Case (II) $\Delta = \delta$, $\underline{\mu} = (0,\delta,2\delta)$.

Here we have $3^{1/2}\rho\,\gamma_{21} + (2b-1)c = 3(b-1)\delta$ and for $i = 1,2,3$

$\rho\varepsilon_{i1}+c$: $(b+1)\delta$, $(2b-1)\delta$, $(b-2)\delta$

$\rho\varepsilon_{i2}+c$: $(2-b)\delta$, $(1-2b)\delta$, $-(b+1)\delta$

By $|2b-1| < |b+1|$, $|2-b|$ we have for all $(i,j) \neq (2,1)$

$\varphi(\varepsilon_{ij})/\varphi(\varepsilon_{21}) \xrightarrow[\delta \to \infty]{} 0$. Thus the proof is completed by noting that

$0 \leq [1-\Phi(\gamma_{ij})]/f(\gamma_{21}) \leq f(\gamma_{21})^{-1} \xrightarrow[\delta \to \infty]{} 0$ holds.

Case (III) $\delta = 0$, $\underline{\mu} = (0,\Delta,\Delta)$.

As in case (I) we have to prove (5.1) only for two pairs: this time for $(i,j) = (1,1)$ and $(2,2)$. By

$$\varphi(\varepsilon_{11})/\varphi(\varepsilon_{21}) = \varphi((-c+\Delta)/\rho)/\varphi((-c-(1-b)\Delta)/\rho)$$

$$= \exp\{(-1/2\rho^2)[b(2-b)\Delta^2 - 2(2-b)c\Delta]\} \xrightarrow[\Delta \to \infty]{} 0 \text{ and}$$

$$0 \leq [1-\Phi(\gamma_{11})]/f(\gamma_{21}) \leq f(\gamma_{21})^{-1} \xrightarrow[\Delta \to \infty]{} 0$$

we see that (5.1) holds for $(i,j) = (1,1)$.

For $(i,j) = (2,2)$ we proceed as in case (I): For sufficiently large Δ, we have $\gamma_{22} > 0$ and $\gamma_{21} < 0$ and therefore

$$0 \leq \varphi(\varepsilon_{22})[1-\Phi(\gamma_{22})]/\varphi(\varepsilon_{21})f(\gamma_{21})$$

$$\leq (-1/\gamma_{22}\gamma_{21})\varphi(\varepsilon_{22})\varphi(\gamma_{22})/\varphi(\varepsilon_{21}).$$

Clearly $-1/\gamma_{22}\gamma_{21} \xrightarrow[\Delta \to \infty]{} 0$. And since the term of $\varphi(\varepsilon_{22})\varphi(\gamma_{22})/\varphi(\varepsilon_{21})$ associated with Δ^2 turns out to be $\exp\{(-1/6\rho^2)(b+1)^2\Delta^2\} \xrightarrow[\Delta \to \infty]{} 0$, (5.1) also holds for $(i,j) = (2,2)$.

This completes the proof of Theorem 2.

Remark. By similar arguments one can prove, that in the general case of $\underline{\mu} = (0,\Delta,\Delta+\delta)$ the l.h.s. of (3.9) tends to infinity if Δ is bounded and $\delta \to \infty$ or if δ is bounded and $\Delta \to \infty$.

6. Comparison of S_0 with S_1.

Theorem 3. *Let* $0 < P^* < 1$ *and* $\underline{\mu} = (0,\Delta,\Delta+\delta) \neq (0,0,0)$ *be fixed. Then the following inequalities hold*:

(6.1) (a) $E_{\underline{\mu}}(|S_1|) < \Phi(2^{-1/2}(c_1-\Delta-\delta))+\Phi(2^{-1/2}(c_1-\delta))+\Phi(2^{-1/2}(c_1+\delta))$

(b) $E_{\underline{\mu}}(|S_0|) > \Phi(2^{-1/2}(c_0-\Delta)) + \Phi(2^{-1/2}(c_0+\Delta))+\Phi(2^{-1/2}(c_0+\Delta+\delta))$

where $c_0 < c_1$. c_0 *and* c_1 *correspond to* S_0 *and* S_1 *and are determined by* (3.1).

Proof: For $Z_i = X_i - \mu_i$, $i = 1,2,3$ being standard normal we have

$$\begin{aligned}
E_{\underline{\mu}}(|S_1|) &= \sum_{i=1}^{3} P_{\underline{\mu}}\{\pi_i \in S_1\} \\
&= P\{Z_1+c_1 \geq Z_2+\Delta, Z_3+\Delta+\delta\} \\
&\quad + P\{Z_2+\Delta+c_1 \geq Z_1, Z_3+\Delta+\delta\} \\
&\quad + P\{Z_3+\Delta+\delta+c_1 \geq Z_1, Z_2+\Delta\} \\
&< P\{Z_1+c_1 \geq Z_3+\Delta+\delta\} \\
&\quad + P\{Z_2+\Delta+c_1 \geq Z_3+\Delta+\delta\} \\
&\quad + P\{Z_3+\Delta+\delta+c_1 \geq Z_2+\Delta\} \\
&= \Phi(2^{-1/2}(c_1-\Delta-\delta))+\Phi(2^{-1/2}(c_1-\delta))+\Phi(2^{-1/2}(c_1+\delta)).
\end{aligned}$$

$$\begin{aligned}
E_{\underline{\mu}}(|S_0|) &= \sum_{i=1}^{3} P_{\underline{\mu}}\{\pi_i \in S_0\} \\
&= 1-P\{Z_1+c_0 \leq Z_2+\Delta, Z_3+\Delta+\delta\} \\
&\quad + 1-P\{Z_2+\Delta+c_0 \leq Z_1, Z_3+\Delta+\delta\} \\
&\quad + 1-P\{Z_3+\Delta+\delta+c_0 \leq Z_1, Z_2+\Delta\} \\
&> 1-P\{Z_1+c_0 \leq Z_2+\Delta\}
\end{aligned}$$

$$+ 1-P\{Z_2+\Delta+c_0 \leq Z_1\}$$
$$+ 1-P\{Z_3+\Delta+\delta+c_0 \leq Z_1\}$$
$$= \Phi(2^{-1/2}(c_0-\Delta)) + \Phi(2^{-1/2}(c_0+\Delta))+\Phi(2^{-1/2}(c_0+\Delta+\delta)).$$

Corollary 2. For P* fixed and Δ or δ or both sufficiently large we have $E_{\underline{\mu}}(|S_1|) < E_{\underline{\mu}}(|S_0|)$.

Proof: For δ bounded we have

$$\lim_{\Delta\to\infty} E_{\underline{\mu}}(|S_1|) = \Phi(2^{-1/2}(c_1-\delta))+\Phi(2^{-1/2}(c_1+\delta)) \quad \text{and}$$

$$\lim_{\Delta\to\infty} E_{\underline{\mu}}(|S_0|) = 2,$$

For Δ bounded we have

$$\lim_{\delta\to\infty} E_{\underline{\mu}}(|S_1|) = 1 \quad \text{and}$$

$$\lim_{\delta\to\infty} E_{\underline{\mu}}(|S_0|) = 1+\Phi(2^{-1/2}(c_0-\Delta))+\Phi(2^{-1/2}(c_0+\Delta)).$$

Finally if Δ and δ are unbounded we have

$$\lim_{\Delta,\delta\to\infty} E_{\underline{\mu}}(|S_1|) = 1 \quad \text{and}$$

$$\lim_{\Delta,\delta\to\infty} E_{\underline{\mu}}(|S_0|) = 2.$$

Corollary 3. Let P* be fixed. If $\underline{\mu} = (0,\Delta,\Delta+\delta)$ satisfies $\Delta,\delta > c_1-c_0$ then $E_{\underline{\mu}}(|S_1|) < E_{\underline{\mu}}(|S_0|)$.

Proof:

$\delta > c_1-c_0$ implies $\Phi(2^{-1/2}(c_1-\Delta-\delta)) < \Phi(2^{-1/2}(c_0-\Delta))$, $\Delta+\delta > c_1-c_0$ implies $\Phi(2^{-1/2}(c_1-\delta)) < \Phi(2^{-1/2}(c_0+\Delta))$ and $\Delta > c_1-c_0$ implies $\Phi(2^{-1/2}(c_1+\delta)) < \Phi(2^{-1/2}(c_0+\Delta+\delta))$. At this point we should remind ourselves that our results derived for n = 1 properly modified (replacing $\underline{\mu}$ by $n^{1/2}\underline{\mu}$) also hold true for $n > 1$. Thus for example the last corollary in the general case reads as follows: "Let P* and $\underline{\mu} = (0,\Delta,\Delta+\delta)$ with $\Delta,\delta > 0$ be fixed. If $n > ((c_1-c_0)/\min(\Delta,\delta))^2$ then we have $E_{\underline{\mu}}(|S_1|) < E_{\underline{\mu}}(|S_0|)$".

Finally, we remark that we are aware of the fact, that we studied a problem (optimality within Seal's class for k = 3), which on the one hand is well known and established in literature, but which on the other hand is only part of the more general problem searching for an optimal function q in (2.6). As a matter of fact there are no results in this direction till now (also none for distributions other than normal distribution) except only in Studden (1967) (cf. corollary (2.1) there), where the k densities, however, are assumed to be known in advance and P* is fixed on the corresponding parameter space. Admittedly solutions of problem (1.1) are hard to find. It should be pointed out that recently some optimality problems have been studied in decision-theoretic and, especially, in Bayesian framework.

REFERENCES

[1] Berger, R. L. (1977). Minimax, admissible and Γ-minimax multiple decision rules. Mimeo Series No. 489, Dept. of Statist., Purdue Univ., W. Lafayette, IN.

[2] Deely, J. J. and Gupta, S. S. (1968). On the properties of subset selection procedures. Sankhyā Ser. A 30, 37-50.

[3] Feller, W. (1968). An Introduction to Probability Theory and Its Applications. Vol. I, 3rd ed., Wiley, New York.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeo Series No. 150, Inst. of Statist., Univ. of North Carolina, Chapel Hill, North Carolina.

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

[6] Seal, K. C. (1955). On a class of decision procedures for ranking means of normal populations. Ann. Math. Statist. 26, 387-398.

[7] Seal, K. C. (1957). An optimum decision rule for ranking means of normal populations. Calcutta Statist. Assoc. Bull. 7, 131-150.

[8] Studden, W. J. (1967). On selecting a subset of k populations containing the best. Ann. Math. Statist. 38,1072-1078.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Mimeograph Series #78-19	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Optimality of subset selection procedures for ranking means of three normal populations	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: Mimeo. Series #78-19
7. AUTHOR(s): Shanti S. Gupta and Klaus J. Miescke	8. CONTRACT OR GRANT NUMBER(s): ONR N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, W. Lafayette, IN 47907	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, DC	12. REPORT DATE: August 1978
	13. NUMBER OF PAGES: 18
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Subset selection, normal populations, expected size, optimality, ranking of means.

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

This paper deals with the classical Gupta (1956,65)-approach ("Minimize the expected subset size under the P*-condition") in the case of three normal populations with a common known variance and equal sample sizes n.

By the method of Lagrangian (undetermined) multipliers a function (involving Φ- and φ-terms only) is derived which is a convenient tool to find optimal procedures within Seal's (1955,57) class. Numerical work together with asymptotical results lead to the conclusion that for every fixed P* and mean vector $\underline{\mu}$, Gupta's

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(1956) means procedure is optimal within Seal's class for sufficiently large sample size n.