BAYES-P* SUBSET SELECTION PROCEDURES FOR THE BEST POPULATION*

by

Shanti S. Gupta
Purdue University

Hwa-Ming Yang
University of Toledo

Technical Report #84-2
(Revised April 1985)

Department of Statistics
Purdue University

February 1984

*This research was supported by the Office of Naval Research Contracts N00014-75-C-0455 and N00014-84-C-0167 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

BAYES-P* SUBSET SELECTION PROCEDURES FOR THE BEST POPULATION*

by

Shanti S. Gupta
Purdue University

Hwa-Ming Yang
University of Toledo

ABSTRACT

Two new selection procedures, called nonrandomized and randomized Bayes-P* procedures are defined for selecting a small nonempty subset of k populations which contains the best population. It is shown that these procedures have some optimal properties. If we restrict attention to the class D(D*) of all nonrandomized (randomized) selection procedures, which satisfy the PP*-condition, that is the posterior probability of a correct selection, for any given observation $\underline{X} = \underline{x}$, is not less than P*, a predetermined number between 1/k and 1, then these two new selection procedures are shown also to be Bayes decision procedures in the class D and D* respectively, provided that some regularity conditions are satisfied. Robustness of these procedures and comparisons with some other selection procedures are studied by using Monte Carlo simulations.

AMS Subject Classification: Primary 62F07; Secondary 62F15.

Key words: Bayes selection procedures; Selection and ranking; Normal populations.

Running Title: Bayes-P* Selection for the Best Population

*This research was supported by the Office of Naval Research Contracts N00014-75-C-0455 and N00014-84-C-0167 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1. INTRODUCTION

Suppose we have k populations $\pi_1,\ldots,\pi_k$ whose distributions are determined by unknown real parameters $\theta_1,\ldots,\theta_k$, respectively. In a subset selection problem, the goal is to select a subset of the populations which includes the population associated with the largest parameter with high probability and, possibly, includes the others with low probabilities. A population π_i will be called the best population if $\theta_i \geq \theta_j$ for all $j \neq i$. The remaining k-1 populations will be called non-best. If there are more than one population satisfying this condition we arbitrarily tag one of them and call it the best one.

A large body of literature exists in the area of subset selection procedures (see Gupta and Panchapakesan (1979)). As pointed out by many authors (see, for example, Bahadur (1950), Bechhofer (1954)) the testing of homogeneity of population means or variances is not a satisfactory solution to a comparison among several populations. Gupta (1956, 1965) gave a maximum-type subset selection procedure. The maximum-type procedures present a direct and relatively efficient way to meet our goal. Gupta and Hsu (1978) studied the performance of maximum-type procedure, average-type procedure (Seal (1955, 1957)) and Bayes procedures. Berger (1979) and Berger and Gupta (1980) proved that the maximum-type procedure is minimax under certain loss functions. The LFC (least favorable configuration) of the maximum-type procedure usually occurs when the distributions are identical, i.e., under the hypothesis of homogeneity. As usual, in many cases, the hypothesis of homogeneity is rejected at some small significance level. It seems then that the maximum-type procedure is still conservative. Therefore we may wish to relax (modify) the so-called P*-condition. On the other hand, in the decision-theoretic approach, Bayes procedure always provides a decision with minimum risk under a certain loss; however, this

does not mean that its quality is good enough to pass a certain level. Suppose the loss function is a linear combination of $L_i(\underline{\theta})$, $i = 1,\ldots,k$, where $L_i(\underline{\theta})$ is the loss if the ith population is selected in the subset, as was assumed by Bahadur and Goodman (1952), Dunnett (1960), Lehmann (1966), Eaton (1967) and Alam (1973). As pointed out by Goel and Rubin (1977), the decision-theoretic procedures mentioned above do not seem to be appropriate, mainly because they ignore a reasonable component of loss which depends on whether or not the selected subset contains the best population and secondly because they specify the subset size in advance, whereas it should depend on the information available from the sample. We may use some loss functions that involve an additional component, such as the loss function given by Gupta and Hsu (1978), which is associated with the probability of incorrect selection, or the one given by Goel and Rubin (1977), which is associated with the distance between the selected subset and the best population, or some others which are proposed by Chernoff and Yahav (1977), Bickel and Yahav (1977) and Kim (1979), to improve the quality of decision. However, the results are quite sensitive to the weights of each components, or equivalently, the ratio of the coefficient of the two components. In practice we always have some difficulties in figuring out the ratio whenever the two components of loss are not comparable, or they are comparable but the ratio is not a constant function or it is not completely known. In these situations we may wish to try some other methods of attack.

For guaranteeing the quality of selection procedures, we would like to have a 'quality control' about the class of all possible selection procedures, that is, any selection procedure with lower

quality will be removed, even though it might be the cheapest one under some loss. By using the PP*-condition (defined in Section 2) as a control condition we get a class of selection procedures D and its randomized version D*. The PP*-condition represents the minimum quality (accuracy) level of a selection procedure under certain prior information. We try to derive one of these procedures, which gives the minimum risk under a large family of loss functions and has properties which we think an optimal selection procedure should have.

In Section 2 we define some notations, the PP*-condition (posterior-P*-condition), class D and D* and propose two selection procedures ψ^B and ψ^{B*}. Some selection procedures close to ψ^B but restricted to normal populations were studied by Roth (1978) and Naik (1978). The optimal properties of ψ^B and ψ^{B*} in D and D*, respectively, such as ordered, justness, most efficient and Bayes with respect to a large family of loss functions are shown in Section 3. Their application to normal distributions is discussed in Section 4. In Section 5 procedure ψ^B is compared with the maximum-type procedure. In Section 6 we discuss their applications to the selection problem for Poisson distributions and Poisson processes and their relation to the selection of gamma distributions. Section 7 deals with comparisons of the performance of selection procedures ψ^{B*}, ψ^B, ψ^M and ψ^{MED}. Here ψ^M and ψ^{MED} are the maximum-type selection procedures based on sample means and sample medians, respectively (see Gupta (1956, 1965) and Gupta and Singh (1980)). The comparisons are based on Monte Carlo studies. Robustness of these four procedures is studied in terms of the expected size of the selected subset and the efficiency (defined in Section 3) where the robustness is in the sense of the effect on the performance of the procedure when the k true distributions are not

normal but, say, logistic, double exponential distribution or the contaminated distribution (gross error model) (Tukey (1960)). Further discussion about the Bayes-P* selection procedures is given in Section 8.

2. NOTATION AND FORMULATION

Assume that we have n_i independent observations X_{ij}, $j = 1,\ldots,n_i$ for population π_i, $i = 1,\ldots,k$. Let $X_i = T_i(X_{i1},\ldots,X_{in_i})$ be a suitable estimator of θ_i, $i = 1,\ldots,k$; assume that X_i's are independently distributed. Usually X_i is a sufficient statistic for θ_i. Let $\underline{\theta} = (\theta_1,\ldots,\theta_k) \in \Theta \subset R^k$ and let $\underline{X} = (X_1,\ldots,X_k)$ with cumulative distribution function (cdf) $F(\underline{x}|\underline{\theta})$ and density (frequency) $f(\underline{x}|\underline{\theta})$. A selection procedure will be denoted by $\psi(\underline{x}) = (\psi_1(\underline{x}),\ldots,\psi_k(\underline{x}))$ where $\psi_i(\underline{x})$: $R^k \to [0,1]$ is the probability that π_i is included in the selected subset when $\underline{X} = \underline{x}$ is observed. A selection procedure ψ is called nonrandomized if all ψ_i's are 0 or 1, otherwise, it is a randomized procedure. A correct selection (CS) is defined to be the selection of any subset that includes the best population. Suppose we have a prior distribution τ for $\underline{\theta} = (\theta_1,\ldots,\theta_k)$ and our control condition is that for any given observation the posterior probability of CS must be greater than or equal to P*, a preassigned value between 1/k and 1. That is

$$(2.1) \qquad P(CS|\psi,\underline{X} = \underline{x}) = \sum_{i=1}^{k} \psi_i(\underline{x})p_i(\underline{x}) \geq P^*, \text{ for all } \underline{x},$$

where

$$(2.2) \qquad p_i(\underline{x}) = P(\pi_i \text{ is the best}|\underline{X} = \underline{x}) = P(\theta_i \text{ is the largest}|\underline{X} = \underline{x}).$$

For convenience, we now assume that posterior cdf of $\underline{\theta}$ is absolutely continuous. Then it is clear that

$$\sum_{i=1}^{k} p_i(\underline{x}) = 1,$$

hence this kind of selection procedures always exist. Let $p_{[1]}(\underline{x}) \leq \ldots \leq p_{[k]}(\underline{x})$ be the ordered $p_i(\underline{x})$'s and let $\pi_{(i)}$ be the population associated with $p_{[i]}(\underline{x})$, $i = 1,\ldots,k$, then a subset selection procedure is completely specified by $\{\psi_{(1)},\ldots,\psi_{(k)}\}$ where $\psi_{(i)}$ is defined by

(2.3) $\quad \psi_{(i)}(\underline{x}) = P(\pi_{(i)}$ is selected$|\psi, \underline{X} = \underline{x})$, $i = 1,\ldots,k$.

Definition 2.1. Given a number P*, $1/k < P^* < 1$, and a prior τ, we say a selection procedure ψ satisfies the PP*-condition (posterior-P*-condition) if

(2.4) $\quad \psi_{(k)}(\underline{x}) = 1$ and $P(CS|\psi, \underline{X} = \underline{x}) \geq P^*$ for all $\underline{x}$.

Remark 1. The PP*-condition implies the expected probability of CS with respect to a given prior is not less than P*. Since the prior information is used in PP*-condition, it is different from the usual so-called P*-condition.

Given a prior τ, let $D = D(\tau,P^*)$ ($D^* = D^*(\tau,P^*)$) be the class of all nonrandomized (randomized) selection procedures in which all procedures satisfy the PP*-condition for any given observation $\underline{X} = \underline{x}$.

Now, we propose two selection procedures ψ^B and ψ^{B*} as follows:

Definition 2.2. Given a number $P^*(1/k < P^* < 1)$, an observation $\underline{X} = \underline{x}$ and a prior τ, the selection procedure ψ^B is defined by $\{\psi^B_{(1)},\ldots,\psi^B_{(k)}\}$ where

$$\psi^B_{(i)}(\underline{x}) = \begin{cases} 1, & \text{if } i \geq j \\ 0, & \text{otherwise} \end{cases}$$

and j is the largest integer between 1 and k, such that

$$\sum_{i=j}^{k} p_{[i]}(\underline{x}) \geq P^*.$$

Definition 2.3. Given a number $P^*(1/k < P^* < 1)$, an observation $\underline{X} = \underline{x}$ and a prior τ, the randomized selection procedure ψ^{B^*} is defined by $\{\psi^{B^*}_{(1)},\ldots,\psi^{B^*}_{(k)}\}$

where

$$\psi^{B^*}_{(k)}(\underline{x}) = 1,$$

and

$$\psi^{B^*}_{(j)}(\underline{x}) = \begin{cases} 1, \text{ if } \sum_{i=j}^{k} p_{[i]}(\underline{x}) \leq P^*, \ j \neq k, \\ v, \text{ if } \begin{cases} \sum_{i=j+1}^{k} p_{[i]}(\underline{x}) < P^* \text{ and} \\ \sum_{i=j}^{k} p_{[i]}(\underline{x}) > P^*, \end{cases} \\ 0, \text{ otherwise,} \end{cases}$$

where the constant v is determined so that

$$v p_{[j]}(\underline{x}) + \sum_{i=j+1}^{k} p_{[i]}(\underline{x}) = P^*, \ 0 < v < 1.$$

Example. If $k = 3$, $P^* = 0.90$ and the posterior probabilities are: $p_1(\underline{x}) = .05$, $p_2(\underline{x}) = .80$, $p_3(\underline{x}) = .15$, then selection procedure ψ^{B^*} will select the population π_2 (corresponding to $p_{[3]}(\underline{x})$) with probability 1, and select π_3 with probability v where v is given by

$$.15v + .80 = .90$$

$$v = .10/.15 = 2/3.$$

By definition ψ^{B} (ψ^{B^*}) is an element of D (D*). In the next section, we will show some properties for both selection procedures ψ^{B} and ψ^{B^*}. They are properties of ordering, justness, efficiency

and Bayes in its class, respectively.

3. OPTIMAL PROPERTIES

In this section some properties of selection procedures ψ^B and ψ^{B*} are studied.

Definition 3.1. A selection procedure ψ is called ordered if for every $\underline{x} \in R^k$, $x_i \leq x_j$ implies $\psi_i(\underline{x}) \leq \psi_j(\underline{x})$. It is called just if for every $i = 1,\ldots,k$, and $\underline{x}, \underline{x}' \in R^k$, $\psi_i(\underline{x}) \leq \psi_i(\underline{x}')$ whenever $x_i \leq x_i'$, $x_j \geq x_j'$ for any $j \neq i$.

Just procedures were defined and investigated in more generality by Nagel (1970) and Gupta and Nagel (1971).

Definition 3.2. A selection procedure is translation invariant if for every $\underline{x} \in R^k$, for every $c \in R$, $\psi_i(\underline{x}+c\underline{1}) = \psi_i(\underline{x})$ for every $i = 1,\ldots,k$, where $\underline{1} = (1,\ldots,1)$.

Lemma 3.1. (Berger and Gupta (1980)) A selection procedure $\psi(\underline{x}) = (\psi_1(\underline{x}),\ldots,\psi_k(\underline{x}))$ is just and translation invariant if and only if the following two conditions hold:

(1) for every $i = 1,\ldots,k$, ψ_i is a function only of the set of differences $\{x_j-x_i \mid j = 1,\ldots,k, j \neq i\}$, and

(2) if $\underline{x}$ and $\underline{y}$ satisfy $x_j-x_i \leq y_j-y_i$ for every $j \neq i$, then $\psi_i(\underline{x}) \geq \psi_i(\underline{y})$.

Let $\wp = (p_1,\ldots,p_k)$, where p_i's are defined by (2.2). If we treat $\wp$ as a selection procedure, then $\psi^B(\psi^{B*})$ is ordered, just and translation invariant if and only if $\wp$ is ordered, just and translation invariant, respectively. Therefore, we have Theorem 3.1.

Theorem 3.1. Selection procedures ψ^B and ψ^{B*} are just and translation invariant if

(1) for every $i = 1,\ldots,k$, $p_i = p_i(\underline{x})$ is a function only of the set of differences $\{x_j - x_i \mid j = 1,\ldots,k,\ j \neq i\}$, and

(2) if $\underline{x}$ and $\underline{y}$ satisfy $x_j - x_i \leq y_j - y_i$ for every $j \neq i$, then $p_i(\underline{x}) \geq p_i(\underline{y})$.

Some sufficient conditions for $\wp$ to be ordered, just and translation invariant are given below:

Theorem 3.2. Let $H(\underline{\theta}|\underline{x})$ be the posterior cdf of $\underline{\theta}$, given $\underline{X} = \underline{x}$. If $H(\underline{\theta}|\underline{x})$ is absolutely continuous and has the generalized stochastic increasing property (GSIP), that is:

(1) $H(\underline{\theta}|\underline{x}) = \prod_{i=1}^{k} H_i(\theta_i|\underline{x})$, $H_i(\cdot|\underline{x})$ = posterior cdf of θ_i.

(2) $H_i(t|\underline{x}) \geq H_j(t|\underline{x})$ for any t, whenever $x_i \leq x_j$.

Then both ψ^B and ψ^{B*} are ordered and just. If, in addition, H_i has location parameter x_i, that is, $H_i(\theta_i|\underline{x}) = H_i(\theta_i - x_i)$ for every $i = 1,\ldots,k$, then both ψ^B and ψ^{B*} are also translation-invariant.

Proof: The first part can be proved by using integration by parts. Since $x_i \leq x_j$, implies $H_i(t|\underline{x}) \geq H_j(t|\underline{x})$ for all t, hence

$$
\begin{aligned}
p_i(\underline{x}) &= P(\theta_i = \theta_{[k]}|\underline{x}) \\
&= \int \prod_{m\neq i} H_m(t|\underline{x})\, dH_i(t|\underline{x}) \\
&\leq \int \prod_{m\neq j} H_m(t|\underline{x})\, dH_i(t|\underline{x}) \\
&= 1 - \int H_i(t|\underline{x})\, d\Big[\prod_{m\neq j} H_m(t|\underline{x})\Big] \\
&\leq 1 - \int H_j(t|\underline{x})\, d\Big[\prod_{m\neq j} H_m(t|\underline{x})\Big]
\end{aligned}
$$

$$= \int \prod_{m \neq j} H_m(t|\underline{x}) dH_j(t|\underline{x})$$

$$= p_j(\underline{x}) \text{ if } x_i \leq x_j.$$

Therefore, both ψ^B and ψ^{B*} are ordered. The proof of justness is similar to the above and hence omitted. The proof of the second part is obvious.

Let G denote the group of all permutations of the components of a k-component vector. A set $S \subset R^k$ is called symmetric if $gS = S$ for all $g \in G$. A distribution H is called symmetric if $H(S) = H(gS)$ for all measurable set S and $g \in G$. A family of distributions $P_{\underline{\theta}}$ is called invariant with respect to G if $P_{\underline{\theta}}(S) = P_{g\underline{\theta}}(gS)$ for all measurable set S and $g \in G$.

Given an observation $\underline{X} = \underline{x}$, suppose set s is the selected subset under a selection procedure ψ in D, then the loss can be described by a non-negative real-valued function $L(\underline{\theta},s)$ which has the properties below:

Definition 3.3. For all $g \in G$, $\underline{\theta} \in \Theta$, a loss function L has property T if and only if

(1) $L(\underline{\theta},s) = L(g\underline{\theta},gs)$,

(2) $L(\underline{\theta},s)$ is non-increasing in θ_i for $i \in s$, and

(3) $L(\underline{\theta},s) \leq L(\underline{\theta},s')$, if $s \subset s'$.

Let property T' indicate that the loss function satisfies the first two conditions of property T, namely, invariance and monotonicity properties. The third condition of ordering of property T is reasonable, because the indirect loss of an incorrect selection is controlled by the PP*-condition and the direct loss of a selected subset is naturally more than its subset.

Example 3.1. The following loss functions satisfy property T:

a. $L(\underline{\theta},s) = |s|$, where $|s|$ is the size of set s.

b. $L(\underline{\theta},s) = \sum_{i\in s} L_i(\underline{\theta})$ where $L_i(\underline{\theta})$, the loss if the ith population is selected in the subset, is invariant and monotonic, i.e. $L_{gi}(g\underline{\theta}) = L_i(\underline{\theta})$ for all $g \in G$, and $L_i(\underline{\theta}) \geq L_{i+1}(\underline{\theta})$ whenever $\theta_i \leq \theta_{i+1}$, $i = 1,\ldots,k-1$. A useful form of this loss is that $L_i(\underline{\theta}) = \ell(q(\underline{\theta}),\theta_i)$ which is non-increasing in θ_i, where $q(\underline{\theta})$ is a real-valued symmetric function of $\underline{\theta}$. For example, $\ell(q(\underline{\theta}),\theta_i) = C(\theta_{[k]}-\theta_i)^a$, $a \geq 0$, $C \geq 0$.

The next theorem (Theorem 3.3.) shows that under some regularity conditions, a selection procedure $\psi^B(\psi^{B*})$ is Bayes in D(D*) if the loss function has property T.

Theorem 3.3. Suppose the prior distribution τ is symmetric on Θ. Given $\underline{\theta} \in \Theta$, $X_1,\ldots,X_k$ are independently distributed and the pdf $f(x_i|\theta_i)$ has monotone likelihood ratio (MLR) property. Then $\psi^B(\psi^{B*})$ is ordered and is a Bayes procedure in D (D*) provided that the loss function has property T.

Proof: First we need to show that the selection procedure ψ^B (ψ^{B*}) or ρ is ordered. For any $i \neq j$, if $x_i \leq x_j$, let

$$\Theta_1 = \{\underline{\theta} \in \Theta | \theta_i < \theta_j\}, \text{ then}$$

$$\begin{aligned} p_j(\underline{x}) - p_i(\underline{x}) &= b\int_{\Theta}(I_{\{\theta_j=\theta_{[k]}\}}(\underline{\theta})-I_{\{\theta_i=\theta_{[k]}\}}(\underline{\theta}))f(\underline{x}|\underline{\theta})d\tau(\underline{\theta}) \\ &= b\int_{\Theta_1}(I_{\{\theta_j=\theta_{[k]}\}}(\underline{\theta})-I_{\{\theta_i=\theta_{[k]}\}}(\underline{\theta}))(f(\underline{x}|\underline{\theta})-f(\underline{x}|\underline{\theta}'))d\tau(\underline{\theta}) \\ &\geq 0, \end{aligned}$$

where b is a normalizing factor and $\underline{\theta}'$ is obtained from $\underline{\theta}$ by interchange the components θ_i and θ_j. The third equation above is an application of

the assumption that τ is symmetric on Θ. The last inequality is based on the fact that $f(\underline{x}|\underline{\theta}) - f(\underline{x}|\underline{\theta}')$ is non-negative by the MLR property of $f(x_i|\theta_i)$ and $I_{\{\theta_j=\theta_{[k]}\}}(\underline{\theta})-I_{\{\theta_i=\theta_{[k]}\}}(\underline{\theta})$ is nonnegative. Because for any given observation $\underline{X} = \underline{x}$, ψ^B always has minimum size of the selected subset, say m, in D. Therefore, under the assumptions and the property T, the selection problem turns into "the mth decision problem" as mentioned in Lemma 1 of Goel and Rubin (1977), then by Theorem 4.1. of Eaton (1967), the result holds. The proof for procedure ψ^{B*} is analogous, and hence is omitted.

Theorem 3.4. Under the assumptions of Theorem 3.2., ψ^B (ψ^{B*}) is Bayes procedure in D(D*) provided that the loss function has property T.

Proof: By Theorem 3.2., $\psi^B(\psi^{B*})$ is ordered. By an argument similar to the argument in Theorem 3.3., the theorem is proved.

For any selection procedure $\psi \in D$, the posterior efficiency of ψ, given observation $\underline{X} = \underline{x}$, is defined by

$$\text{eff}(\psi|\underline{x}) = P(CS|\psi,\underline{x})/E(S|\psi,\underline{x})$$

where $E(S|\psi,\underline{x})$ is the posterior expected size of the selected subset. The expectation of $\text{eff}(\psi|\underline{x})$ is the efficiency of procedure and is denoted by $\text{eff}(\psi)$. A selection procedure $\psi \in D$ is called most efficient (ME) in D(D*) if $\text{eff}(\psi) \geq \text{eff}(\psi')$ for all $\psi' \in D(D^*)$.

Theorem 3.5. The selection procedures ψ^B and ψ^{B*} are ME in D and D*, respectively.

Proof: In D, given any observation $\underline{X} = \underline{x}$, since ψ^B has minimum size of selected subset, say m, any selection procedure ψ' in D should have its size of selected subset equal to m+c for some $0 \leq c \leq k-1$. Now,

$$\begin{aligned} \text{eff}(\psi'|\underline{x}) &= \{p_{[k-m-c+1]}(\underline{x}) + \ldots + p_{[k]}(\underline{x})\}/(m+c) \\ &\leq \{cp_{[k-m]}(\underline{x}) + p_{[k-m+1]}(\underline{x}) + \ldots + p_{[k]}(\underline{x})\}/(m+c) \\ &\leq \{p_{[k-m+1]}(\underline{x}) + \ldots + p_{[k]}(\underline{x})\}/m \\ &= \text{eff}(\psi|\underline{x}), \end{aligned}$$

the first part is proved. For ψ^{B*} in D*, the proof is similar hence is omitted.

4. EXTENSION AND APPLICATIONS

In this section, the formulas for the posterior probabilities which are necessary to carry out the selection procedures ψ^B and ψ^{B*} are given under various assumptions.

Suppose we have k populations $\pi_1, \ldots, \pi_k$; π_i has normal distribution $N(\mu_i, \sigma_i^2)$ where μ_i's are unknown. Assume that we have sample $X_{i1}, \ldots, X_{in_i}$ for each population π_i. Let X_i be the sample mean and let $\underline{X} = (X_1, \ldots, X_k)$. Suppose we are interested in selecting a subset containing the population having the largest population mean under the PP*-condition, with respect to some prior distribution τ of $\underline{\mu} = (\mu_1, \ldots, \mu_k)$.

A. No Prior Information

Under the situation where very little is known a priori, we may use a 'locally uniform' prior (see Box and Tiao (1973)) provided that

the unknown parameters are locally independent a priori (see Guttman and Tiao (1964)).

A.1. Common Variance σ^2 (Known) and Common Sample Size n

By using the non-informative prior which in this case is locally uniform prior, $\tau(\underline{\mu}) \propto c$, we have

$$(4.1) \qquad p_{[i]}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(t + \sqrt{n}\,\sigma^{-1}(x_{[i]}-x_{[j]}))d\Phi(t), \quad i = 1,\ldots,k.$$

A.2. Unequal Variance σ_i^2's (Known) and Unequal Sample Size n_i's

By using the same prior as above in A.1, i.e. $\tau(\underline{\mu}) \propto c$, we have

$$(4.2) \qquad p_i(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(tv_i/v_j+(x_i-x_j)/v_j)d\Phi(t)$$

where $v_i = \sigma_i/\sqrt{n_i}$, $i = 1,\ldots,k$.

A.3. Unequal and Unknown Variance σ_i^2's and Unequal Sample Size n_i's

By using the locally uniform prior $\tau(\mu_i,\sigma_i) \propto \sigma_i^{-1}$ for each population, we have

$$(4.3) \qquad p_i(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} T_{v_j}\left(t\,\frac{s_i/\sqrt{n_i}}{s_j/\sqrt{n_j}} + \frac{x_i-x_j}{s_j/\sqrt{n_j}}\right)dT_{v_j}(t)$$

where $v_i = n_i-1$,

$$(4.4) \qquad v_i s_i^2 = \sum_{j=1}^{n_i} (x_{ij}-x_i)^2, \quad i = 1,\ldots,k.$$

and T_v is the cdf of t-distribution with v degrees of freedom. For large v, it can be approximated by the normal distribution.

B. Independent Normal Prior

B.1. Identical Prior

Assume μ_i's have common distribution $N(\theta_0,\sigma_0^2)$ and given μ_i, X_i has distribution $N(\mu_i,\sigma^2/n)$, then

$$(4.5) \qquad p_{[i]}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(t+bn\sigma^{-2}(x_{[i]}-x_{[j]}))d\Phi(t),$$

where $b^2 = (\sigma_0^{-2} + n\sigma^{-2})^{-1}$, $i = 1,\ldots,k$.

B.2. Non-Identical Prior

Assume μ_i's have independent normal prior distribution $N(\theta_i,\sigma_{0i}^2)$. Since given μ_i, X_i has normal distribution $N(\mu_i,\sigma_{1i}^2/n_i)$. Let

$$(4.6) \qquad z(x_i) = b_i^2(\sigma_{0i}^{-2}\theta_i + n_i\sigma_{1i}^{-2}x_i)$$

$$(4.7) \qquad b_i^2 = (\sigma_{0i}^{-2} + \sigma_{1i}^{-2}n_i)^{-1}$$

we have, for $i = 1,\ldots,k$,

$$(4.8) \qquad p_i(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi[(tb_i + (z(x_i)-z(x_j))/b_j]d\Phi(t).$$

In Case A.1. ψ^B and ψ^{B*} are just, translation invariant and ordered, hence these will be a Bayes procedure in D and D*, respectively, provided loss function has property T.

C. General Normal Model

Suppose we have k normal populations with common known variance $\sigma^2 > 0$ and common sample size n. The observation can be reduced to $\underline{X} = (X_1,\ldots,X_k)$ where X_i is the sample mean for population π_i. Assume that, given $\underline{\mu}$, $\underline{X}$ has normal distribution $N(\underline{\mu},vI)$, where $v = \sigma^2/n$ and $\underline{\mu}$ itself has normal distribution $N(\theta_0\underline{1},rI + wU)$ with $\theta_0 \in R$, $r = \sigma_0^2 > 0$, $w > -r/k$, $\underline{1} = (1,\ldots,1)$ and $U = \underline{1}'\underline{1}$. Note that here $r > 0$ and $w > -r/k$ are sufficient and necessary for $rI + wU$ to be positive definite. This model was chosen by Chernoff and Yahav (1977) ($t > 0$), Gupta and Hsu (1978) and Miescke (1979). In this model, the $p_i(\underline{x})$'s are exactly the same as that of the independent prior case B.1.

5. RELATION BETWEEN PROCEDURES ψ^B AND ψ^M IN THE NORMAL LOCATION PARAMETER CASE

Suppose there are k independent normal populations with a common known variance σ^2 and common sample size n. For this case Gupta (1956) proposed and studied the maximum-type procedure

$$\psi^M: \text{ Select } \pi_i \text{ iff } X_i > X_{[k]} - d\sigma/\sqrt{n}, \quad i = 1,\ldots,k;$$

where $d = d(k,P^*) > 0$ is determined by the P*-condition, that is,

$$(5.1) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}(t+d)\,d\Phi(t) = P^*.$$

When the prior τ is the non-informative prior, the following theorem shows that $\psi^M \in D(\tau,P^*)$ and gives for ψ^M a lower bound for the posterior probability of CS. Let

$$(5.2) \qquad \mathcal{X} = \{\text{all possible observed values}\} = R^k,$$

$$(5.3) \qquad \mathcal{X}_1 = \{\underline{x} \in \mathcal{X} \mid x_{[k]} - d\sigma/\sqrt{n} \le x_{[1]}\},$$

$$(5.4) \qquad \mathcal{X}_i = \{\underline{x} \in \mathcal{X} \mid x_{[i-1]} < x_{[k]} - d\sigma/\sqrt{n} \le x_{[i]}\}, \quad 2 \le i \le k,$$

$$(5.5) \qquad \mathcal{X}_i^{(1)} = \{\underline{x} \in \mathcal{X} \mid x_{[1]} = x_{[i-1]} < x_{[k]} - d\sigma/\sqrt{n} \le x_{[i]}\} \subset \mathcal{X}_i$$

$$(5.6) \qquad \mathcal{X}_i^{(2)} = \{\underline{x} \in \mathcal{X} \mid x_{[1]} = x_{[i-1]} < x_{[k]} - d\sigma/\sqrt{n} = x_{[i]} = x_{[k-1]}\} \subset \mathcal{X}_i^{(1)}.$$

It is clear that $\{\mathcal{X}_1,\ldots,\mathcal{X}_k\}$ is a partition of the sample space $\mathcal{X}$.

Theorem 5.1. Given $P^*(1/k < P^* < 1)$ and a non-informative prior τ, if the observation $\underline{X} = \underline{x} \in \mathcal{X}_i$, then

$$(5.7) \qquad P(CS \mid \psi^M, \underline{X} = \underline{x}) \ge Q^*(i)$$

where

$$(5.8) \qquad Q^*(i) = P^* + (1-P^*)((k-i)/(k-1)).$$

Therefore, $\psi^M \in D(\tau,P^*)$.

Proof: Without loss of generality we can assume $\sigma/\sqrt{n} = 1$. Since $P_{[j]}(\underline{x})$ is nonincreasing for all $x_{[m]}$, $m \le j-1$, given $\underline{x} \in \Omega_i$, we have

$$
\begin{aligned}
P(CS|\psi^M, \underline{x}) &\ge \inf_{\underline{x}\in\Omega_i} P(CS|\psi^M, \underline{x}) \\
&= \inf_{\underline{x}\in\Omega_i} \sum_{j=i}^{k} P_{[j]}(\underline{x}) \\
&= \inf_{\underline{x}\in\Omega_i^{(1)}} \sum_{j=i}^{k} P_{[j]}(\underline{x}) \\
&= 1-\sup_{\underline{x}\in\Omega_i^{(1)}} \sum_{m=1}^{i-1} P_{[m]}(\underline{x}) \\
&= 1-\sup_{\underline{x}\in\Omega_i^{(1)}} \sum_{m=1}^{i-1} \int_{-\infty}^{\infty} \prod_{j\neq m} \Phi(t+x_{[m]}-x_{[j]})\, d\Phi(t) \\
&= 1-\sup_{\underline{x}\in\Omega_i^{(1)}} \sum_{m=1}^{i-1} \int_{-\infty}^{\infty} \{\prod_{j\ge i} \Phi(t+x_{[m]}-x_{[j]})\}\Phi^{i-2}(t)\, d\Phi(t) \\
&= 1-\sum_{m=1}^{i-1} \int_{-\infty}^{\infty} \Phi(t-d)\Phi^{k-2}(t)\, d\Phi(t) \\
&= 1-(i-1)\int_{-\infty}^{\infty} \Phi(t-d)\Phi^{k-2}(t)\, d\Phi(t) \\
&= Q^*(i).
\end{aligned}
$$

The supremum occurs when $\underline{x} \in \Omega_i^{(2)}$. The last equality follows from the identity

$$
\begin{aligned}
(5.9)\qquad & (k-1)\int_{-\infty}^{\infty} \Phi^{k-2}(t)\Phi(t-d)\, d\Phi(t) \\
&= 1-\int_{-\infty}^{\infty} \Phi^{k-1}(t+d)\, d\Phi(t)
\end{aligned}
$$

which can be proved by integration by parts.

Remark 5.1. If the procedure ψ^M selects one population only, i.e. $\underline{X} = \underline{x} \in \mathcal{X}_k$, then by Theorem 5.1. We have $p_{[k]} \geq P^*$, so both procedures ψ^B and ψ^{B^*} will select the same population. In other word, when the maximum-type selection procedure does an excellent job, so does the Bayes-P* selection procedures. However, the converse is not true. In general, under PP*-condtion, the subset selected by ψ^B or ψ^{B^*} is always smaller than the one selected by ψ^M, see Roth (1978) for some discussion.

Remark 5.2. For the case $k = 2$, $\psi^B = \psi^M$ a.e. for any given $\underline{X} = \underline{x}$; if $\underline{x} \in \mathcal{X}_2$, then $p_{[2]}(\underline{x}) > P^*$, hence both ψ^M and ψ^B select the population $\pi_{(2)}$ which is associated with $x_{[2]}$. If $x \in \mathcal{X}_1$ and $x_{[2]} - d\sigma/\sqrt{n} < x_{[1]}$ then ψ^M and ψ^B select both populations π_1 and π_2. Since $P(X_{[2]} - d\sigma/\sqrt{n} = x_{[1]}) = 0$, we have $\psi^B = \psi^M$ a.e..

Remark 5.3. Under the assumption of Theorem 5.1., and from the proof of it we have a lower bound on the value of $p_{[i]}(\underline{x}) + \ldots + p_{[k]}(\underline{x})$ for any observation $\underline{x} \in \mathcal{X}_i$.

6. Application to Poisson Distributions and Poisson Processes

6.1. Poisson Distributions Case

Suppose that $\pi_1, \ldots, \pi_k$ are k independent Poisson populations, where the independent observations $X_{i1}, \ldots, X_{in_i}$ from π_i have the Poisson distribution with parameter λ_i denoted by $P(\cdot|\lambda_i)$, $i = 1, \ldots, k$. Under the non-informative prior $\tau(\lambda) \propto \lambda^{-\frac{1}{2}}$ for each population, if the best population is associated with the maximum parameter, we have

$$p_i(\underline{x}) = P(\lambda_i = \lambda_{[k]}|\underline{x})$$

$$= \int_0^\infty \prod_{j \neq i} \chi^2_{m_j}(yn_i/n_j)\, d\chi^2_{m_i}(y) \tag{6.1}$$

$$(6.2) \qquad = \int_0^\infty \prod_{j\neq i} \chi^2_{m_j}(y)\, d\chi^2_{m_i}(y), \text{ if } n_1 = \ldots = n_k$$

where

$$(6.3) \qquad m_i = 2n_i x_i + 1, \; x_i = \sum_{j=1}^{n_i} x_{ij}/n_i.$$

where χ^2_m is the cdf of chi-squared distribution with m degrees of freedom.

On the other hand, if we are interested in selecting the population with the smallest parameter λ, then

$$(6.4) \qquad p_i(\underline{x}) = \int_0^\infty \prod_{j\neq i} [1-\chi^2_{m_j}(yn_j/n_i)]\, d\chi^2_{m_i}(y)$$

$$(6.5) \qquad = \int_0^\infty \prod_{j\neq i} [1-\chi^2_{m_j}(y)]\, d\chi^2_{m_i}(y) \quad \text{if} \quad n_1 = \ldots = n_k.$$

For this case, the simulation results for procedures ψ^B and ψ^{B*} are tabulated in Table 5.

6.2. Poisson Processes Case

Suppose we have k independent Poisson processes $\{X^{(1)}(t)\},\ldots,\{X^{(k)}(t)\}$ with expected arrival times equal to $1/\lambda_1,\ldots,1/\lambda_k$, respectively. Hence for the processes $\{X^{(i)}(t)\}$, the probability that there are m_i arrivals until time t_i is

$$(6.6) \qquad P(X^{(i)}(t_i) = m_i \mid \lambda_i, t_i) = (t_i\lambda_i)^{m_i} \exp(-t_i\lambda_i)/m_i!.$$

If there exists no prior information, then we use the non-informative prior $\tau(\lambda_i) \propto \lambda_i^{-\frac{1}{2}}$ for all processes. Let $\underline{m} = (m_1,\ldots,m_k)$ and $\underline{t} = (t_1,\ldots,t_k)$, it can be shown that the ith Poisson processes has the maximum parameter, i.e. the minimum expected waiting time, given $(\underline{m},\underline{t})$ is

$$(6.7)\qquad p_i(\underline{m},\underline{t}) = \int_0^\infty \prod_{j\neq i} \chi^2_{2m_j+1}(yt_j/t_i)\, d\chi^2_{2m_i+1}(y),\ i = 1,\ldots,k.$$

Here we list two special cases which are of interest.

(a) Observations $\underline{m}$ of all processes are obtained in common length time intervals $[s_i, t_0+s_i]$. Since Poisson process is stationary, we can assume that $s_i = 0$. In this case $p_i(\underline{m},\underline{t})$ is independent of $\underline{t}$.

(b) All m_i's are equal to m_0, i.e. we fix m_0 first, then get observations, the waiting time, $\underline{t}$. Let T_i be the waiting time of the m_0th arrival in the ith process, then T_i has a gamma distribution with pdf given by

$$(6.8)\qquad f(t_i) = \frac{\lambda_i}{\Gamma(m_i)} (\lambda_i t_i)^{m_i-1} e^{-\lambda_i t_i},\ t_i > 0.$$

By using the same non-informative prior for λ as before, we get the same formulas for $p_i(\underline{m},\underline{t})$. That is under case (b) the selection problem is identical with the selection problem on populations with gamma or exponential distribution.

Remark 6.1. Under non-informative prior, in comparing the subset selection problem for k Poisson distributions with the problem for k Poisson processes, it is easily seen that Poisson distribution model is a special case of Poisson processes model, with, $t_i = n_i$ and $m_i = n_i x_i$.

7. COMPARISON OF THE PERFORMANCE OF ψ^{B*}, ψ^{B}, ψ^{M} AND ψ^{MED}

Let π_i, $i = 1,\ldots,k$ be k independent populations, where π_i has the associated cdf $F_i(x,\theta_i) = F(x-\theta_i)$ with unknown location parameter θ_i. Let $f_i(x,\theta_i) = f(x-\theta_i)$ be the pdf. Suppose the goal is to find a small (nontrivial) subset which contains the best.

The following subset selection procedure ψ^{MED} based on sample medians is due to Gupta and Singh (1980).

ψ^{MED}: Select π_i if and only if $Y_i \geq Y_{[k]} - d'$

where Y_i is the median of the 2m+1 random observations from population π_i and $Y_{[k]} = \max Y_i$. The value of d' is determined by the following equation so that the P*-condition is met.

$$(7.1) \qquad \int W^{k-1}(u+d')A(u)du = P^*$$

where

$$(7.2) \qquad A(u) = ((2m+1)!/(m!)^2)[F(u)]^m[1-F(u)]^m f(u)$$

$$(7.3) \qquad W(u) = I_{F(u)}(m+1, m+1)$$

where $I_y(a,b)$ is the incomplete beta function.

In this section we use Monte Carlo simulation techniques to compare the performance of selection procedures ψ^B, ψ^{B*}, ψ^M and ψ^{MED} in the normal means problem. Because selection procedures ψ^M and ψ^{MED} are not based on any prior information about the unknown parameters, we assume that the prior distribution τ for both procedures ψ^B and ψ^{B*} is locally uniform. Since procedure ψ^M satisfies both the P*-condition and the PP*-condition with respect to the non-informative prior, it makes sense to compare the Bayes-P* procedures ψ^B and ψ^{B*} with ψ^M and compare ψ^M with ψ^{MED} in terms of efficiency. Furthermore, we assume the true distributions to be non-normal distributions, namely, the logistic, Laplace (the double exponential) and the gross error model (the contaminated) distribution, but keep the selection procedure unchanged (i.e. still based on the normal assumption) and compute the efficiency. The Monte Carlo simulation results for both equal distances of the parameters and slippage cases are tabulated. In the

simulation study all generated random variables are adjusted to have variance one. Each time five random numbers with indicated distribution were generated for each population by using IMSL random number generators. All four procedures are applied to the same data. The simulation process was repeated one thousand times. The relative frequency of selecting population π_i is used as an approximation to the probability of selecting population π_i. The relative frequency of the sum of selecting each population is treated as an approximation of the expected selected size. The efficiency (EFF) of each selection procedure is approximated by the relative frequency of the ratio of selecting the best one to the expected selection size.

The simulation results indicate that in all cases we have the performance

$$\psi^{B*} \succ \psi^{B} \succ \psi^{M}, \tag{7.4}$$

where the symbol "$\succ$" stands for better than.

For small sample size, the efficiency of rule ψ^{M} tends to be larger than ψ^{MED} under the P*-condition.

Remark 7.1. The gross error model we used has the density function

$$f(x-\theta) = (1-c)\varphi(x-\theta) + \frac{c}{4}\varphi\left(\frac{x-\theta}{4}\right), \quad c = .15 \tag{7.5}$$

for which φ is the pdf of $N(0,1)$ and the variance for this distribution is $(1-c) + 16c = 3.25$.

In the tables, the efficiency (EFF) of a procedure ψ, given parameter $\underline{\theta}$, is defined by

$$EFF_{\underline{\theta}}(\psi) = P_{\underline{\theta}}(CS|\psi)/E_{\underline{\theta}}(S|\psi) \tag{7.6}$$

where $E_{\underline{\theta}}(S|\psi)$ is the expected selected size.

Discussion of the Tables:

For Table 1 and Table 2 (equal distances case) the value of P* is .99 and .90 respectively, the common sample size n = 5, and k = 5. If the k populations have normal distributions with the unknown parameter configuration $(\theta, \theta+\Delta,\ldots,\theta + (k-1)\Delta)$ and common variance one, then from both tables the performance based on either the efficiency or the expected selected size is

$$\psi^{B*} \succ \psi^{B} \succ \psi^{M}, \tag{7.7}$$

if the PP*-condition is considered; and

$$\psi^{M} \succ \psi^{MED} \tag{7.8}$$

under the P*-condition.

When the true distributions are not normal, but the logistic, Laplace or the gross error model, the simulation results are very close to the normal case. This suggests the four procedures are reasonably robust. From Table 2 all efficiencies are larger than the corresponding ones in Table 1. This is to be expected because the value of P* is smaller in the second table.

For Table 3 and Table 4 (slippage case) the value of P* is .99 and .90 respectively, the common sample size n = 5, and k = 5. If the k populations have normal distributions with unknown parameter configurations $(\theta,\ldots,\theta, \theta+\Delta)$ and common variance one, then from both tables the performance is the same as the equal distances case.

Note that in both equal distances and slippage cases when $\Delta\sqrt{n} > 1$, that means the largest population mean and the second largest population mean are not very close, the Bayes-P* selection procedures ψ^{B} and ψ^{B*}, with respect to the locally uniform priors, always satisfy not only the PP*-condition but also the estimated $P(CS) \geq P^*$, and the expected

selected size of the Bayes-P* procedures is much less than the selection procedures ψ^{M} and ψ^{MED}. For example, in the normal equal distance case, P* = .99, k = 5 and $\Delta\sqrt{n} = 4$, we have

$$E(S|\psi^{MED}) - E(S|\psi^{B}) = .361;$$

In the normal slippage case, P* = .99, k = 5 and $\Delta\sqrt{n} = 4$, we have

$$E(S|\psi^{MED}) - E(S|\psi^{B}) = 1.469.$$

8. DISCUSSION:

The Bayes-P* selection procedures ψ^{B} and ψ^{B*} are highly-efficient and have the following advantages.

a. These procedures can be applied to any family of distributions, even their mixtures, and do not need equal sample size for each population.

b. Good prior information will not be ignored. Even under non-informative situation, they perform well.

c. They are robust in terms of the loss function. We do not even need to specify or to know the exact form of the loss function before we make a decision. There will automatically be a Bayes decision procedure under the control condition and the assumptions given by Theorem 3.3. and Theorem 3.4.

d. The weight or contribution of each population in the selected subset is known.

e. Based on the simulation results of Section 7, Bayes-P* selection procedure is robust if the true family of distributions for each population is symmetric.

The only disadvantage in using the proposed selection procedures is that the computation of the posterior probabilities needs more work than that for the classical selection procedures for which necessary constants are available in published tables; however, this disadvantage can be offset by the use of computers. In fact, we need not evaluate all p_i's, but only a few of the large ones.

TABLE 1

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta,\theta+\Delta,\ldots,\theta+(k-1)\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .99

		NORMAL		LOGISTIC		LAPLACE		GROSS ERROR	
$\Delta\sqrt{n}$		EFF	ES	EFF	ES	EFF	ES	EFF	ES
.5	ψ^{B*}	0.2754 (.0029)	3.7545 (.0241)	0.2712 (.0026)	3.7522 (.0234)	0.2755 (.0029)	3.7220 (.0241)	0.2689 (.0032)	3.8065 (.0253)
	ψ^{B}	0.2567 (.0027)	4.0900 (.0262)	0.2532 (.0023)	4.0810 (.0251)	0.2575 (.0027)	4.0440 (.0260)	0.2498 (.0030)	4.1500 (.0273)
	ψ^{M}	0.2144 (.0016)	4.7680 (.0173)	0.2135 (.0014)	4.7710 (.0167)	0.2164 (.0018)	4.7550 (.0183)	0.2151 (.0019)	4.7680 (.0190)
	ψ^{MED}	0.2103 (.0016)	4.8480 (.0149)	0.2048 (.0005)	4.9080 (.0098)	0.2036 (.0006)	4.9340 (.0089)	0.2004 (.0002)	4.9910 (.0030)
1	ψ^{B*}	0.3798 (.0044)	2.8759 (.0247)	0.3752 (.0041)	2.8759 (.0236)	0.3724 (.0041)	2.8629 (.0238)	0.3647 (.0040)	2.9178 (.0236)
	ψ^{B}	0.3469 (.0042)	3.1710 (.0273)	0.3441 (.0040)	3.1680 (.0264)	0.3420 (.0038)	3.1470 (.0261)	0.3349 (.0038)	3.2120 (.0262)
	ψ^{M}	0.2659 (.0028)	4.0340 (.0285)	0.2722 (.0031)	3.9550 (.0289)	0.2680 (.0028)	3.9970 (.0286)	0.2648 (.0030)	4.0410 (.0286)
	ψ^{MED}	0.2407 (.0022)	4.3600 (.0251)	0.2411 (.0023)	4.3470 (.0247)	0.2277 (.0019)	4.5490 (.0223)	0.2211 (.0012)	4.6230 (.0189)
2	ψ^{B*}	0.5807 (.0061)	1.8753 (.0161)	0.5990 (.0065)	1.8340 (.0167)	0.5847 (.0061)	1.8667 (.0163)	0.5748 (.0060)	1.8913 (.0160)
	ψ^{B}	0.5318 (.0062)	2.0810 (.0189)	0.5484 (.0066)	2.0330 (.0193)	0.5313 (.0061)	2.0770 (.0185)	0.5283 (.0061)	2.0900 (.0188)
	ψ^{M}	0.4473 (.0053)	2.4680 (.0218)	0.4638 (.0059)	2.4140 (.0227)	0.4485 (.0052)	2.4560 (.0219)	0.4406 (.0053)	2.5120 (.0224)
	ψ^{MED}	0.3928 (.0046)	2.8020 (.0245)	0.3880 (.0046)	2.8310 (.0241)	0.3785 (.0036)	2.8210 (.0206)	0.3643 (.0027)	2.8750 (.0183)
4	ψ^{B*}	0.8877 (.0058)	1.1931 (.0104)	0.8773 (.0059)	1.2121 (.0108)	0.8943 (.0056)	1.1813 (.0101)	0.8829 (.0059)	1.2030 (.0107)
	ψ^{B}	0.8455 (.0073)	1.3090 (.0146)	0.8338 (.0075)	1.3330 (.0150)	0.8555 (.0072)	1.2890 (.0143)	0.8445 (.0073)	1.3110 (.0146)
	ψ^{M}	0.7635 (.0079)	1.4750 (.0160)	0.7488 (.0080)	1.5070 (.0162)	0.7697 (.0079)	1.4620 (.0159)	0.7480 (.0079)	1.5060 (.0160)
	ψ^{MED}	0.6730 (.0077)	1.6700 (.0164)	0.6528 (.0075)	1.7090 (.0158)	0.6345 (.0071)	1.7350 (.0144)	0.6112 (.0066)	1.7810 (.0135)

TABLE 2

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta,\theta+\Delta,\ldots,\theta+(k-1)\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

$k = 5, \quad P^* = .90$

		NORMAL		LOGISTIC		LAPLACE		GROSS ERROR	
$\Delta\sqrt{n}$		EFF	ES	EFF	ES	EFF	ES	EFF	ES
.5	ψ^{B*}	0.3788 (.0071)	2.3686 (.0226)	0.3917 (.0068)	2.3955 (.0226)	0.3840 (.0066)	2.3927 (.0227)	0.3857 (.0066)	2.4089 (.0233)
	ψ^{B}	0.3568 (.0064)	2.7630 (.0265)	0.3663 (.0061)	2.7810 (.0265)	0.3549 (.0058)	2.8160 (.0265)	0.3554 (.0058)	2.8110 (.0269)
	ψ^{M}	0.2729 (.0044)	3.9520 (.0328)	0.2733 (.0044)	3.9950 (.0326)	0.2705 (.0041)	4.0130 (.0326)	0.2646 (.0040)	4.0360 (.0327)
	ψ^{MED}	0.2513 (.0035)	4.1420 (.0308)	0.2513 (.0034)	4.2710 (.0292)	0.2366 (.0031)	4.4640 (.0272)	0.2224 (.0020)	4.6450 (.0217)
1	ψ^{B*}	0.5799 (.0082)	1.7930 (.0188)	0.5826 (.0081)	1.7638 (.0182)	0.5682 (.0079)	1.8109 (.0185)	0.5555 (.0078)	1.7912 (.0173)
	ψ^{B}	0.5278 (.0079)	2.1140 (.0233)	0.5318 (.0080)	2.0990 (.0232)	0.5157 (.0075)	2.1610 (.0236)	0.5040 (.0073)	2.1300 (.0218)
	ψ^{M}	0.4023 (.0064)	2.8930 (.0324)	0.4069 (.0065)	2.8680 (.0325)	0.3959 (.0063)	2.9290 (.0316)	0.3907 (.0063)	2.9080 (.0308)
	ψ^{MED}	0.3544 (.0057)	3.2480 (.0332)	0.3566 (.0060)	3.2910 (.0342)	0.3274 (.0047)	3.4000 (.0308)	0.3196 (.0038)	3.4260 (.0284)
2	ψ^{B*}	0.8288 (.0070)	1.2584 (.0109)	0.8167 (.0070)	1.2736 (.0108)	0.8242 (.0069)	1.2586 (.0104)	0.8274 (.0070)	1.2614 (.0111)
	ψ^{B}	0.7728 (.0082)	1.4570 (.0169)	0.7565 (.0082)	1.4890 (.0167)	0.7577 (.0081)	1.4880 (.0167)	0.7697 (.0082)	1.4700 (.0174)
	ψ^{M}	0.6524 (.0082)	1.7770 (.0206)	0.6318 (.0080)	1.8320 (.0207)	0.6418 (.0082)	1.8070 (.0210)	0.6363 (.0081)	1.8170 (.0208)
	ψ^{MED}	0.5840 (.0081)	2.0050 (.0239)	0.5694 (.0077)	2.0400 (.0228)	0.5533 (.0070)	2.0480 (.0206)	0.5393 (.0065)	2.0630 (.0190)
4	ψ^{B*}	0.9811 (.0025)	1.0284 (.0039)	0.9821 (.0024)	1.0266 (.0037)	0.9847 (.0023)	1.0229 (.0035)	0.9796 (.0027)	1.0296 (.0040)
	ψ^{B}	0.9670 (.0039)	1.0660 (.0079)	0.9685 (.0038)	1.0630 (.0077)	0.9750 (.0034)	1.0500 (.0069)	0.9675 (.0039)	1.0650 (.0078)
	ψ^{M}	0.9160 (.0059)	1.1680 (.0118)	0.9170 (.0059)	1.1660 (.0118)	0.9295 (.0055)	1.1410 (.0110)	0.9335 (.0054)	1.1330 (.0107)
	ψ^{MED}	0.8417 (.0074)	1.3180 (.0149)	0.8625 (.0071)	1.2750 (.0141)	0.8850 (.0067)	1.2300 (.0133)	0.8910 (.0065)	1.2180 (.0131)

TABLE 3

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta,\theta,\ldots,\theta+\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .99

		NORMAL		LOGISTIC		LAPLACE		GROSS ERROR	
$\Delta\sqrt{n}$		EFF	ES	EFF	ES	EFF	ES	EFF	ES
.5	ψ^{B*}	0.2278	4.1757	0.2201	4.1722	0.2269	4.1552	0.2262	4.2047
		(.0025)	(.0202)	(.0023)	(.0193)	(.0025)	(.0208)	(.0025)	(.0222)
	ψ^{B}	0.2190	4.4940	0.2129	4.5150	0.2178	4.4880	0.2187	4.5090
		(.0021)	(.0214)	(.0019)	(.0202)	(.0021)	(.0218)	(.0022)	(.0227)
	ψ^{M}	0.2032	4.9360	0.2021	4.9360	0.2043	4.9170	0.2044	4.8840
		(.0005)	(.0089)	(.0007)	(.0096)	(.0011)	(.0109)	(.0013)	(.0136)
	ψ^{MED}	0.2031	4.9370	0.2008	4.9730	0.2009	4.9790	0.2000	4.9990
		(.0011)	(.0094)	(.0004)	(.0053)	(.0003)	(.0050)	(.0001)	(.0010)
1	ψ^{B*}	0.2448	4.0727	0.2488	4.0381	0.2432	4.0505	0.2446	4.1141
		(.0021)	(.0208)	(.0025)	(.0219)	(.0023)	(.0210)	(.0025)	(.0221)
	ψ^{B}	0.2297	4.3970	0.2344	4.3710	0.2291	4.3720	0.2302	4.4350
		(.0017)	(.0218)	(.0021)	(.0235)	(.0019)	(.0219)	(.0022)	(.0229)
	ψ^{M}	0.2045	4.9190	0.2057	4.8870	0.2044	4.9060	0.2062	4.8820
		(.0006)	(.0100)	(.0009)	(.0126)	(.0007)	(.0108)	(.0012)	(.0130)
	ψ^{MED}	0.2033	4.9190	0.2039	4.9380	0.2009	4.9670	0.2006	4.9880
		(.0007)	(.0104)	(.0010)	(.0096)	(.0005)	(.0063)	(.0002)	(.0037)
2	ψ^{B*}	0.3015	3.6352	0.3132	3.5174	0.3079	3.5926	0.3002	3.6828
		(.0040)	(.0280)	(.0044)	(.0292)	(.0044)	(.0289)	(.0044)	(.0291)
	ψ^{B}	0.2765	3.9780	0.2849	3.8680	0.2806	3.9430	0.2752	4.0390
		(.0038)	(.0302)	(.0041)	(.0312)	(.0041)	(.0310)	(.0042)	(.0311)
	ψ^{M}	0.2254	4.6700	0.2290	4.6210	0.2305	4.6280	0.2319	4.6310
		(.0026)	(.0230)	(.0028)	(.0239)	(.0031)	(.0249)	(.0034)	(.0250)
	ψ^{MED}	0.2153	4.7680	0.2144	4.8030	0.2079	4.8880	0.2024	4.9700
		(.0016)	(.0183)	(.0020)	(.0176)	(.0014)	(.0133)	(.0009)	(.0074)
4	ψ^{B*}	0.6415	1.9744	0.6361	2.0027	0.6404	1.9671	0.6519	1.9470
		(.0092)	(.0312)	(.0092)	(.0319)	(.0091)	(.0308)	(.0092)	(.0316)
	ψ^{B}	0.5908	2.2360	0.5869	2.2560	0.5874	2.2350	0.5988	2.2120
		(.0097)	(.0360)	(.0097)	(.0363)	(.0096)	(.0356)	(.0097)	(.0365)
	ψ^{M}	0.4715	2.9100	0.4696	2.9500	0.4738	2.9150	0.4701	2.9310
		(.0093)	(.0445)	(.0094)	(.0451)	(.0094)	(.0447)	(.0093)	(.0451)
	ψ^{MED}	0.3556	3.6050	0.3376	3.7610	0.3124	3.9140	0.2931	4.0280
		(.0074)	(.0428)	(.0072)	(.0420)	(.0063)	(.0398)	(.0055)	(.0371)

TABLE 4

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta,\theta,\ldots,\theta+\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .90

		NORMAL		LOGISTIC		LAPLACE		GROSS ERROR	
$\Delta\sqrt{n}$		EFF	ES	EFF	ES	EFF	ES	EFF	ES
.5	ψ^{B*}	0.2660 (.0061)	2.7314 (.0224)	0.2767 (.0062)	2.7754 (.0220)	0.2721 (.0059)	2.7695 (.0228)	0.2749 (.0059)	2.8053 (.0243)
	ψ^{B}	0.2536 (.0053)	3.1510 (.0257)	0.2650 (.0056)	3.1960 (.0253)	0.2626 (.0054)	3.1810 (.0265)	0.2641 (.0054)	3.2170 (.0281)
	ψ^{M}	0.2151 (.0026)	4.4790 (.0258)	0.2248 (.0031)	4.4960 (.0257)	0.2241 (.0031)	4.4600 (.0269)	0.2191 (.0030)	4.4670 (.0277)
	ψ^{MED}	0.2103 (.0023)	4.5330 (.0241)	0.2164 (.0024)	4.6500 (.0213)	0.2121 (.0023)	4.7440 (.0202)	0.2037 (.0011)	4.9190 (.0114)
1	ψ^{B*}	0.3532 (.0067)	2.5723 (.0236)	0.3591 (.0068)	2.5943 (.0237)	0.3558 (.0069)	2.6086 (.0240)	0.3460 (.0064)	2.6783 (.0246)
	ψ^{B}	0.3207 (.0057)	2.9760 (.0265)	0.3321 (.0061)	2.9970 (.0271)	0.3288 (.0061)	3.0260 (.0273)	0.3212 (.0059)	3.0840 (.0281)
	ψ^{M}	0.2442 (.0034)	4.2790 (.0302)	0.2490 (.0041)	4.2880 (.0307)	0.2487 (.0039)	4.3190 (.0308)	0.2440 (.0045)	4.3280 (.0321)
	ψ^{MED}	0.2378 (.0032)	4.3550 (.0277)	0.2351 (.0033)	4.4700 (.0267)	0.2177 (.0021)	4.6950 (.0207)	0.2093 (.0019)	4.8580 (.0162)
2	ψ^{B*}	0.5418 (.0082)	2.1015 (.0252)	0.5553 (.0083)	2.0727 (.0256)	0.5412 (.0082)	2.1065 (.0255)	0.5381 (.0084)	2.1113 (.0261)
	ψ^{B}	0.4829 (.0080)	2.4640 (.0299)	0.4998 (.0083)	2.4140 (.0304)	0.4874 (.0082)	2.4540 (.0302)	0.4854 (.0082)	2.4620 (.0307)
	ψ^{M}	0.3546 (.0072)	3.5630 (.0415)	0.3607 (.0075)	3.5430 (.0425)	0.3562 (.0075)	3.5710 (.0419)	0.3462 (.0070)	3.5960 (.0413)
	ψ^{MED}	0.3159 (.0062)	3.8080 (.0390)	0.3045 (.0060)	3.9580 (.0390)	0.2736 (.0052)	4.2210 (.0346)	0.2550 (.0041)	4.3540 (.0311)
4	ψ^{B*}	0.9129 (.0056)	1.1682 (.0124)	0.9037 (.0059)	1.1849 (.0129)	0.9261 (.0053)	1.1434 (.0117)	0.9191 (.0057)	1.1547 (.0126)
	ψ^{B}	0.8787 (.0072)	1.2770 (.0176)	0.8664 (.0074)	1.3040 (.0182)	0.8948 (.0068)	1.2410 (.0168)	0.8912 (.0069)	1.2510 (.0173)
	ψ^{M}	0.7689 (.0094)	1.6630 (.0317)	0.7709 (.0095)	1.6680 (.0322)	0.7931 (.0091)	1.5820 (.0298)	0.8033 (.0090)	1.5530 (.0299)
	ψ^{MED}	0.6261 (.0103)	2.2180 (.0407)	0.6523 (.0102)	2.0910 (.0388)	0.7060 (.0100)	1.8970 (.0369)	0.7192 (.0099)	1.8350 (.0350)

TABLE 5

For procedures ψ^{B*} and ψ^{B} and the parameter configurations (.5,1,...,.5k) of k Poisson populations, this table gives the values (based on simulation) of the probability or selecting the population with parameter .5i, i = 1,...,k and the expected selected size ES. The prior distribution for each population is $\tau(\lambda) \propto \lambda^{-\frac{1}{2}}$.

n = 10

k		P*							
		.99		.95		.90		.75	
		ψ^{B*}	ψ^{B}	ψ^{B*}	ψ^{B}	ψ^{B*}	ψ^{B}	ψ^{B*}	ψ^{B}
2	1	.933	1.000	.966	.990	.950	.990	.946	.990
	2	.670	.820	.450	.670	.324	.490	.110	.190
	ES	1.663	1.820	1.415	1.660	1.274	1.480	1.005	1.180
3	1	.998	1.000	.990	1.000	.967	1.000	.929	1.000
	2	.741	.850	.338	.630	.321	.510	.171	.280
	3	.205	.330	.120	.130	.065	.110	.021	.050
	ES	1.944	2.180	1.498	1.760	1.354	1.620	1.121	1.330
4	1	.988	1.000	.981	1.000	.966	.990	.911	.950
	2	.754	.820	.560	.730	.250	.400	.175	.290
	3	.369	.490	.070	.140	.067	.100	.025	.050
	4	.075	.110	.041	.060	.024	.040	0	0
	ES	2.187	2.420	1.653	1.930	1.316	1.530	1.112	1.290
5	1	.996	1.000	.981	.990	.985	1.000	.950	.980
	2	.737	.850	.503	.650	.367	.550	.128	.260
	3	.355	.470	.102	.160	.133	.210	.013	.030
	4	.067	.090	.015	.030	.013	.020	.006	.010
	5	0	0	.006	.010	0	0	0	0
	ES	2.154	2.410	1.607	1.840	1.498	1.780	1.097	1.280

REFERENCES

Alam, K. (1973). On a multiple decision rule. Ann. Statist. 1, 750-755.

Bahadur, R. R. (1950). On the problem in the theory of k populations. Ann. Math. Statist. 21, 362-375.

Bahadur, R. R. and Goodman, L. A. (1952). Impartial decision rules and sufficient statistics. Ann. Math. Statist. 23, 553-562.

Berger, R. L. (1979). Minimax subset selection for loss measured by subset size. Ann. Statist. 7, 1333-1338.

Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with application to unequal variance problems. Scand. J. Statist. 7, 21-26.

Box, G.E.P. and Tiao, G. C. (1973). Bayesian Inference in Statistical Analysis. Addison Wesley, Massachusetts.

Dunnett, C. W. (1960). On selecting the largest of k normal population means. J. Roy. Statist. Soc. Ser. B, 22, 1-40.

Eaton, M. L. (1967). Some optimal properties of ranking procedures. Ann. Math. Statist. 38, 124-137.

Goel, P. K. and Rubin, H. (1977). On selecting a subset containing the best population - A Bayes approach. Ann. Statist. 5, 969-983.

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

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

Gupta, S. S. and Hsu, J. C. (1978). On the performance of some subset selection procedures. Commun. Statist.-Simula. Computa. B7(6), 561-591.

Gupta, S. S. and Nagel, K. (1971). On some contributions to multiple decision theory. Statistical Decision Theory and Related Topics (Eds. S. S. Gupta and J. Yackel), Academic Press, New York, pp. 79-102.

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

Gupta, S. S. and Singh, A. K. (1980). On rules based on sample medians for selection of the largest location parameter. Commun. Statist.-Theor. Meth. A9(12), 1277-1298.

Guttman, I. and Tiao, G. C. (1964). A Bayes approach to some best population problems. Ann. Math. Statist. 35, 825-835.

Lehmann, E. L. (1955). Ordered families of distributions. Ann. Statist. 3, 413-428.

Lehmann, E. L. (1966). On a theorem of Bahadur and Goodman. Ann. Math. Statist. 37, 1-6.

Miescke, K. J. (1979). Bayesian subset selection for additive and linear loss functions. Commun. Statist.-Theor. Meth. A8(12), 1205-1226.

Nagel, K. (1970). On Subset Selection Rules with Certain Optimality Properties. Ph.D. Thesis (Mimeo. Ser. No. 222). Dept. of Statist., Purdue Univ., W. Lafayette, Indiana.

Naik, U. D. (1978). On selection procedures based on the exceedance probability. Calcutta Statist. Assoc. Bull. 29-40.

Roth, A. J. (1978). A new procedure for selecting a subset containing the best normal population. J. Amer. Statist. Assoc. 73, 613-617.

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

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

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. Contributions to Probability and Statistics (Eds. I. Olkin and others). Stanford Univ. Press, 39, 448-458.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report #84-2	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): BAYES-P* SUBSET SELECTION PROCEDURES FOR THE BEST POPULATION	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: Technical Report #84-2
7. AUTHOR(s): Shanti S. Gupta and Hwa-Ming Yang	8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0455 and N00014-84-C-0167
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, West 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: February 1984
	13. NUMBER OF PAGES: 32
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)

Bayes procedures, subset selection, posterior probability, correct selection, efficiency.

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

Two new selection procedures, called nonrandomized and randomized Bayes-P* procedures are defined for selecting a small nonempty subset of k populations which contains the best population. It is shown that these procedures have some optimal properties. If we restrict attention to the class D(D*) of all nonrandomized (randomized) selection procedures, which satisfy the PP*-condition, that is the posterior probability of a correct selection, for any given observation $\underline{X} = \underline{x}$, is not less than P*, a predetermined number between 1/k and 1, then these two new selection procedures are shown also to be (over)

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE 1

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta, \theta+\Delta, \ldots, \theta+(k-1)\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the parenthesis are the standard errors of the corresponding estimates.

k = 5, P* = .99

		NORAML		LOGISTIC		LAPLACE		GROSS ERROR	
$\Delta\sqrt{n}$		EFF	ES	EFF	ES	EFF	ES	EFF	ES
.5	ψ^{B*}	0.2754	3.7545	0.2712	3.7522	0.2755	3.7220	0.2689	3.8065
		(.0029)	(.0241)	(.0026)	(.0234)	(.0029)	(.0241)	(.0032)	(.0253)
	ψ^{B}	0.2567	4.0900	0.2532	4.0810	0.2575	4.0440	0.2498	4.1500
		(.0027)	(.0262)	(.0023)	(.0251)	(.0027)	(.0260)	(.0030)	(.0273)
	ψ^{M}	0.2144	4.7680	0.2135	4.7710	0.2164	4.7550	0.2151	4.7680
		(.0016)	(.0173)	(.0014)	(.0167)	(.0018)	(.0183)	(.0019)	(.0190)
	ψ^{MED}	0.2103	4.8480	0.2048	4.9080	0.2036	4.9340	0.2004	4.9910
		(.0016)	(.0149)	(.0005)	(.0098)	(.0006)	(.0089)	(.0002)	(.0030)
1	ψ^{B*}	0.3798	2.8759	0.3752	2.8759	0.3724	2.8629	0.3647	2.9178
		(.0044)	(.0247)	(.0041)	(.0236)	(.0041)	(.0238)	(.0040)	(.0236)
	ψ^{B}	0.3469	3.1710	0.3441	3.1680	0.3420	3.1470	0.3349	3.2120
		(.0042)	(.0273)	(.0040)	(.0264)	(.0038)	(.0261)	(.0038)	(.0262)
	ψ^{M}	0.2659	4.0340	0.2722	3.9550	0.2680	3.9970	0.2648	4.0410
		(.0028)	(.0285)	(.0031)	(.0289)	(.0028)	(.0286)	(.0030)	(.0286)
	ψ^{MED}	0.2407	4.3600	0.2411	4.3470	0.2277	4.5490	0.2211	4.6230
		(.0022)	(.0251)	(.0023)	(.0247)	(.0019)	(.0223)	(.0012)	(.0189)
2	ψ^{B*}	0.5807	1.8753	0.5990	1.8340	0.5847	1.8667	0.5748	1.8913
		(.0061)	(.0161)	(.0065)	(.0167)	(.0061)	(.0163)	(.0060)	(.0160)
	ψ^{B}	0.5318	2.0810	0.5484	2.0330	0.5313	2.0770	0.5283	2.0900
		(.0062)	(.0189)	(.0066)	(.0193)	(.0061)	(.0185)	(.0061)	(.0188)
	ψ^{M}	0.4473	2.4680	0.4638	2.4140	0.4485	2.4560	0.4406	2.5120
		(.0053)	(.0218)	(.0059)	(.0227)	(.0052)	(.0219)	(.0053)	(.0224)
	ψ^{MED}	0.3928	2.8020	0.3880	2.8310	0.3785	2.8210	0.3643	2.8750
		(.0046)	(.0245)	(.0046)	(.0241)	(.0036)	(.0206)	(.0027)	(.0183)
4	ψ^{B*}	0.8877	1.1931	0.8773	1.2121	0.8943	1.1813	0.8829	1.2030
		(.0058)	(.0104)	(.0059)	(.0108)	(.0056)	(.0101)	(.0059)	(.0107)
	ψ^{B}	0.8455	1.3090	0.8338	1.3330	0.8555	1.2890	0.8445	1.3110
		(.0073)	(.0146)	(.0075)	(.0150)	(.0072)	(.0143)	(.0073)	(.0146)
	ψ^{M}	0.7635	1.4750	0.7488	1.5070	0.7697	1.4620	0.7480	1.5060
		(.0079)	(.0160)	(.0080)	(.0162)	(.0079)	(.0159)	(.0079)	(.0160)
	ψ^{MED}	0.6730	1.6700	0.6528	1.7090	0.6345	1.7350	0.6112	1.7810
		(.0077)	(.0164)	(.0075)	(.0158)	(.0071)	(.0144)	(.0066)	(.0135)

TABLE 2

Efficiency (EFF) and expected selected size (ES) (based on simulation) of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta, \theta+\Delta, \ldots, \theta+(k-1)\Delta$); the common variance = 1, common sample size n = 5 and the prior for ψ^{B*} and ψ^{B} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .90

	NORAML		LOGISTIC		LAPLACE		GROSS ERROR	
	EFF	ES	EFF	ES	EFF	ES	EFF	ES
ψ^{B*}	0.3788	2.3686	0.3917	2.3955	0.3840	2.3927	0.3857	2.4089
	(.0071)	(.0226)	(.0068)	(.0226)	(.0066)	(.0227)	(.0066)	(.0233)
ψ^{B}	0.3568	2.7630	0.3663	2.7810	0.3549	2.8160	0.3554	2.8110
	(.0064)	(.0265)	(.0061)	(.0265)	(.0058)	(.0265)	(.0058)	(.0269)
ψ^{M}	0.2729	3.9520	0.2733	3.9950	0.2705	4.0130	0.2646	4.0360
	(.0044)	(.0328)	(.0044)	(.0326)	(.0041)	(.0326)	(.0040)	(.0327)
ψ^{MED}	0.2513	4.1420	0.2513	4.2710	0.2366	4.4640	0.2224	4.6450
	(.0035)	(.0308)	(.0034)	(.0292)	(.0031)	(.0272)	(.0020)	(.0217)
ψ^{B*}	0.5799	1.7930	0.5826	1.7638	0.5682	1.8109	0.5555	1.7912
	(.0082)	(.0188)	(.0081)	(.0182)	(.0079)	(.0185)	(.0078)	(.0173)
ψ^{B}	0.5278	2.1140	0.5318	2.0990	0.5157	2.1610	0.5040	2.1300
	(.0079)	(.0233)	(.0080)	(.0232)	(.0075)	(.0236)	(.0073)	(.0218)
ψ^{M}	0.4023	2.8930	0.4069	2.8680	0.3959	2.9290	0.3907	2.9080
	(.0064)	(.0324)	(.0065)	(.0325)	(.0063)	(.0316)	(.0063)	(.0308)
ψ^{MED}	0.3544	3.2480	0.3566	3.2910	0.3274	3.4000	0.3196	3.4260
	(.0057)	(.0332)	(.0060)	(.0342)	(.0047)	(.0308)	(.0038)	(.0284)
ψ^{B*}	0.8288	1.2584	0.8167	1.2736	0.8242	1.2586	0.8274	1.2614
	(.0070)	(.0109)	(.0070)	(.0108)	(.0069)	(.0104)	(.0070)	(.0111)
ψ^{B}	0.7728	1.4570	0.7565	1.4890	0.7577	1.4880	0.7697	1.4700
	(.0082)	(.0169)	(.0082)	(.0167)	(.0081)	(.0167)	(.0082)	(.0174)
ψ^{M}	0.6524	1.7770	0.6318	1.8320	0.6418	1.8070	0.6363	1.8170
	(.0082)	(.0206)	(.0080)	(.0207)	(.0082)	(.0210)	(.0081)	(.0208)
ψ^{MED}	0.5840	2.0050	0.5694	2.0400	0.5533	2.0480	0.5393	2.0630
	(.0081)	(.0239)	(.0077)	(.0228)	(.0070)	(.0206)	(.0065)	(.0190)
ψ^{B*}	0.9811	1.0284	0.9821	1.0266	0.9847	1.0229	0.9796	1.0296
	(.0025)	(.0039)	(.0024)	(.0037)	(.0023)	(.0035)	(.0027)	(.0040)
ψ^{B}	0.9670	1.0660	0.9685	1.0630	0.9750	1.0500	0.9675	1.0650
	(.0039)	(.0079)	(.0038)	(.0077)	(.0034)	(.0069)	(.0039)	(.0078)
ψ^{M}	0.9160	1.1680	0.9170	1.1660	0.9295	1.1410	0.9335	1.1330
	(.0059)	(.0118)	(.0059)	(.0118)	(.0055)	(.0110)	(.0054)	(.0107)
ψ^{MED}	0.8417	1.3180	0.8625	1.2750	0.8850	1.2300	0.8910	1.2180
	(.0074)	(.0149)	(.0071)	(.0141)	(.0067)	(.0133)	(.0065)	(.0131)

TABLE 3

Efficiency (EFF) and expected selected size (ES) (based on simulation of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption, when the unknown means of the k populations are $\theta,\ldots,\theta,\theta+\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .99

$\Delta\sqrt{n}$		NORAML EFF	NORAML ES	LOGISTIC EFF	LOGISTIC ES	LAPLACE EFF	LAPLACE ES	GROSS ERROR EFF	GROSS ERROR ES
0.5	ψ^{B*}	0.2278	4.1757	0.2201	4.1722	0.2269	4.1552	0.2262	4.2047
		(.0025)	(.0202)	(.0023)	(.0193)	(.0025)	(.0208)	(.0025)	(.0222)
	ψ^{B}	0.2190	4.4940	0.2129	4.5150	0.2178	4.4880	0.2187	4.5090
		(.0021)	(.0214)	(.0019)	(.0202)	(.0021)	(.0218)	(.0022)	(.0227)
	ψ^{M}	0.2032	4.9360	0.2021	4.9360	0.2043	4.9170	0.2044	4.8840
		(.0005)	(.0089)	(.0007)	(.0096)	(.0011)	(.0109)	(.0013)	(.0136)
	ψ^{MED}	0.2031	4.9370	0.2008	4.9730	0.2009	4.9790	0.2000	4.9990
		(.0011)	(.0094)	(.0004)	(.0053)	(.0003)	(.0050)	(.0001)	(.0010)
1	ψ^{B*}	0.2448	4.0727	0.2488	4.0381	0.2432	4.0505	0.2446	4.1141
		(.0021)	(.0208)	(.0025)	(.0219)	(.0023)	(.0210)	(.0025)	(.0221)
	ψ^{B}	0.2297	4.3970	0.2344	4.3710	0.2291	4.3720	0.2302	4.4350
		(.0017)	(.0218)	(.0021)	(.0235)	(.0019)	(.0219)	(.0022)	(.0229)
	ψ^{M}	0.2045	4.9190	0.2057	4.8870	0.2044	4.9060	0.2062	4.8820
		(.0006)	(.0100)	(.0009)	(.0126)	(.0007)	(.0108)	(.0012)	(.0130)
	ψ^{MED}	0.2033	4.9190	0.2039	4.9380	0.2009	4.9670	0.2006	4.9880
		(.0007)	(.0104)	(.0010)	(.0096)	(.0005)	(.0063)	(.0002)	(.0037)
2	ψ^{B*}	0.3015	3.6352	0.3132	3.5174	0.3079	3.5926	0.3002	3.6828
		(.0040)	(.0280)	(.0044)	(.0292)	(.0044)	(.0289)	(.0044)	(.0291)
	ψ^{B}	0.2765	3.9780	0.2849	3.8680	0.2806	3.9430	0.2752	4.0390
		(.0038)	(.0302)	(.0041)	(.0312)	(.0041)	(.0310)	(.0042)	(.0311)
	ψ^{M}	0.2254	4.6700	0.2290	4.6210	0.2305	4.6280	0.2319	4.6310
		(.0026)	(.0230)	(.0028)	(.0239)	(.0031)	(.0249)	(.0034)	(.0250)
	ψ^{MED}	0.2153	4.7680	0.2144	4.8030	0.2079	4.8880	0.2024	4.9700
		(.0016)	(.0183)	(.0020)	(.0176)	(.0014)	(.0133)	(.0009)	(.0074)
4	ψ^{B*}	0.6415	1.9744	0.6361	2.0027	0.6404	1.9671	0.6519	1.9470
		(.0092)	(.0312)	(.0092)	(.0319)	(.0091)	(.0308)	(.0092)	(.0316)
	ψ^{B}	0.5908	2.2360	0.5869	2.2560	0.5874	2.2350	0.5988	2.2120
		(.0097)	(.0360)	(.0097)	(.0363)	(.0096)	(.0356)	(.0097)	(.0365)
	ψ^{M}	0.4715	2.9100	0.4696	2.9500	0.4738	2.9150	0.4701	2.9310
		(.0093)	(.0445)	(.0094)	(.0451)	(.0094)	(.0447)	(.0093)	(.0451)
	ψ^{MED}	0.3556	3.6050	0.3376	3.7610	0.3124	3.9140	0.2931	4.0280
		(.0074)	(.0428)	(.0072)	(.0420)	(.0063)	(.0398)	(.0055)	(.0371)

TABLE 4

Efficiency (EFF) and Expected selected Size (ES) (based on simulation of ψ^{B*}, ψ^{B}, ψ^{M} and ψ^{MED}, under normal assumption when the unknown means of the k populations are $\theta,\ldots,\theta,\theta+\Delta$; the common variance = 1, common sample size n = 5 and the prior for ψ^{B} and ψ^{B*} is the locally uniform prior. The numbers in the paranthesis are the standard errors of the corresponding estimates.

k = 5, P* = .90

$\sqrt{n}\Delta$		NORAML		LOGISTIC		LAPLACE		GROSS ERROR	
		EFF	ES	EFF	ES	EFF	ES	EFF	ES
0.5	ψ^{B*}	0.2660	2.7314	0.2767	2.7754	0.2721	2.7695	0.2749	2.8053
		(.0061)	(.0224)	(.0062)	(.0220)	(.0059)	(.0228)	(.0059)	(.0243)
	ψ^{B}	0.2536	3.1510	0.2650	3.1960	0.2626	3.1810	0.2641	3.2170
		(.0053)	(.0257)	(.0056)	(.0253)	(.0054)	(.0265)	(.0054)	(.0281)
	ψ^{M}	0.2151	4.4790	0.2248	4.4960	0.2241	4.4600	0.2191	4.4670
		(.0026)	(.0258)	(.0031)	(.0257)	(.0031)	(.0269)	(.0030)	(.0277)
	ψ^{MED}	0.2103	4.5330	0.2164	4.6500	0.2121	4.7440	0.2037	4.9190
		(.0023)	(.0241)	(.0024)	(.0213)	(.0023)	(.0202)	(.0011)	(.0114)
1	ψ^{B*}	0.3532	2.5723	0.3591	2.5943	0.3558	2.6086	0.3460	2.6783
		(.0067)	(.0236)	(.0068)	(.0237)	(.0069)	(.0240)	(.0064)	(.0246)
	ψ^{B}	0.3207	2.9760	0.3321	2.9970	0.3288	3.0260	0.3212	3.0840
		(.0057)	(.0265)	(.0061)	(.0271)	(.0061)	(.0273)	(.0059)	(.0281)
	ψ^{M}	0.2442	4.2790	0.2490	4.2880	0.2487	4.3190	0.2440	4.3280
		(.0034)	(.0302)	(.0041)	(.0307)	(.0039)	(.0308)	(.0045)	(.0321)
	ψ^{MED}	0.2378	4.3550	0.2351	4.4700	0.2177	4.6950	0.2093	4.8580
		(.0032)	(.0277)	(.0033)	(.0267)	(.0021)	(.0207)	(.0019)	(.0162)
2	ψ^{B*}	0.5418	2.1015	0.5553	2.0727	0.5412	2.1065	0.5381	2.1113
		(.0082)	(.0252)	(.0083)	(.0256)	(.0082)	(.0255)	(.0084)	(.0261)
	ψ^{B}	0.4829	2.4640	0.4998	2.4140	0.4874	2.4540	0.4854	2.4620
		(.0080)	(.0299)	(.0083)	(.0304)	(.0082)	(.0302)	(.0082)	(.0307)
	ψ^{M}	0.3546	3.5630	0.3607	3.5430	0.3562	3.5710	0.3462	3.5960
		(.0072)	(.0415)	(.0075)	(.0425)	(.0075)	(.0419)	(.0070)	(.0413)
	ψ^{MED}	0.3159	3.8080	0.3045	3.9580	0.2736	4.2210	0.2550	4.3540
		(.0062)	(.0390)	(.0060)	(.0390)	(.0052)	(.0346)	(.0041)	(.0311)
4	ψ^{B*}	0.9129	1.1682	0.9037	1.1849	0.9261	1.1434	0.9191	1.1547
		(.0056)	(.0124)	(.0059)	(.0129)	(.0053)	(.0117)	(.0057)	(.0126)
	ψ^{B}	0.8787	1.2770	0.8664	1.3040	0.8948	1.2410	0.8912	1.2510
		(.0072)	(.0176)	(.0074)	(.0182)	(.0068)	(.0168)	(.0069)	(.0173)
	ψ^{M}	0.7689	1.6630	0.7709	1.6680	0.7931	1.5820	0.8033	1.5530
		(.0094)	(.0317)	(.0095)	(.0322)	(.0091)	(.0298)	(.0090)	(.0299)
	ψ^{MED}	0.6261	2.2180	0.6523	2.0910	0.7060	1.8970	0.7192	1.8350
		(.0103)	(.0407)	(.0102)	(.0388)	(.0100)	(.0369)	(.0099)	(.0350)