AN ESTIMATION PROBLEM RELATING TO SUBSET SELECTION FOR NORMAL POPULATIONS

S. Jeyaratnam and S. Panchapakesan[1]
Department of Mathematics
Southern Illinois University
Carbondale, Illinois, U.S.A.

Technical Report #83-38

Department of Statistics
Purdue University

August 1983

[1]The research of this author was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University and reproduction for any purpose of the United States Government is permitted.

AN ESTIMATION PROBLEM RELATING TO SUBSET SELECTION FOR NORMAL POPULATIONS

S. Jeyaratnam and S. Panchapakesan[1]
Department of Mathematics
Southern Illinois University
Carbondale, Illinois, U.S.A.

I. INTRODUCTION

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k independent normal populations with unknown means $\theta_1, \theta_2, \ldots, \theta_k$, respectively, and a common known variance τ^2. The population associated with the largest θ_i is called the <u>best</u> population. In the subset selection approach, we want to select a nonempty subset of the k populations so that it includes the best population with a minimum guaranteed probability $P^*(1/k < P^* < 1)$. The basic idea of the subset selection approach is that the number of populations to be selected should depend upon the evidence supplied by the data. The size of the selected subset depends on the sample size and the confidence level P^* associated with the claim that a <u>correct selection</u> (i.e. selection of any subset that contains the best) occurs. It can be said that subject to the P^*-value and the sample size, we cannot make finer distinction among the populations that are selected in seeking the best. In this case, one may decide to use the selected populations in equal proportions in the future. In this sense, the average worth of the selected subset is given by

[1]The research of this author was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University and reproduction for any purpose of the United States Government is permitted.

(1.1) $$M = \sum_{i \in S} \theta_i \Big/ \sum_{i \in S} I_i$$

where S denotes the set of indices of the selected populations, and $I_i = 1$ or 0 according as π_i is or is not included in the selected subset. Our interest is to estimate M, which we call the mean of the selected subset. It is important to note that M is a random variable.

In this paper, we consider the subset selection rule of Gupta [4], [5], which has known optimality properties; see Gupta and Panchapakesan [6], and Berger and Gupta [1]. Let $Y_1, Y_2, \ldots, Y_k$ be the sample means based on n independent observations from each population. The rule R of Gupta [5] is:

"Select π_i if and only if $Y_i \geq \max_{1 \leq j \leq k} Y_j - d\tau/\sqrt{n}$"

where $d > 0$ is to be determined such that the probability of a correct selection (PCS) is P*. This value of d is shown to be given by

(1.2) $$\int_{-\infty}^{\infty} \Phi^{k-1}(t+d)\, \varphi(t)\, dt = P^*,$$

where (here and in the sequel) Φ and φ denote the standard normal cdf and density function respectively.

Our present investigations relate to only the case of $k = 2$. The following notations hold for the entire paper:

(1.3) $$\begin{cases} \sigma^2 = 2\tau^2/n \;; \quad c = d/\sqrt{2} \;; \quad \delta = (\theta_1-\theta_2)/\sigma \;; \\ \theta = \theta_1-\theta_2 \;; \quad \theta^* = \theta_1+\theta_2 \;; \\ Y = Y_1-Y_2 \;; \quad Y^* = Y_1+Y_2 \;; \\ I_1 = I_{\{Y>c\sigma\}} \;; \quad I_2 = I_{\{Y<-c\sigma\}} \;; \\ \text{where } I_A \text{ denotes the indicator function of the set A.} \end{cases}$$

Now, for k = 2, we get

$$c = \Phi^{-1}(P^*), \tag{1.4}$$
$$M = \theta_1 I_1 + \theta_2 I_2 + \frac{\theta_1+\theta_2}{2} \; (1-I_1-I_2) \; .$$

When c = 0, the rule R selects the population that yields the largest sample mean and M is the mean of the selected population. Of course, in this case, the minimum PCS cannot be guaranteed for $P^* > 1/k$ unless additional modifications are made in the formulation of the selection problem. This is the aspect not considered by those who discussed the estimation of M in this case; these are Sarkadi [8] and Dahiya [3] for k = 2 and known τ; Hsieh [7] for k = 2 and unknown τ; and Cohen and Sackrowitz [2] for $k \geq 2$ and known τ.

For any estimator $\hat{M}$ of M, the bias $B(\hat{M}) = E(T-M)$ and the mean squared error $MSE(\hat{M}) = E(\hat{M}-M)^2$. It can be shown (Theorem 2.1) that no unbiased estimator of M, having a finite variance, exists. In Section II we define the 'natural' estimator T and three classes of estimators $T_{1\lambda}$, $T_{2\lambda}$ and $T_{3\lambda}$, obtained by making adjustments for the bias of T. The biases and the MSEs of $T_{1\lambda}$, $T_{2\lambda}$ and $T_{3\lambda}$ are discussed in Sections II and III respectively. Numerical comparisons of the performances of these estimators are made in Section IV.

II. THE ESTIMATORS AND THEIR BIASES

Since $Y \sim N(\theta,\sigma^2)$, it is easy to see that

$$E(M) = \theta_1\{1-\Phi(c-\delta)\} + \theta_2\{1-\Phi(c+\delta)\} + \frac{\theta_1+\theta_2}{2}\{\Phi(c+\delta)+\Phi(c-\delta)-1\} \tag{2.1}$$

$$= \frac{1}{2}[\theta^* + \delta\sigma\{\Phi(c+\delta) - \Phi(c-\delta)\}].$$

Theorem 2.1. No unbiased estimator of M, with finite variance, exists.

Proof. See Appendix A1. □

Let us now consider the 'natural' estimator

$$(2.2) \qquad T = \begin{cases} \max(Y_1, Y_2) & \text{if } |Y| < c\sigma, \\ \frac{Y^*}{2} & \text{otherwise.} \end{cases}$$

Since $\max(Y_1,Y_2) = \frac{1}{2}\{Y^* + |Y|\}$, we obtain

$$E(T) = \frac{\theta^*}{2} + \frac{1}{2} \int_{-\infty}^{-c\sigma} |y| \frac{1}{\sigma} \varphi\left(\frac{y-\theta}{\sigma}\right) dy + \frac{1}{2} \int_{c\sigma}^{\infty} |y| \frac{1}{\sigma} \varphi\left(\frac{y-\theta}{\sigma}\right) dy .$$

By changing the variable of integration by setting $t = (y-\theta)/\sigma$ and using (A2.1) in Appendix 2, E(T) simplifies to

$$(2.3) \qquad E(T) = \frac{\theta^*}{2} + \frac{\delta\sigma}{2} \{\Phi(c+\delta) - \Phi(c-\delta)\} + \frac{\sigma}{2} \{\varphi(c+\delta) + \varphi(c-\delta)\}$$

$$= E(M) + \frac{\sigma}{2} \{\varphi(c+\delta) + \varphi(c-\delta)\} .$$

Thus $\quad B(T) = \frac{\sigma}{2} \{\varphi(c+\delta) + \varphi(c-\delta)\}$.

2.1 Estimators $T_{1\lambda}$

Since the bias of T is positive, we define $T_{1\lambda}$ by

$$(2.4) \qquad T_{1\lambda} = T - \frac{\lambda\sigma}{2} \{\varphi(c+\frac{Y}{\sigma}) + \varphi(c-\frac{Y}{\sigma})\} , \quad \lambda \geq 0.$$

The bias of $T_{1\lambda}$ is

$$(2.5)\qquad B(T_{1\lambda}) = E[T-M - \frac{\lambda\sigma}{2}\{\varphi(c+\frac{Y}{\sigma}) + \varphi(c-\frac{Y}{\sigma})\}]$$

$$= \frac{\sigma}{2}\{\varphi(c+\delta) + \varphi(c-\delta)\} - \frac{\lambda\sigma}{2\sqrt{2}}\{\varphi(\frac{c+\delta}{\sqrt{2}}) + \varphi(\frac{c-\delta}{\sqrt{2}})\},$$

by using (2.3) and (A2.7). It should be noted that $T_{1\lambda}$ becomes T when $\lambda = 0$, and that it reduces to $\hat{M}_\lambda$ of Dahiya [3] when $c = 0$.

2.2 Estimators $T_{2\lambda}$

To motivate the definition of $T_{2\lambda}$, $\lambda \geq 0$, consider the following estimator

$$(2.6)\qquad U = \frac{Y^*}{2} + \frac{Y}{2}[\Phi\{\lambda(c+\frac{Y}{\sigma})\} - \Phi\{\lambda(c-\frac{Y}{\sigma})\}].$$

For $\lambda = 1$, U is the maximum likelihood estimator (MLE) of E(M).

$$E(U) = \frac{\theta^*}{2} - \frac{\theta}{2} + \frac{1}{2}E[Y\Phi\{\lambda(c+\frac{Y}{\sigma})\}] + \frac{1}{2}E[Y\Phi\{\lambda(\frac{Y}{\sigma} - c)\}]$$

and using (A2.8) this is simplified to

$$(2.7)\qquad E(U) = \frac{\theta^*}{2} + \frac{\theta}{2}\left[\Phi\left\{\frac{\lambda(\delta+c)}{\sqrt{1+\lambda^2}}\right\} - \Phi\left\{\frac{\lambda(c-\delta)}{\sqrt{1+\lambda^2}}\right\}\right]$$

$$+ \frac{\sigma\lambda}{2\sqrt{1+\lambda^2}}\left[\varphi\left\{\frac{\lambda(\delta+c)}{\sqrt{1+\lambda^2}}\right\} + \varphi\left\{\frac{\lambda(\delta-c)}{\sqrt{1+\lambda^2}}\right\}\right].$$

We note that the sum of the first two terms in (2.7) tends to E(M) as $\lambda \to \infty$. Also, it is seen clearly from (A2.7) that the last term in (2.7) is unbiasedly estimated by $\frac{\sigma\lambda}{2}[\varphi\{\lambda(\frac{Y}{\sigma}+c)\}+\varphi\{\lambda(\frac{Y}{\sigma}-c)\}]$. By subtracting this unbiased estimator from U, we define

$$(2.8) \qquad T_{2\lambda} = \frac{Y^*}{2} + \frac{Y}{2}[\Phi\{\lambda(c+\frac{Y}{\sigma})\} - \Phi\{\lambda(c-\frac{Y}{\sigma})\}] - \frac{\sigma\lambda}{2}[\varphi\{\lambda(c+\frac{Y}{\sigma})\} + \varphi\{\lambda(c-\frac{Y}{\sigma})\}].$$

The bias of $T_{2\lambda}$ is

$$(2.9) \qquad B(T_{2\lambda}) = \frac{\theta}{2}\left[\Phi\left\{\frac{\lambda(\delta+c)}{\sqrt{1+\lambda^2}}\right\} + \Phi\left\{\frac{\lambda(\delta-c)}{\sqrt{1+\lambda^2}}\right\} - \Phi(\delta+c) - \Phi(\delta-c)\right],$$

which tends to zero as λ tends to infinity. Finally, we note that $T_{2\lambda}$ corresponds to t_c of Dahiya [3] when c = 0; Dahiya's c corresponds to our λ.

2.3 Estimators $T_{3\lambda}$

Let us first consider T_3, the MLE of E(M); this is same as U in (2.6) with $\lambda = 1$. Thus, from (2.7) with $\lambda = 1$,

$$(2.10) \qquad E(T_3) = \frac{\theta^*}{2} + \frac{\theta}{2}\left[\Phi(\frac{\delta+c}{\sqrt{2}}) - \Phi(\frac{c-\delta}{\sqrt{2}})\right] + \frac{\sigma}{2\sqrt{2}}\left[\varphi(\frac{\delta+c}{\sqrt{2}}) + \varphi(\frac{\delta-c}{\sqrt{2}})\right].$$

Hence the bias of T_3 is

$$(2.11)\qquad B(T_3) = \frac{\theta}{2}\left[\Phi\left(\frac{\delta+c}{\sqrt{2}}\right) + \Phi\left(\frac{\delta-c}{\sqrt{2}}\right) - \Phi(\delta+c) - \Phi(\delta-c)\right]$$

$$+ \frac{\sigma}{2\sqrt{2}}\left[\varphi\left(\frac{\delta+c}{\sqrt{2}}\right) + \varphi\left(\frac{\delta-c}{\sqrt{2}}\right)\right].$$

Noting that the last term in (2.11) is unbiasedly estimated by

$\frac{\sigma}{2}\left[\varphi\left(\frac{Y}{\sigma}+c\right) + \varphi\left(\frac{Y}{\sigma} - c\right)\right]$, we define

$$(2.12)\qquad T_{3\lambda} = T_3 - \lambda\left[\frac{Y}{2}\left\{\Phi\left(\frac{c}{\sqrt{2}} + \frac{Y}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{Y}{\sqrt{2}\sigma} - \frac{c}{\sqrt{2}}\right) - \Phi\left(\frac{Y}{\sigma} + c\right) - \Phi\left(\frac{Y}{\sigma} - c\right)\right\}\right.$$

$$\left. + \frac{\sigma}{2}\left\{\varphi\left(\frac{Y}{\sigma} + c\right) + \varphi\left(\frac{Y}{\sigma} - c\right)\right\}\right].$$

The bias of $T_{3\lambda}$ is

$$(2.13)\qquad B(T_{3\lambda}) = B(T_3) - \frac{\lambda\sigma}{2\sqrt{2}}\left[\varphi\left(\frac{\delta+c}{\sqrt{2}}\right) + \varphi\left(\frac{\delta-c}{\sqrt{2}}\right)\right]$$

$$- \frac{\lambda}{2} E\left[Y\left\{\Phi\left(\frac{c}{\sqrt{2}} + \frac{Y}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{Y}{\sqrt{2}\sigma} - \frac{c}{\sqrt{2}}\right) - \Phi\left(\frac{Y}{\sigma} + c\right) - \Phi\left(\frac{Y}{\sigma} - c\right)\right\}\right].$$

Now, using (A2.8) to evaluate the expectation in (2.13) and carrying out routine manipulations, we obtain

$$(2.14)\qquad B(T_{3\lambda}) = \frac{\theta}{2}\Big[(1+\lambda)\Big\{\Phi(\frac{\delta+c}{\sqrt{2}}) + \Phi(\frac{\delta-c}{\sqrt{2}})\Big\} - \lambda\Big\{\Phi(\frac{\delta+c}{\sqrt{3}}) + \Phi(\frac{\delta-c}{\sqrt{3}})\Big\}$$

$$- \{\Phi(\delta+c) + \Phi(\delta-c)\}\Big]$$

$$+ \frac{\sigma}{2\sqrt{2}}\Big[\Big\{\varphi(\frac{\delta+c}{\sqrt{2}}) + \varphi(\frac{\delta-c}{\sqrt{2}})\Big\} - \lambda\sqrt{\frac{2}{3}}\Big\{\varphi(\frac{\delta+c}{\sqrt{3}}) + \varphi(\frac{\delta-c}{\sqrt{3}})\Big\}\Big].$$

It should be noted that $T_{3\lambda}$ is slightly different from the estimator T_λ of Dahiya [3] when $c = 0$ because he does not estimate part of $B(T_3)$ unbiasedly as we do in (2.12).

III. MEAN SQUARED ERRORS

We give here the expressions for the MSEs of T and $T_{1\lambda}$. For $T_{2\lambda}$ and $T_{3\lambda}$ the derivations become more tedious. For the numerical comparisons of the MSEs we use numerical integration as will be explained later in this section.

3.1 Mean Squared Error of T

We note that M and T can be written as

$$(3.1)\qquad \begin{cases} M = \frac{1}{2}\{\theta^* + \theta(I_1-I_2)\}, \\ T = \frac{1}{2}\{Y^* + Y(I_1-I_2)\}. \end{cases}$$

Now, $M - T = \frac{1}{2}\{(Y^*-\theta^*) + (Y-\theta)(I_1-I_2)\}$. Since Y and Y* are independent and $E(Y^*-\theta^*) = 0$, it is easy to see by direct evaluation that

(3.2) $$\mathrm{MSE}(T) = \frac{\sigma^2}{4} + \frac{\sigma^2}{4}\int_{c-\delta}^{\infty} t^2\varphi(t)\,dt + \frac{\sigma^2}{4}\int_{-\infty}^{-c-\delta} t^2\varphi(t)\,dt$$

$$= \frac{\sigma^2}{4}\{2 + (c-\delta)\varphi(c-\delta) + (c+\delta)\varphi(c+\delta) + \Phi(\delta-c) - \Phi(\delta+c)\},$$

using (A2.3).

3.2 Mean Squared Error of $T_{1\lambda}$

Letting $V = \varphi(c + \frac{Y}{\sigma}) + \varphi(c - \frac{Y}{\sigma})$, we have

(3.3) $$\mathrm{MSE}(T_{1\lambda}) = E[(T-M) - \frac{\lambda\sigma}{2} V]^2$$

$$= \mathrm{MSE}(T) + (\lambda\sigma^2/4)\, E(V^2) - \lambda\sigma E[(T-M)V].$$

By repeated applications of (A2.6), $E(V^2)$ can be evaluated in a straightforward manner to yield

(3.4) $$E(V^2) = \frac{1}{\sqrt{6\pi}}\left[\varphi\left\{\sqrt{\tfrac{2}{3}}\,(\delta+c)\right\} + \varphi\left\{\sqrt{\tfrac{2}{3}}\,(\delta-c)\right\} + 2\varphi\left\{\sqrt{\frac{2(\delta^2+3c^2)}{3}}\right\}\right].$$

Again, noting that V (which is a function of Y) and Y* are independent, it is easy to see that

(3.5) $$E[(T-M)V] = \frac{1}{2}E[(Y-\theta)I_1V] - \frac{1}{2}E[(Y-\theta)I_2V].$$

The right-hand side of (3.5) can be written as a sum of four integrals each of which is either of the form in (A2.4) or in (A2.5). Thus we get $4\,E[(T-M)V] = B_1 + B_2 + B_3 + B_4$, where

$$
(3.6)\quad \begin{cases}
B_1 = \sigma\,\varphi\left(\frac{c+\delta}{\sqrt{2}}\right)\left[\varphi\left(\frac{3c-\delta}{\sqrt{2}}\right) + \frac{c+\delta}{\sqrt{2}}\,\Phi\left(\frac{3c-\delta}{\sqrt{2}}\right) - \frac{c+\delta}{\sqrt{2}}\right], \\
B_2 = \sigma\,\varphi\left(\frac{\delta-c}{\sqrt{2}}\right)\left[\varphi\left(\frac{\delta-c}{\sqrt{2}}\right) - \frac{\delta-c}{\sqrt{2}}\,\Phi\left(\frac{\delta-c}{\sqrt{2}}\right)\right], \\
B_3 = \sigma\,\varphi\left(\frac{c+\delta}{\sqrt{2}}\right)\left[\varphi\left(\frac{c+\delta}{\sqrt{2}}\right) + \frac{c+\delta}{\sqrt{2}}\,\Phi\left(\frac{-c-\delta}{\sqrt{2}}\right)\right], \\
B_4 = \sigma\,\varphi\left(\frac{\delta-c}{\sqrt{2}}\right)\left[\varphi\left(\frac{3c+\delta}{\sqrt{2}}\right) + \frac{\delta-c}{\sqrt{2}}\,\Phi\left(\frac{-3c-\delta}{\sqrt{2}}\right)\right].
\end{cases}
$$

Now we can obtain MSE $(T_{1\lambda})$ by using (3.2), (3.4), and (3.6) in (3.3).

Remark 3.1. For $i = 1, 2, 3$, $T_{i\lambda} - M$ is of the form $\frac{1}{2}(Y^*-\theta^*) + g(Y)$, where g is some known function. Hence, $MSE(T_{i\lambda}) = \sigma^2/4 + E[g^2(Y)]$; the expectation of the product term is zero. After suitable change of variable, $E[g^2(Y)]$ can be numerically evaluated using the Gauss-Hermite quadrature formula. Our results in the next section were obtained by using the 20-point formula.

IV. COMPARISON OF THE ESTIMATORS

In this section we make some comparisons of the performances of $T_{1\lambda}$, $T_{2\lambda}$ and $T_{3\lambda}$ based on the values of their biases and the MSEs computed in units of σ and σ^2, respectively, for $\delta = 0\ (0.5)\ 4\ (1)\ 7$; $c = 1.645, 1.9600, 2.576$; $\lambda = 0, \sqrt{2}\exp\{-c^2/4\}, 1, \sqrt{2}$. These biases and MSEs are given in Tables 1A through 1C and Tables 2A through 2C respectively. For convenience, $\sqrt{2}\exp\{-c^2/4\}$ is denoted by λ_c in the tables.

Remark 4.1. The choices for values of λ are based on the following considerations. For $\lambda = 0$, $T_{1\lambda}$ becomes T. The value $\lambda = 1$ corresponds to using MLEs in constructing the estimators. Further, $B(T_{1\lambda})$ decreases in λ and if we let λ_δ to be the value for which $B(T_{1\lambda}) = 0$, $\lambda_\delta \leq \sqrt{2}$. Also, λ_δ tends to 0 or $\sqrt{2}\ \exp\{-c^2/4\}$ according as δ tends to infinity or zero.

Remark 4.2. The c-values chosen here correspond to $P^* = 0.90, 0.95, 0.99$, the usual values of interest in selection problems. The value of $c = 0$ ($P^* = 0.50$) is not of interest in our selection problems. This is the case considered by Dahiya [3]. As pointed out earlier, $T_{1\lambda}$ and $T_{2\lambda}$ in this case coincide with $\hat{M}_\lambda$ and t_c of Dahiya. For $T_{1\lambda}$, our choices for values of λ are included in Dahiya's tables. For $T_{2\lambda}$, $\lambda = 1$ is the only common choice. However, we do not report our values in this case, as it is not of main interest here. Finally, Dahiya defines a hybrid estimator H_c for his problem; however, H_c is really our T. But his c-values are chosen arbitrarily and they do not correspond to P*-values of common interest.

Now, considering the biases, we see that $T_{2\lambda}$ performs better than $T_{1\lambda}$ and $T_{2\lambda}$ for small values of δ ($\delta \leq 1$ for $P^* = .90$ and $\delta \leq .5$ for $P^* = .95, .99$). As δ increases, $T_{2\lambda}$ becomes increasingly bad without adjustment for bias (i.e. $\lambda = 0$); however, with λ increasing the bias of $T_{2\lambda}$ is very much reduced. For large δ ($\delta \geq 5$ for $P^* = .90$, $\delta \geq 6$ for $P^* = .95, .99$), $T_{1\lambda}$ performs better than $T_{3\lambda}$ and is generally better than $T_{2\lambda}$ as well. For small δ ($\delta \leq 1$ for $P^* = .90, .95$) and $\delta \leq 2$ for $P^* = .99$) $T_{1\lambda}$ performs better than $T_{3\lambda}$ for $\lambda \leq \lambda_c$. For moderate δ, $T_{3\lambda}$ is better than $T_{1\lambda}$ for $\lambda \leq \lambda_c$.

From the point of view of MSE, an overall picture emerges as follows: For $\delta \leq 1$, $T_{2\lambda}$ is the best. For $1.5 \leq \delta \leq 3$, $T_{1\lambda}$ is the best. When $3.5 \leq \delta \leq 4$, $T_{1\lambda}$ with small λ or $T_{3\lambda}$ with large λ is the best. Finally, for $\delta \geq 5$, $T_{3\lambda}$ with λ away from zero is the best.

Table 1A. Biases of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in units of σ

$P^* = .90$; $c = 1.645$; $\lambda_c = .7190$

δ \ λ	0	λ_c	1	$\sqrt{2}$
0.0	.1031 0 .1434	-.0000 0 .0379	-.0403 0 -.0033	-.0997 0 -.0641
0.5	.1236 -.0275 .1549	.0184 .0091 .0495	-.0227 .0086 .0083	-.0833 .0062 -.0524
1.0	.1680 -.1277 .1707	.0590 .0183 .0683	.0164 .0190 .0284	-.0464 .0142 -.0306
1.5	.1988 -.3311 .1556	.0894 -.0063 .0648	.0466 .0034 .0293	-.0164 .0047 -.0230
2.0	.1876 -.6386 .0973	.0856 -.0732 .0292	.0458 -.0444 .0026	-.0129 -.0260 -.0367
2.5	.1384 -1.0046 .0218	.0526 -.1504 -.0161	.0190 -.0977 -.0309	-.0304 -.0583 -.0527
3.0	.0797 -1.3684 -.0329	.0151 -.1952 -.0410	-.0101 -.1227 -.0441	-.0473 -.0700 -.0488
3.5	.0357 -1.6944 -.0507	-.0073 -.1907 -.0371	-.0242 -.1105 -.0319	-.0490 -.0580 -.0241
4.0	.0125 -1.9815 -.0421	-.0129 -.1517 -.0185	-.0228 -.0774 -.0092	-.0375 -.0360 .0044
5.0	.0007 -2.4990 -.0126	-.0054 -.0619 .0062	-.0077 -.0211 .0136	-.0112 -.0067 .0245
6.0	.0000 -3.0000 -.0019	-.0009 -.0165 .0053	-.0012 -.0031 .0080	-.0017 -.0005 .0121
7.0	.0000 -3.5000 -.0002	-.0001 -.0031 .0015	-.0001 -.0003 .0021	-.0002 -.0000 .0030

Table 1B. Biases of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in terms of σ

$P^* = .95$; $c = 1.96$; $\lambda_c = .5413$

δ \ λ	0	λ_c	1	$\sqrt{2}$
	.0584	.0000	-.0495	-.0943
0.0	0	0	0	0
	.1080	.0422	.0134	-.0637
	.0784	.0168	-.0355	-.0826
0.5	-.0163	.0144	.0112	.0073
	.1250	.0570	-.0007	-.0527
	.1283	.0591	.0005	-.0524
1.0	-.0835	.0387	.0317	.0209
	.1595	.0874	.0263	-.0289
	.1799	.1037	.0391	-.0192
1.5	-.2419	.0308	.0321	.0216
	.1729	.1014	.0408	-.0140
	.1994	.1216	.0556	-.0040
2.0	-.5159	-.0380	-.0072	-.0035
	.1366	.0755	.0237	-.0231
	.1724	.1009	.0403	-.0144
2.5	-.8817	-.1511	-.0719	-.0440
	.0602	.0192	-.0155	-.0469
	.1162	.0577	.0082	-.0365
3.0	-1.2762	-.2553	-.1232	-.0731
	-.0152	-.0318	-.0459	-.0586
	.0609	.0187	-.0171	-.0494
3.5	-1.6419	-.3056	-.1336	-.0744
	-.0556	-.0510	-.0471	-.0436
	.0249	-.0021	-.0250	-.0456
4.0	-1.9586	-.2947	-.1078	-.0544
	-.0580	-.0404	-.0256	-.0122
	.0020	-.0056	-.0120	-.0178
5.0	-2.4970	-.1830	-.0365	-.0134
	-0.0225	-.0036	.0124	.0268
	.0001	-.0012	-.0023	-.0033
6.0	-2.9999	-.0818	-.0063	-.0014
	-.0040	.0044	.0116	.0180
	.0000	-.0001	-.0002	-.0003
7.0	-3.5000	-.0288	-.0006	-.0001
	-.0004	.0018	.0036	.0053

Table 1C. Biases of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in units of σ

$P^* = .99$; $c = 2.576$; $\lambda_c = .2692$

δ \ λ	0	λ_c	1	$\sqrt{2}$
0.0	.0145 0 .0537	.0000 0 .0332	-.0392 0 -.0225	-.0615 0 -.0541
0.5	.0249 -.0045 .0709	.0084 .0162 .0479	-.0364 .0096 -.0144	-.0618 .0053 -.0497
1.0	.0579 -.0287 .1163	.0360 .0537 .0874	-.0236 .0347 .0090	-.0574 .0200 -.0355
1.5	.1119 -.1057 .1682	.0828 .0782 .1333	.0040 .0603 .0386	-.0406 .0363 -.0151
2.0	.1690 -.2823 .1896	.1338 .0411 .1527	.0384 .0590 .0527	-.0157 .0367 -.0039
2.5	.1989 -.5871 .1519	.0578 .0109 .0323	.1609 -.0889 .1197	-.0006 .0069 -.0172
3.0	.1823 -.9963 .0653	.1460 -.2909 .0438	.0474 -.0696 -.0145	-.0085 -.0432 -.0476
3.5	.1302 -1.4389 -.0243	.0995 -.4978 -.0323	.0162 -.1383 -.0539	-.0310 -.0832 -.0662
4.0	.0724 -1.8456 -.0746	.0495 -.6442 -.0706	-.0126 -.1595 -.0596	-.0478 -.0905 -.0534
5.0	.0106 -2.4808 -.0565	.0018 -.7028 -.0429	-.0219 -.0890 -.0058	-.0353 -.0406 .0152
6.0	.0006 -2.9991 -.0148	-.0015 -.5980 -.0060	-.0070 -.0223 .0178	-.0101 -.0068 .0313
7.0	.0000 -3.5000 -.0020	-.0003 -.4602 .0010	-.0010 -.0031 .0091	-.0015 -.0005 .0138

Table 2A. MSEs of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in units of σ^2

$P^* = .90$; $c = 1.645$; $\lambda_c = .7190$

δ \ λ	0	λ_c	1	$\sqrt{2}$
	.3825	.3605	.3580	.3606
0.0	.2500	.2932	.2925	.2935
	.3195	.3704	.3962	.4403
	.3841	.3596	.3564	.3582
0.5	.2614	.2917	.2942	.2997
	.3196	.3726	.3999	.4466
	.3836	.3566	.3529	.3545
1.0	.3288	.2992	.3100	.3273
	.3256	.3795	.4085	.4594
	.3778	.3575	.3566	.3625
1.5	.5318	.3418	.3622	.3916
	.3525	.3952	.4214	.4699
	.3783	.3764	.3822	.3973
2.0	.9402	.4332	.4570	.4882
	.4107	.4265	.4434	.4791
	.3983	.4188	.4320	.4568
2.5	1.5628	.5508	.5652	.5818
	.4878	.4697	.4732	.4891
	.4337	.4691	.4865	.5157
3.0	2.3499	.6464	.6398	.6310
	.5516	.5061	.4974	.4941
	.4676	.5051	.5218	.5485
3.5	3.2456	.6857	.6552	.6244
	.5780	.5198	.5039	.4877
	.4885	.5184	.5311	.5509
4.0	4.2284	.6710	.6243	.5854
	.5697	.5128	.4952	.4738
	.4996	.5100	.5142	.5206
5.0	6.4994	.5842	.5412	.5187
	.5247	.4909	.4789	.4627
	.5000	.5019	.5027	.5038
6.0	9.2500	.5264	.5071	.5017
	.5042	.4916	.4868	.4800
	.5000	.5002	.5003	.5004
7.0	12.5002	.5058	.5007	.5001
	.5004	.4973	.4961	.4943

Table 2B. MSEs of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in units of σ^2

$P^* = .95$, $c = 1.96$, $\lambda_c = .5413$

δ \ λ	0	λ_c	1	$\sqrt{2}$
0.0	.3147 .2500 .2975	.3090 .2802 .3204	.3104 .2760 .3449	.3166 .2755 .3709
0.5	.3277 .2545 .3042	.3191 .2806 .3316	.3187 .2785 .3609	.3240 .2798 .3920
1.0	.3545 .2902 .3201	.3396 .2834 .3574	.3355 .2882 .3975	.3387 .2947 .4405
1.5	.3726 .4244 .3422	.3538 .2992 .3846	.3480 .3148 .4328	.3509 .3286 .4859
2.0	.3750 .7500 .3779	.3606 .3508 .4119	.3592 .3744 .4565	.3663 .3920 .5092
2.5	.3779 1.3282 .4358	.3766 .4510 .4470	.3851 .4691 .4743	.4004 .4793 .5130
3.0	.3983 2.1404 .5052	.4123 .5776 .4883	.4315 .5697 .4915	.4547 .5605 .5081
3.5	.4337 3.1082 .5585	.4572 .6822 .5202	.4820 .6337 .5026	.5081 .6021 .4983
4.0	.4676 4.1626 .5764	.4918 .7296 .5298	.5149 .6415 .5010	.5378 .5966 .4836
5.0	.4972 6.4941 .5404	.5087 .6797 .5061	.5189 .5660 .4807	.5285 .5061 .4608
6.0	.5000 9.2500 .5087	.5026 .5904 .4941	.5049 .5140 .4824	.5070 .5042 .4724
7.0	.5000 12.5002 .5010	.5003 .5359 .4969	.5006 .5017 .4935	.5009 .5002 .4905

Table 2C. MSEs of $T_{1\lambda}$(top entry), $T_{2\lambda}$(middle entry), and $T_{3\lambda}$(bottom entry) expressed in units of σ^2

$P^* = .99$; $c = 2.576$; $\lambda_c = .2692$

δ \ λ	0	λ_c	1	$\sqrt{2}$
0.0	.2729 .2500 .2687	.2717 .2637 .2721	.2723 .2587 .2830	.2750 .2574 .2902
0.5	.2852 .2514 .2753	.2829 .2658 .2802	.2816 .2612 .2960	.2839 .2600 .3065
1.0	.3167 .2667 .2922	.3114 .2707 .3011	.3045 .2685 .3305	.3055 .2688 .3505
1.5	.3518 .3399 .3153	.3423 .2821 .3287	.3281 .2853 .3751	.3275 .2904 .4079
2.0	.3722 .5600 .3482	.3602 .3210 .3626	.3426 .3272 .4189	.3424 .3405 .4618
2.5	.3750 1.0313 .4030	.3650 .4211 .4117	.3546 .4124 .4599	.3597 .4312 .5033
3.0	.3779 1.8026 .4836	.3748 .5957 .4800	.3825 .5353 .5007	.3970 .5474 .5322
3.5	.3983 2.8230 .5685	.4039 .8087 .5509	.4317 .6536 .5352	.4556 .6440 .5472
4.0	.4337 3.9832 .6227	.4452 .9916 .5951	.4848 .7167 .5492	.5128 .6801 .5421
5.0	.4885 6.4663 .5989	.4986 1.1131 .5711	.5284 .6534 .5107	.5468 .5989 .4862
6.0	.4996 9.2491 .5302	.5030 1.0119 .5151	.5128 .5456 .4783	.5186 .5197 .4602
7.0	.5000 12.5001 .5047	.5006 .8689 .4994	.5023 .5072 .4856	.5033 .5016 .4783

Taking the bias and MSE together into account, the relative performances of these estimators can be largely summarized as follows: For $\delta \leq 2$, $T_{2\lambda}$ is the best. For $2 < \delta < 5$, one of $T_{1\lambda}$ and $T_{3\lambda}$ is always better than or as good as $T_{2\lambda}$. When $\delta \geq 5$, $T_{2\lambda}$ is never the best. However, for $\delta > 2$, there is no clear choice between $T_{1\lambda}$ and $T_{3\lambda}$ because the bias and the MSE pull each in opposite directions.

APPENDIX 1

Proof of Theorem 2.1. Since Y* and Y are jointly sufficient for E(M), we restrict our attention to unbiased estimators that are functions of Y* and Y. Let $\eta(Y^*,Y)$ be an unbiased estimator of M such that $V(\eta) < \infty$. Let Z be such that it is independent of Y* and Y, and $Z \sim N(\delta/b^2, 1/b^2)$, where $b > 0$ is known. Applying the Neyman Factorization Theorem to the joint distribution of Y*, Y and Z, we see that Y* and $W = Z + Y/\sigma$ are sufficient for θ^* and δ. Thus $\eta_1 = E[\eta|(Y^*,W)]$ is a function of Y* and W only. Also, $E(\eta_1) = E(\eta)$ and $V(\eta_1) \leq V(\eta)$ by Rao-Blackwell Theorem. Now, let

$$\eta_2 = [\frac{Y^*}{2} + \frac{b^2W}{1+b^2}\{\Phi(bc^* + DW) - \Phi(bc^* - DW)\}]$$

$$-[D\varphi\{bc^* + DW\} + D\varphi\{bc^* - DW\}]$$

$$= R_1 - R_2, \text{ say,}$$

where $bc^* = c\sqrt{1+b^2}$ and $D = b^2/\sqrt{1+b^2}$. Using (A2.7) and (A2.8), it can be seen that $E(\eta_2) = E(M)$. Since $E(\eta_1 - \eta_2) = 0$, it follows that $\eta_1 = \eta_2$ with probability 1 by the completeness of (Y^*,W). Thus $V(\eta_2) = V(\eta_1) \leq V(\eta) < \infty$.

To complete the proof, we obtain a contradiction by showing that $V(\eta_2)$ can be increased indefinitely by letting $b \to \infty$. To see this, we first note that $V(R_1) < \infty$ which implies that $V(R_2) < \infty$. It is also easy to see that $0 < E(R_2) < \infty$. Hence, $E(R_2^2) < \infty$ which implies that $E[D^2\varphi^2\{bc^* + DW\}] < \infty$. On the other hand, using (A2.7) and the fact that $\varphi^2(a) = \varphi(\sqrt{2}a)/\sqrt{2\pi}$, we can see that

$$E[D^2\varphi^2\{bc^* + DW\}] = \frac{1}{\sqrt{2\pi}} \frac{b^4}{1+b^2} \frac{1}{\sqrt{1+2b^2}} \varphi\left\{\frac{\sqrt{2}b(\delta + c^*)}{\sqrt{1 + 2b^2}}\right\},$$

a quantity which tends to infinity as $b \to \infty$. □

APPENDIX 2

We state a few results relating to normal distribution. We omit the proofs, which are straightforward.

(A2.1) $$\int_a^b t\,\varphi(t)\,dt = \varphi(a) - \varphi(b).$$

(A2.2) $$\int_{-\infty}^{\infty} \varphi\{\alpha(t+\beta)\}\,\varphi(t)\,dt = \frac{1}{\sqrt{1+\alpha^2}}\,\varphi\left(\frac{\alpha\beta}{\sqrt{1+\alpha^2}}\right).$$

(A2.3) $$\int_a^b t^2\,\varphi(t)\,dt = a\,\varphi(a) - b\,\varphi(b) + \Phi(b) - \Phi(a).$$

(A2.4) $$\int_a^{\infty} t\,\varphi(t)\,\varphi(t+\alpha)\,dt = \frac{1}{2}\,\varphi\left(\frac{\alpha}{\sqrt{2}}\right)\left[\varphi\left(\sqrt{2}a + \frac{\alpha}{\sqrt{2}}\right) + \frac{\alpha}{\sqrt{2}}\,\Phi\left(\sqrt{2}a + \frac{\alpha}{\sqrt{2}}\right) - \frac{\alpha}{\sqrt{2}}\right].$$

(A2.5) $$\int_{-\infty}^{-b} t\,\varphi(t)\,\varphi(t+\alpha)\,dt = -\frac{1}{2}\,\varphi\left(\frac{\alpha}{\sqrt{2}}\right)\left[\varphi\left(\sqrt{2}\,b - \frac{\alpha}{\sqrt{2}}\right) + \frac{\alpha}{\sqrt{2}}\,\Phi\left(\frac{\alpha}{\sqrt{2}} - \sqrt{2}\,b\right)\right].$$

(A2.6) $$\int_{-\infty}^{\infty} \varphi(t+\alpha)\,\varphi(t+\beta)\,\varphi(t)\,dt = \frac{1}{\sqrt{6\pi}}\,\varphi\left\{\sqrt{\frac{2(\alpha^2+\beta^2-\alpha\beta)}{3}}\right\}.$$

Let $Y \sim N(\theta,\sigma^2)$. Then

(A2.7) $$E\left[\varphi\left\{\alpha\left(\frac{Y}{\sigma} + c\right)\right\}\right] = \frac{1}{\sqrt{1+\alpha^2}}\,\varphi\left\{\frac{\alpha(\delta+c)}{\sqrt{1+\alpha^2}}\right\},$$

(A2.8) $$E\left[Y\Phi\left\{\alpha\left(\frac{Y}{\sigma} + c\right)\right\}\right] = \theta\,\Phi\left\{\frac{\alpha(\delta+c)}{\sqrt{1+\alpha^2}}\right\} + \frac{\sigma\alpha}{\sqrt{1+\alpha^2}}\,\varphi\left\{\frac{\alpha(\delta+c)}{\sqrt{1+\alpha^2}}\right\}.$$

REFERENCES

[1] Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. Scand. J. Statist. 7, 21-26.

[2] Cohen, A. and Sackrowitz, H. B. (1982). Estimating the mean of the selected population. Statistical Decision Theory and Related Topics III, Vol. 1, S. S. Gupta and J. O. Berger (eds.). Academic Press, New York.

[3] Dahiya, R. C. (1974). Estimation of the mean of the selected population. J. Amer. Statist. Assoc. 69, 226-230.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeo. Ser. No. 150), Institute of Statistics, University of North Carolina, Chapel Hill.

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

[6] Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. John Wiley, New York.

[7] Hsieh, H.-K. (1981). On estimating the mean of the selected population with unknown variance. Comm. Statist. A - Theory Methods 10, 1869-1878.

[8] Sarkadi, K. (1967). Estimation after selection. Studia Sci. Math. Hungar. 2, 341-350.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report #83-38	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): AN ESTIMATION PROBLEM RELATING TO SUBSET SELECTION FOR NORMAL POPULATIONS	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: Technical Report #83-38
7. AUTHOR(s): S. Jeyaratnam and S. Panchapakesan	8. CONTRACT OR GRANT NUMBER(s): 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, D.C.	12. REPORT DATE: 1983
	13. NUMBER OF PAGES
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, mean of selected subset, estimation, normal, natural estimator, adjustments for bias, mean squared errors, numerical comparisons

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

Let $\pi_1, \ldots, \pi_k$ be k independent normal populations with unknown means $\theta_1, \ldots, \theta_k$, respectively, and a common known variance τ^2. Let $Y_1, \ldots, Y_k$ denote the means of independent samples of size n from $\pi_1, \ldots, \pi_k$. For selecting a nonempty subset containing the best population (the one associated

DD FORM 1 JAN 73 1473

with the largest θ_i) we consider the rule of Gupta which selects π_i if and only if $Y_i \geq \max_{1 \leq j \leq k} Y_j - d\tau/\sqrt{n}$, where $d > 0$ is determined such that the probability of including the best population in the selected subset is at least equal to a preassigned level P^* $(1/k < P^* < 1)$. We are interested in estimating M, the mean of all the θ_is corresponding to the populations that are selected. In this paper we consider the case of $k = 2$. In Section II we define the 'natural' estimator T and three classes of estimators $T_{1\lambda}$, $T_{2\lambda}$ and $T_{3\lambda}$, obtained by making adjustments for the bias of T. The biases and the MSEs of $T_{1\lambda}$, $T_{2\lambda}$ and $T_{3\lambda}$ are discussed in Sections II and III respectively. Numerical comparisons of the performances of these estimators are made in Section IV.