

AD-A100 948

PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS

F/6 12/1

SOME RESULTS IN THE THEORY OF SUBSET SELECTION PROCEDURES. (U)

AUG 80 H YANG

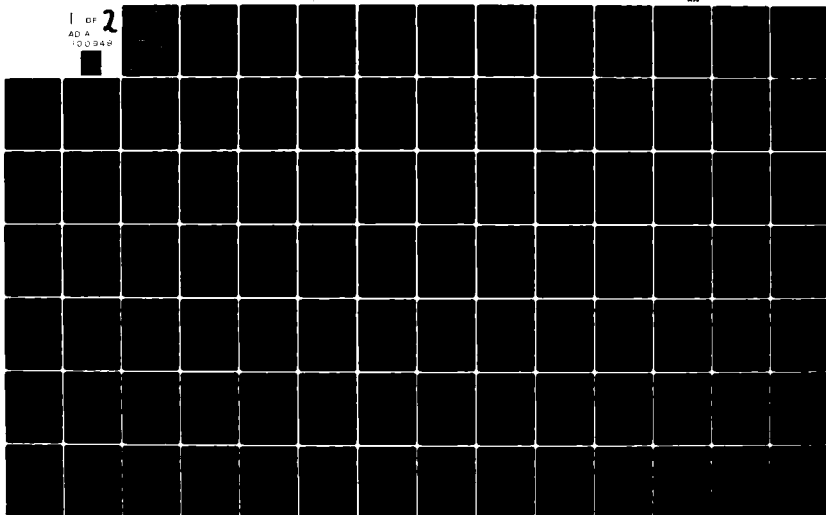
N00014-75-C-0455

UNCLASSIFIED

MMS-80-22

AM

1 OF 2
AD A
100948

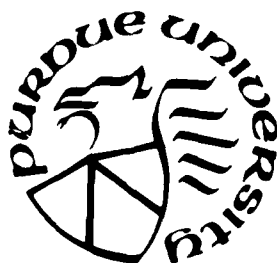


AD A100948

LEVEL II

①

PURDUE UNIVERSITY



DTIC
ELECTE
JUL 0 6 1981
S **D**
E

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

DTIC FILE COPY

81 6 29 243

LEA II

①

Some Results in the Theory
of Subset Selection Procedures

by

Hwa-Ming Yang
Purdue University

Department of Statistics
Division of Mathematical Sciences
Micrograph Series 0-22
August, 1980

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

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

Accession For	
NTIS GRA&I	X
DTIC TAB	
Unannounced	
Justification	
By	
Distribution	
Availability	
Dissem.	
A	

INTRODUCTION

Selection and ranking (ordering) problems in statistical inference arise mainly because the classical tests of homogeneity are often inadequate in certain situations where the experimenter is interested in comparing k (≥ 2) populations, treatments or processes with the goal of selecting one or more worthwhile (good) populations. Mosteller (1947), Paulson (1949), Bahadur (1950) and Bahadur and Pabbins (1950) were among the earliest research workers to recognize this inadequacy and to formulate the problem as a multiple decision problem aimed at the selection and ranking of the k populations.

In the thirty years since these early papers, selection and ranking problems have become an active area of statistical research. There have been two approaches to these problems, the 'indifference zone' approach and the 'subset selection' approach. In the first approach, due to Bechhofer (1954), the experimenter wishes to select one population (or a fixed number $t \geq 1$ of population) which is guaranteed to be the one of interest to him with a fixed probability P^* whenever the unknown parameters lie outside some subspace of the parameter space, the so-called indifference zone. Important contributions using this approach have been made by Bechhofer and Sobel (1954), Bechhofer, Dunnett and Sobel (1954), Sobel and Huyett (1957), Sobel (1967), Bechhofer, Fretter and Sobel (1968), Mahabunulu (1967), Dosu and Sobel (1968, 1971) and

Tamhane and Bechhofer (1977, 1979) among others. A quite complete bibliography may be found in Gupta and Panchapakesan (1979) (see also Gibbons, Olkin and Sobel (1977)).

The second approach pioneered by Gupta (1956, 1963, 1965) assumes no a priori information about the parameter space. A single population is not necessarily chosen; rather a subset of the given k populations is selected depending on the outcome of the experiment. It is guaranteed to contain the population(s) of interest with probability which is at least equal to P^* (the basic probability requirement) regardless of the true unknown configurations of the parameters. Some recent contributors in the category of subset selection include: Deely (1965), Gnanadesikan (1966), Gnanadesikan and Gupta (1970), Gupta (1967), Gupta and Studden (1970), Nagel (1970), Gupta and Nagel (1971), Gupta and Panchapakesan (1972), Rizvi and Sobel (1967), McDonald (1969), Gupta and McDonald (1970), Santner (1975), W. T. Huang (1972), D. Y. Huang (1975), Gupta and Huang (1975a, 1975b) and Gupta and Huang (1976).

Subset selection procedures can also be thought of as screening procedures which enable the experimenter to select a subset of populations (under study) which contains the populations of interest so that the populations in the selected subset can be further studies.

Sequential and multistage aspects of the ranking and selection problems, have been explored, based on the indifference zone approach by Bechhofer, Dunnett and Sobel (1954), Bechhofer (1958), Paulson (1962, 1963, 1964, 1967) and Bechhofer, Fiefer and Sobel (1968). Parson and Gupta (1972), Huang (1972), Gupta and Huang (1975), Gupta and Miescke (1979) and Carroll (1974) have investigated subset selection procedures, based on sequential sampling.

Contributions to optimum properties of subset selection procedures have been made by Lehmann (1961), Studden (1967), Deely and Gupta (1968), Berger (1977, 1979), Gupta and Hsu (1978), Gupta and Miescke (1978), Berger and Gupta (1980).

In the decision-theoretic approach to the subset selection problem, Goel and Rubin (1977), Chernoff and Yahav (1977), Bickel and Yahav (1977), Gupta and Hsu (1978), Miescke (1979), Gupta and Kim (1980), Gupta and Hsiao (1980) have given different formulations under different loss functions and carried out investigations which indicate that the Gupta type maximum (minimum) means procedures are quite 'optimal' and 'robust'.

The main purpose of this thesis is to study some problems using the subset selection approach and provide procedures and results for some unsolved problems.

Chapter I considers the problem of selecting a subset containing all populations better than a control under an ordering prior. Here, by an ordering prior we mean that there exists a known simple or partial order relationship among the unknown parameters of the treatments (excluding the control). Three new selection procedures are proposed and studied. These procedures do meet the usual requirement that the probability of a correct selection is greater than or equal to a pre-determined number P^* . Two of the three procedures use the isotonic regression over the sample means of the k treatments with respect to (wrt) the given ordering prior. Tables which are necessary to carry out the selection procedures with isotonic approach for the selection of unknown means of normal populations and gamma populations are given. Monte Carlo comparisons on the performance of several procedures for the normal or gamma means problem were carried out in

several selected cases; these are given in Table V and Table VI at the end of Chapter I. In each case ten thousand simulations were performed. The results of this study seem to indicate that the procedures based on isotonic estimators always have superior performance, especially, when there are more than one bad populations (in comparison with the control).

Chapter II deals with a new 'Bayes- P^* ' approach about the problem of selecting a subset which contains the 'best' of k populations. Here, by best we mean the (unknown) population with the largest unknown mean. The (non-randomized) Bayes- P^* rule refers to a rule with minimum risk in the class of (non-randomized) rules which satisfy the condition that the posterior probability of selecting the best is at least equal to P^* . Given the priors of the unknown parameters, two 'Bayes- P^* ' subset selection procedures ϕ^B and ϕ_{NR}^B (randomized and non-randomized, respectively) under certain loss functions are obtained and compared with the classical maximum-type means procedure ϕ^M . The comparisons of the performance of ϕ^B with ϕ_{NR}^B and ϕ^M , based on Monte Carlo studies, indicate that the procedure ϕ^B always has higher 'efficiency' and smaller expected selected size of the selected subset. Also ϕ^B appears to be robust when the true distributions are not normal but are some other symmetric distributions such as, the logistic, the double exponential, Laplace, and the gross error model (the contaminated distribution).

CHAPTER I
SELECTION PROCEDURES FOR POPULATIONS
BETTER THAN A CONTROL UNDER ORDERING PRIOR

1.1. Introduction

In this chapter, three new selection procedures are given for the problem of selecting a subset which contains all populations better than a standard or control under simple or partial ordering prior. Here by simple or partial ordering prior we mean that there exist known simple or partial order relationships (defined more specifically later in Section 1.2) among unknown parameters. The procedures described do meet the usual requirement that the probabilities of a correct selection are greater than or equal to a predetermined number P^* , the so-called P^* condition.

Many authors have considered the problem of comparing populations with a control under different types of formulations (see Gupta and Panchapakesan (1979)). Dunnett (1955) considered the problem of separating those treatments which are better than the control from those that are worse. Gupta and Sobel (1958), Gupta (1965), Haik (1974), Broström (1977) studied the problem of selecting a subset containing all populations better than the control. Lehmann (1961) discussed similar problems with emphasis on the derivation of a restricted minimax procedure. Kim (1979), Hsiao (1979) studied the problem of

selecting populations close to a control. In all these papers it is assumed that all populations are independent and that there is no information about the order of unknown parameters. However, in many situations, we may know something about the unknown parameters. What we know is always not the prior distributions but some partial or incomplete prior information, such as the simple or partial order relationship among the unknown parameters. This type of information about the ordering prior may come from the past experiences; or it may arise in the experiments where, for example, higher dose level of some drugs always has larger effect (side-effect) on the patients.

In Section 1.2 definitions and notations used in this chapter are introduced. In Section 1.3 we consider the problem for location parameters. We propose three types of selection procedures for the cases when the control parameter is known or not known (the scale parameter may or may not be assumed known). Some equivalent forms of the procedures are given, and their properties are discussed. In Section 1.4 the problem for scale parameters of the gamma distributions is considered and three analogous selection procedures are proposed. In both Section 1.3 and 1.4 simple ordering priors are assumed and some theorems in the theory of random walks are used. In Section 1.5 a selection procedure is given for the problem of selecting all populations better than the control under partial ordering prior. Section 1.6 deals with the use of Monte Carlo techniques to make comparisons among the selection procedures proposed in Section 1.3 and those in Section 1.4, respectively.

1.2. Notations and Definitions

Suppose we have $k + 1$ populations $\pi_0, \pi_1, \dots, \pi_k$. The population treatment π_0 is called the control or standard population. Assume that the random variables X_{ij} associated with $F(\cdot; \theta_i)$ and $X_{i1}, \dots, X_{in_i}$, $i = 1, \dots, k$, is an independent sample from π_i . Assume that we have an ordering prior of $\pi_1, \dots, \pi_k$. First we assume that the ordering prior is the simple order, so that without loss of generality, we may assume that, $\theta_1 \leq \dots \leq \theta_k$. In Section 1.5 we will consider the partial ordering prior case. Note that the values of θ_i 's are unknown.

Suppose our goal is to find a non-trivial (small) subset which contains all populations with parameter larger (smaller) than the control θ_0 (known or unknown) with probability not less than a given value $1 - \alpha$.

The action space $\mathcal{A}$ is the class of all subsets of set $\{1, 2, \dots, k\}$. An action A is the selection of some subset of the k populations. $i \in A$ means that π_i is included in the selected subset.

Let $\theta = (\theta_0, \theta_1, \dots, \theta_k)$. Then the parameter space is denoted by Ω , where $\Omega = \{\theta \in \mathbb{R}^{k+1} \mid \theta_1 \leq \theta_2 \leq \dots \leq \theta_k; -\infty < \theta_0 < \infty\}$ is a subset of $k + 1$ dimensional Euclidean space $\mathbb{R}^{k+1}$.

The sample space is denoted by $\mathcal{X}$ where

$$\mathcal{X} = \{x \in \mathbb{R}^{n_1 + \dots + n_k} \mid x = (x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{kn_1}, \dots, x_{kn_1})\}.$$

Definition 1.2.1. A (non-randomized) selection procedure (rule) δ is a mapping from $\mathcal{X}$ to $\mathcal{A}$.

A population π_i ($i = 1, \dots, k$) is called a good population if $\pi_i \neq \pi_0$, and we say a selection procedure δ make a correct selection (CS) if the selected subset contains all good populations. A selection procedure δ satisfies the P^* -condition if

$$P_{\pi_i}(\text{CS}|\delta) \geq P^* \text{ for all } \pi_i \in \pi$$

that is

$$\inf_{\pi \in \pi} P_{\pi}(\text{CS}|\delta) \geq P^*. \quad (1.2.1)$$

Let $\mathcal{A} = \{\delta | \inf_{\pi \in \pi} P_{\pi}(\text{CS}|\delta) \geq P^*\}$ be a collection of all selection procedures satisfying the P^* -condition.

In the sequel we will use the isotonic estimators (see Barlow, Bartholomew, Bremner and Brunk (1972)). Hence we give the following definitions and theorems.

Definition 1.2.2. Let the set $\mathcal{J}$ be a finite set. A binary relation " $\leq$ " on $\mathcal{J}$ is called a simple order if it is

- (1) reflexive: $x \leq x$ for $x \in \mathcal{J}$
- (2) transitive: $x, y, z \in \mathcal{J}$ and $x \leq y, y \leq z$ imply $x \leq z$
- (3) antisymmetric: $x, y \in \mathcal{J}$ and $x \leq y, y \leq x$ imply $x = y$
- (4) every two elements are comparable: $x, y \in \mathcal{J}$ imply either $x \leq y$ or $y \leq x$.

A partial order on $\mathcal{J}$ is a binary relation " $\leq$ " on $\mathcal{J}$, such that it is (1) reflexive, (2) transitive, and (3) antisymmetric. Thus every simple order is a partial order. We use poset $(\mathcal{J}, \leq)$ to denote the set $\mathcal{J}$ that has a partial order binary relation " $\leq$ " on it.

Definition 1.2.3. A real-valued function f is called isotonic on poset $(\mathcal{J}, \preceq)$ if and only if (1) f is defined on $\mathcal{J}$, (2) if $x, y \in \mathcal{J}$, $x \preceq y$, imply $f(x) \leq f(y)$.

Definition 1.2.4. Let q be a real-valued function on $\mathcal{J}$ and let W be a given positive function on $\mathcal{J}$. A function q^* on $\mathcal{J}$ is called an isotonic regression of q with weights W if and only if:

- (1) q^* is an isotonic function on poset $(\mathcal{J}, \preceq)$
- (2) $\int_{x \in \mathcal{J}} [q(x) - q^*(x)]^2 W(x) = \min_{f \in \mathcal{G}} \int_{x \in \mathcal{J}} [q(x) - f(x)]^2 W(x)$,

where $\mathcal{G}$ is the class of all isotonic functions on poset $(\mathcal{J}, \preceq)$.

From Barlow, et. al. (1972), (see their Theorems 1.2, 1.6 and the corollary there), we have the following theorems.

Theorem 1.2.1. There exists one and only one isotonic regression q^* of q with weight W on poset $(\mathcal{J}, \preceq)$.

Definition 1.2.5. A set S is convex if s_1 and $s_2 \in S$ and $0 \leq \alpha \leq 1$ then $\alpha s_1 + (1-\alpha)s_2 \in S$.

Definition 1.2.6. A set S is a cone if $s \in S$ then for any non-negative real number c , $cs \in S$.

Definition 1.2.7. A poset $(\mathcal{J}, \preceq)$ is a lattice if $\inf H$ and $\sup H$ exist for any finite non-empty subset H of $\mathcal{J}$.

If f and g are two isotonic functions on poset $(\mathcal{J}, \preceq)$, we define $f \wedge g$ and $f \vee g$ as

$$(f \wedge g)(t) = f(t) \wedge g(t) = \min(f(t), g(t))$$

and

$$(f \vee q)(t) = f(t) \vee q(t) = \max(f(t), q(t)).$$

Then we state the following:

Theorem 1.2.2. The class $\mathcal{C}$ of all isotonic functions on poset $(\mathcal{L}, \leq)$ is a convex cone and a lattice.

There are some algorithms, such as the "pool-adjacent-violation" algorithm (see page 13 of Barlow, et. al. (1972)) or Ayer, Brunk, Iwamoto, Reid and Silverman (1955) or the "up-and-down blocks" algorithm, Serfling (1964), which show how to calculate the isotonic regression under simple order.

The following max-min formulas were given by Ayer et. al. (1955).

Theorem 1.2.3. (max-min formulas)

Assume that we have poset $(\mathcal{L}, \leq)$ where $\mathcal{L} = \{1, \dots, k + 1\}$ and that function $q: \mathcal{L} \rightarrow \mathbb{R}$, then the isotonic regression q^* of q with weight W has the following formulas:

$$\begin{aligned} q^*(i) &= \max_{s \leq i} \min_{t \geq i} Av(s, t) \\ &= \max_{s \leq i} \min_{t \geq s} Av(s, t) \\ &= \min_{t \geq i} \max_{s \leq i} Av(s, t) \\ &= \min_{t \geq i} \max_{s \leq t} Av(s, t) \end{aligned}$$

where

$$Av(s, t) = \frac{\sum_{i=s}^t q(i) W(i)}{\sum_{i=s}^t W(i)}.$$

Corollary 1.2.1. $(q + c)^* = q^* + c$,
 $(aq)^* = aq^*$, if $a \geq 0$.

Corollary 1.2.2. $[(q^*)q + \zeta(q^*)]^* = (q^*)q^* + \zeta(q^*)$, where ζ is a non-negative function and ζ is an arbitrary function.

1.3. Proposed Selection Procedures for the Location Parameter Problem

To discuss some more general results, we assume that population π_i has an absolutely continuous location-scale distribution function $F(x; \mu_i, \sigma_i^2) = F(\frac{x - \mu_i}{\sigma_i})$, where $0 < F(x) < 1$ for all x , $\mu_i \in \mathbb{R}$ and the values of $\mu_1, \dots, \mu_k$ are unknown, but their ordering, say, $\mu_1 \leq \dots \leq \mu_k$ is known. Note that in this case we replace μ in the parameter space Ω by μ , all other quantities remaining the same.

Let us define the subspace $\pi_i = \{\mu \in \mathbb{R}^1 : \mu_{k-i+1} = \mu_0, \mu_{i+1} = \mu_0\}$, $i = 1, \dots, k-1$ and let subspace $\pi_k = \{\mu \in \mathbb{R}^1 : \mu_0 = \mu_1\}$ and subspace $\pi_0 = \{\mu \in \mathbb{R}^1 : \mu_k = \mu_0\}$, then we have $\Omega = \bigcup_{i=0}^k \pi_i$. Note that the control π_0 could be known or unknown. If π_0 is unknown, we assume that the distribution of population π_0 is $F(x; \mu_0, \sigma_0^2)$ and we take independent observations $X_{01}, \dots, X_{0n_0}$ from π_0 and the sample space $\mathcal{X}$ turns to

$\{x \in \mathbb{R}^{n_0 + \dots + n_k} : x = (X_{01}, \dots, X_{0n_0}, X_{11}, \dots, X_{kn_k})\}$. Using the partition $\{\pi_0, \dots, \pi_k\}$ of parameter space Ω , we have

$$\inf_{\mu \in \Omega} P_\mu(CS_1^*) = \inf_{1 \leq i \leq k} \inf_{\mu \in \pi_i} P_\mu(CS_1^*),$$

for any selection procedure $\delta \in \mathcal{A}$. Hence the P^* -condition is equivalent to

$$\inf_{\mu \in \mu_i} P_{\mu}(CS \leq \alpha) \leq P^*, \text{ for } i = 1, \dots, k.$$

Note that $\inf_{\mu \in \mu_0} P_{\mu}(CS \leq \alpha) = 1$ for any selection procedure $\cdot$ since there exist no good population in this case.

Suppose $X_i = x_i$ is the outcome of the sample mean of population μ_i , $i = 1, \dots, k$. Let J denote the set $\{\mu_1, \mu_2, \dots, \mu_k\}$ where $\mu_1 \leq \dots \leq \mu_k$, and let $W(\mu_i) = n_i \sigma^{-2} = w_i$, $g(\mu_i) = x_i$, $i = 1, \dots, k$. Then by the max-min formulas, the isotonic regression of g is g^* , where

$$g^*(\mu_i) = \max_{1 \leq s \leq i} \min_{s \leq t \leq k} \frac{\sum_{j=s}^t x_j w_j}{\sum_{j=s}^t w_j}, \quad i = 1, \dots, k.$$

The isotonic estimator of μ_i is denoted by $X_{i:k}$, $i = 1, \dots, k$ where

$$\begin{aligned} X_{i:k} &= \max_{1 \leq s \leq i} \min_{s \leq t \leq k} \frac{\sum_{j=s}^t X_j w_j}{\sum_{j=s}^t w_j} \\ &= \max_{1 \leq j \leq i} \bar{X}_{j:k} \end{aligned} \quad (1.3.1)$$

where

$$\bar{X}_{j:k} = \min \left\{ X_j, \frac{X_j w_j + X_{j+1} w_{j+1}}{w_j + w_{j+1}}, \dots, \frac{X_j w_j + \dots + X_k w_k}{w_j + \dots + w_k} \right\}. \quad (1.3.2)$$

1.3.1. Proposed Selection Procedure $\cdot_1$

Case I. μ_0 known, common variance σ^2 known, and common sample size n .

Definition 1.3.1. We define the procedure Δ_1 as follows:

Step 1. Select μ_1 , $i = 1, \dots, k$ and stop, if

$$\hat{x}_{1:k} \geq \mu_0 - d_{1:k}^{(1)} \frac{1}{\sqrt{n}},$$

otherwise reject μ_1 and go to step 2.

Step 2. Select μ_2 , $i = 2, \dots, k$ and stop, if

$$\hat{x}_{2:k} \geq \mu_0 - d_{2:k}^{(1)} \frac{1}{\sqrt{n}},$$

otherwise reject μ_2 and go to step 3.

⋮

Step $k-1$. Select μ_{k-1} , $i = k-1, k$ and stop, if

$$\hat{x}_{k-1:k} \geq \mu_0 - d_{k-1:k}^{(1)} \frac{1}{\sqrt{n}},$$

otherwise reject μ_{k-1} and go to step k .

Step k . Select μ_k and stop, if

$$\hat{x}_{k:k} \geq \mu_0 - d_{k:k}^{(1)} \frac{1}{\sqrt{n}},$$

otherwise reject μ_k .

Here $d_{i:k}^{(1)}$'s are the smallest values such that $\mu_1 \in \Delta$, that is, μ_1 satisfies the p^* -condition.

1.1.2. On the Evaluation of $\inf_{\mu \in \mathcal{C}_i} P_\mu(CS|\delta_1)$ and the Value of the

Constants $d_{1:k}^{(1)}, \dots, d_{k:k}^{(1)}$

For any $\mu \in \mathcal{C}_i$, $1 \leq i \leq k$, let Z_i 's i.i.d. $F(\cdot; 0, 1)$ then

$$\begin{aligned} P_\mu(CS|\delta_1) &= P_\mu\left(\bigcup_{j=1}^{k-i+1} \{X_{j:k} \geq \mu_0 - d_{j:k}^{(1)} \frac{\mu_0}{\sqrt{n}}\}\right) \\ &= P_\mu\left(\bigcup_{j=1}^{k-i+1} \bigcup_{r=1}^j \{X_{r:k} \geq \mu_0 - d_{j:k}^{(1)} \frac{\mu_0}{\sqrt{n}}\}\right) \\ &= P_\mu\left(\bigcup_{j=1}^{k-i+1} \bigcup_{r=1}^j \{Z_{r:k} + \frac{\mu_0 - \mu}{\sqrt{n}} \geq -d_{j:k}^{(1)}\}\right) \end{aligned}$$

which is decreasing in μ_r , $r = 1, \dots, k-i+1$.

Hence

$$\inf_{\mu \in \mathcal{C}_i} P_\mu(CS|\delta_1) = P(Z_{k-i+1:k} \geq -d_{k-i+1:k}^{(1)})$$

On the other hand,

$$\begin{aligned} \inf_{\mu \in \mathcal{C}_i^*} P_\mu(CS|\delta_1) &= P_{\mu^*}\left(\bigcup_{j=1}^{k-i+1} \{X_{j:k} \geq \mu_0 - d_{j:k}^{(1)} \frac{\mu_0}{\sqrt{n}}\}\right) \\ &= P(Z_{k-i+1:k} \geq -d_{k-i+1:k}^{(1)}) \end{aligned}$$

whenever $\mu^* = (\underbrace{\mu_0, \dots, \mu_0}_{i}, \underbrace{0, \dots, 0}_{k-i}) \in \mathcal{C}_i^*$

Thus, we have

$$\inf_{\mu \in \mathcal{C}_i} P_\mu(CS|\delta_1) = P(Z_{k-i+1:k} \geq -d_{k-i+1:k}^{(1)}).$$

Since $\hat{Z}_{k-i+1:k} = \min \{Z_{k-i+1}, \dots, \frac{Z_{k-i+1} + \dots + Z_k}{i}\}$ has the same distributions as

$$\hat{Z}_{1:i} = \min \{Z_1, \dots, \frac{Z_1 + \dots + Z_i}{i}\},$$

let $V_i = \hat{Z}_{1:i}$ (1.3.3)

$$= \min_{1 \leq r \leq i} \frac{1}{r} \sum_{j=1}^r Z_j,$$

we have

$$\inf_{\mu \in \mathcal{C}_i} P_\mu(CS|s_1) = P(V_i \leq -d_{k-i+1:k}^{(1)}), \quad i = 1, \dots, k. \quad (1.3.4)$$

Theorem 1.3.1. In case I, (μ_0 known, common known σ^2 and common sample size n), if $d_{k-i+1:k}^{(1)}$ is the solution of equation

$$P(V_i \leq -x) = P^* \quad (1.3.5)$$

where

$$V_i = \min_{1 \leq r \leq i} \frac{1}{r} \sum_{j=1}^r Z_j \quad \text{and} \quad Z_j \text{ are i.i.d. } F(\cdot),$$

$i = 1, \dots, k$ then π_1 satisfies the P^* -condition.

Proof. For any $i, 1 \leq i \leq k$,

$$\inf_{\mu \in \mathcal{C}_i} P_\mu(CS|s_1) = P(V_i \leq -d_{k-i+1:k}^{(1)}) = P^*,$$

so π_1 satisfies the P^* -condition.

Therefore, the problem of finding the $d_{i:k}^{(1)}$'s reduces to finding the distributions of $V_1, \dots$, and V_k . This is achieved by using some theorems in the theory of random walk.

1.3.3. Some Theorems in the Theory of Random Walk

Suppose $Y_1, Y_2, \dots$ are independent random variables with a common distribution H not concentrated on a half-axis, i.e. $0 < P(Y_1 \leq 0) < P(Y_1 \geq 0) < 1$. The induced random walk is the sequence of random variables

$$S_0 = 0, S_n = Y_1 + \dots + Y_n, \quad n = 1, 2, \dots$$

Let

$$r_n = P(S_1 \leq 0, \dots, S_{n-1} \leq 0, S_n > 0) \quad (1.3.6)$$

and

$$r(s) = \sum_{n=1}^{\infty} r_n s^n, \quad 0 \leq s \leq 1. \quad (1.3.7)$$

Then we have the following theorem which was discovered by Andersen (1943). Feller (1971) gave an elegant short proof.

Theorem 1.3.2.

$$\log \frac{1}{1-r(s)} = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n = 0). \quad (1.3.8)$$

Theorem 1.3.3. (Feller (1971))

Let

$$p_n = P(S_1 > 0, \dots, S_n > 0),$$

then

$$p(s) = \sum_{n=1}^{\infty} p_n s^n = \frac{1}{1-r(s)}, \quad (1.3.9)$$

hence

$$\log p(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n = 0). \quad (1.3.10)$$

By symmetry, the probabilities

$$q_n = P(S_1 \leq 0, \dots, S_n \leq 0) \quad (1.3.11)$$

have the generating function q given by

$$\log q(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n \leq 0). \quad (1.3.12)$$

Note: The above two theorems remain valid if the signs $\leq$ and $\geq$ are replaced by $>$ and $<$, respectively.

Now, let

$$U_j^1 = \max_{1 \leq r \leq j} \frac{1}{r} \sum_{i=1}^r Z_i^1, \quad j = 1, 2, \dots, \quad (1.3.13)$$

and

$$V_j^1 = \min_{1 \leq r \leq j} \frac{1}{r} \sum_{i=1}^r Z_i^1, \quad j = 1, 2, \dots, \quad (1.3.14)$$

where Z_i^1 's are i.i.d. with absolutely continuous c.d.f. $G(\cdot)$. We would like to apply Theorem 1.3.3 to get the distribution of U_j^1 and V_j^1 , $j = 1, 2, \dots$.

Remark 1.3.1. The distribution of U_j^1 , $j = 1, \dots, k$ for some $k \geq 1$, will be used whenever our goal is changed to select a subset containing no population with parameter smaller than the control.

Theorem 1.3.4. The generating function $q(s)$ of $P(U_j^1 \leq x)$, $j = 1, \dots$ is

$$\sum_{j=1}^{\infty} s^j P(U_j^1 \leq x) = e^{-\sum_{n=1}^{\infty} \frac{1-s^n}{n} P(\bar{c}_n \leq 0)} \quad (1.3.15)$$

where

$$S_n = \sum_{i=1}^n (Z_i' - x), \quad n = 1, 2, \dots,$$

if the distribution of $Y_1 = Z_1' - x$ is not concentrated on a half-axis.

Proof. Since the distribution of random variable $Y_i = Z_i' - x$ is not concentrated on a half-axis, and Y_i 's are i.i.d. let $S_r = \sum_{i=1}^r (Z_i' - x)$, $r = 1, \dots, k$. Then

$$(U_j' - x) = \left(\max_{1 \leq r \leq j} \frac{1}{r} S_r \leq 0 \right) = (S_1 \leq 0, \dots, S_j \leq 0).$$

By Feller's Theorem 1.3.3, we complete the proof.

Similarly, $(V_j' - x) = (S_i \geq 0, \quad i = 1, 2, \dots, j)$,

where

$$S_i = \sum_{r=1}^i (Z_r' - x).$$

Theorem 1.3.5. The generating function $p(s)$ of $P(V_j' \geq x)$ is

$$\sum_{j=1}^{\infty} s^j P(V_j' \geq x) = \exp \left(\sum_{n=1}^{\infty} \frac{1}{n} s^n P(S_n \geq 0) \right), \quad (1.3.16)$$

if the distribution of $Y_1 = Z_1' - x$ is not concentrated on a half-axis.

Corollary 1.3.1. Both Theorem 1.3.4 and Theorem 1.3.5 hold for all x such that $0 < G(x) < 1$.

Proof. Let $Y_1 = Z_1^2 - x$, then

$$P(Y_1 < 0) = G(x)$$

and

$$0 < G(x) < 1,$$

hence Y_1 is not concentrated on a half-axis.

Corollary 1.3.2. Both Theorem 1.3.4 and Theorem 1.3.5 hold for all x whenever $G = \Phi$, c.d.f. of $N(0,1)$, or $G = F$ which is defined at the beginning of Section 1.3.

Proof. Followed immediately by Corollary 1.3.1.

Note that in the case of location parameter of normal population,

$$P(U_n^1 \geq -x) = P(V_n^1 \geq x).$$

Let

$$\Delta_j(x) = \Delta_j = P(S_j \geq 0), \quad j = 1, 2, \dots,$$

$$a(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} \Delta_n,$$

we have

$$p(s) = \sum_{j=1}^{\infty} s^j P(V_j^1 \geq x) = \exp(a(s)).$$

$$\text{Lemma 1.3.1. } p^{(n+1)}(s) = \sum_{j=0}^n \binom{n}{j} p^{(j)}(s) a^{(n+1-j)}(s), \quad \text{on.}$$

Proof. Since $p'(s) = p(s) \cdot a'(s)$, the result can be proved by induction on n .

Theorem 1.3.6. Under the assumption of Theorem 1.3.5

$$\begin{aligned}
 P(V_{n+1}^1 \leq x) &= \frac{1}{(n+1)!} \lim_{s \rightarrow 0^+} \frac{d^{n+1}}{ds^{n+1}} p(s) \\
 &= \frac{1}{n+1} \sum_{j=0}^n P(V_j^1 \leq x) \lambda_{n-j+1}, \quad n = 0, 1, 2, \dots \quad (1.3.17)
 \end{aligned}$$

where

$$P(V_0^1 \leq x) = 1, \quad \forall x.$$

Proof. By Lemma 1.3.1, we have

$$\begin{aligned}
 P(V_{n+1}^1 \leq x) &= \frac{1}{(n+1)!} \lim_{s \rightarrow 0^+} p^{(n+1)}(s) \\
 &= \sum_{j=0}^n \frac{1}{(n+1)!} \frac{n!}{j!(n-j)!} p^{(j)}(0) [(n-j)! \cdot \lambda_{n+1-j}] \\
 &= \frac{1}{n+1} \sum_{j=0}^n \frac{p^{(j)}(0)}{j!} \lambda_{n+1-j} \\
 &= \frac{1}{n+1} \sum_{j=0}^n P(V_j^1 \leq x) \lambda_{n+1-j}.
 \end{aligned}$$

Similarly, we have

$$P(U_{n+1}^1 \leq x) = \frac{1}{n+1} \sum_{j=0}^n P(U_{n-i+1}^1 \leq x) P(S_i = 0). \quad (1.3.18)$$

1.3.4. Limiting Distributions of U_n^1 and V_n^1

Let $F_n(x) = P(U_n^1 \leq x)$ and $F_+(x)$ denote the limiting distribution function as $n \rightarrow \infty$ of U_n^1 . Suppose the distribution of random variable $Z_1 = Z_1^1 \leq x$ is not concentrated on a half axis, then we have

$$1 - F_n(x) = P(S_1 > 0) + \sum_{r=2}^n P(S_1 = 0, \dots, S_{r-1} = 0, S_r = n),$$

$$1 - F_n(x) = \lim_{s \rightarrow 1^-} \varphi(s),$$

and apply Andersen-Feller Theorem 1.3.2, we have

$$F_n(x) = \exp \left\{ - \sum_{r=1}^n \frac{1}{r} P(S_r > 0) \right\}. \quad (1.3.19)$$

Similarly,

$$G_n(x) = \exp \left\{ - \sum_{r=1}^n \frac{1}{r} P(S_r = 0) \right\}. \quad (1.3.20)$$

where

$$G_n(x) = P(V_n' > x).$$

Let

$$G_n(-d_{1:i}^{(1)}) = P^*, \quad (1.3.21)$$

If Z_i , $i = 1, \dots, k$, are independent identically distributed

$N(0,1)$, then we can use the recurrence formula of Theorem 1.1.6 to

solve the equations $P(V_i > -d_{k-i+1:k}^{(1)}) = P^*$, $i = 1, \dots, k$. Hence in Case I, $\varphi_j(x) = 1(-x_j)$.

Remark 1.3.2. From formula (1.3.4) we know that $d_{k-i+1:k}^{(1)}$ ($i = 1, \dots, k$)

does not depend on k . And we have $d_{k-i+1:k}^{(1)} = d_{1:i}^{(1)}$. These values for

$k = 1(1)6, 10, \dots$ and $P^* = .99, .975, .95, .925, .9, .85, .8, .75,$

.7, .65 are tabulated in Table I.

1.3.5. Some Other Forms of Selection Procedure

Lemma 1.3.2. $d_{1:i}^{(1)}$ is increasing in i .

Proof. By Remark 1.3.2 and the fact

$$V_{i+1} = \min(V_i, \frac{iV_i + Z_{i+1}}{i+1}).$$

Lemma 1.3.3. If $c_j, 1 \leq j \leq i \leq k$ is decreasing in j , then

$$\bigcup_{j=1}^i \{X_{j:k} \geq -c_j\} = \bigcup_{j=1}^i \{X_{j:k} \geq -c_j\}.$$

Proof. $\bigcup_{j=1}^i \{X_{j:k} \geq -c_j\} = \bigcup_{j=1}^i \{X_{j:k} \geq -c_j\},$

since

$$X_{j:k} = \hat{X}_{j:k}, \quad 1 \leq j \leq k.$$

On the other hand, if

$$X_{r:k} \geq -c_r \quad \text{for some } r, \quad 1 \leq r \leq i$$

then

$$X_{s:k} \geq -c_r \quad \text{for some } s, \quad 1 \leq s \leq r,$$

since

$$X_{r:k} = \max_{1 \leq s \leq r} X_{s:k}.$$

Because c_j is decreasing in j , this implies $X_{s:k} \geq -c_s$ for some $s,$

$1 \leq s \leq r.$

Hence we have

$$\bigcup_{j=1}^i X_{j:k} \cap \{c_j\} = \bigcup_{j=1}^i X_{j:k} \cap \{c_j\},$$

therefore the lemma is proved.

Definition 1.3.2. We define a selection procedure δ_i^1 by replacing the inequality in the i th step of procedure δ_1 by the inequality

$$X_{i:k} \cap \{c\} = d_{i:k}^{(1)} \sqrt{n}, \quad i = 1, \dots, k$$

where $d_{i:k}^{(1)}, \dots, d_{k:k}^{(1)}$ are the smallest values such that δ_i^1 satisfies the P^* -condition.

Theorem 1.3.7. The selection procedure δ_1 and δ_i^1 are identical and $d_{i:k}^{(1)} = d_{i:k}^*$, $i = 1, \dots, k$.

Proof. For any i , $1 \leq i \leq k$, by Theorem 1.3.1

$$P^* = \inf_{\delta \in \delta_1} P\{(CS)_1^{(1)}\} = P\{Z_{k-i+1:k}^{(1)} \leq d_{k-i+1:k}^{(1)}\}.$$

On the other hand, using the same arguments as Section 1.3.1, we have

$$P^* = \inf_{\delta \in \delta_1} P\{(CS)_1^{(1)}\} = P\{Z_{k-i+1:k}^{(1)} \leq d_{k-i+1:k}^*\}.$$

Hence we have $d_{i:k}^{(1)} = d_{i:k}^*$, $i = 1, \dots, k$.

Since $X_{1:k} = X_{i:k}$, the first step of δ_1 and δ_i^1 are identical. For $i = 2, \dots, k$, the event

$$\begin{aligned}
\text{select } \pi_1, \dots, \pi_{k-1} &= \left\{ \bigcap_{j=1}^i (X_{j:k} = 0) = \bigcap_{j=1}^i (V_j = 0) \right\} \\
&= \left\{ \bigcap_{j=1}^i (X_{j:k} = 0) = \bigcap_{j=1}^i (V_j = 0) \right\} \\
&= \text{select } \pi_1, \dots, \pi_k
\end{aligned}$$

by Lemma 1.3.2 and Lemma 1.3.3. Hence selection procedures π_1 and π_2 are identical.

1.3.6. Some Other Proposed Selection Procedures π_2, π_3, π_4

In Case I, we proposed some other selection procedures:

Definition 1.3.3. We define a selection procedure π_d by

$$\pi_d: \text{Select } \pi_i \text{ if and only if } X_{i:k} = 0 = d_{\min} \quad i = 1, \dots, k$$

where d is the smallest value such that π_d satisfies the P^* -condition.

Theorem 1.3.3. Under assumptions of Case I, and selection procedure π_d , if we select population π_i , then we will select populations π_j for all $j \geq i$.

Proof. Since $X_{i:k} = X_{j:k}$ for all $j \geq i$.

Evaluation the Value d of π_d

For any $i, 1 \leq i \leq k$, we have

$$\begin{aligned}
\inf_{\pi_i} P_i(CS_{\pi_d}) &= \inf_{\pi_i} P_i(V_{k-i+1:k} = 0 = d_{\min}) \\
&= P(V_i = 0)
\end{aligned}$$

by the same argument for selection procedure π_1 and here

$$V_i = Z_{1:i} = \min_{1 \leq j \leq i} \left(\frac{1}{r} \sum_{j=1}^r Z_{ij} \right).$$

We need the constant d such that $P(V_i \geq d) = P^*$ holds for all i , $1 \leq i \leq k$. By Lemma 1.3.2 we have $d = d_{1:k}^{(1)}$. Hence we have the following theorem.

Theorem 1.3.3. Selection procedure π_2 satisfies the P^* -condition with $d = d_{1:k}^{(1)}$.

Corollary 1.3.3. If S_1 and S_2 are the selected subsets associated with selection procedures π_1 and π_2 , respectively, then $S_1 = S_2$.

Proof. Proof follows from Lemma 1.3.2.

Definition 1.3.4. The procedure π_3 is defined as follows:

Step 1. Select π_1 , $i = 1$ and stop, if

$$\bar{X}_1 \geq \mu_0 + d_1 \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_1 and go to step 2.

Step 2. Select π_2 , $i = 2$ and stop, if

$$\bar{X}_2 \geq \mu_0 + d_2 \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_2 and go to step 3.

⋮

Step $k-1$. Select π_i , $i = k-1$ and stop, if

$$\bar{X}_{k-1} \geq \mu_0 + d_{k-1} \frac{\sigma}{\sqrt{n}},$$

otherwise reject μ_{k-1} and go to step k.

Step k. Select μ_k and stop, if

$$\bar{X}_k - 0 \leq d_k \sqrt{\frac{1}{n}},$$

otherwise reject μ_k .

Here $X_j = \max(X_1, \dots, X_j)$ and d_i 's are the smallest values such that μ_i satisfies the P^* -condition.

Evaluation of d_i 's

For any i , $1 \leq i \leq k$,

$$\begin{aligned} \inf_{\mu \in \mu_i} P_{\mu}(CS_3) &= \inf_{\mu \in \mu_i} P\left(\bigcap_{j=1}^{k-i+1} X_j - \mu \leq d_j \sqrt{\frac{1}{n}}\right) \\ &= P(Z_{k-i+1} \leq d_{k-i+1} \sqrt{\frac{1}{n}}) \\ &= P(Z_{k-i+1} \leq d_{k-i+1}) \\ &= P^*, \quad Z_i \sim F(\cdot; 0, 1). \end{aligned}$$

This implies $d_{k-i+1} = d$ for all i , and

$$\begin{aligned} d &= F^{-1}(1-P^*), \\ &= F^{-1}(P^*), \quad \text{if } F \text{ is symmetric} \\ &= F^{-1}(P^*), \quad \text{if } X_i \sim N(\mu_i, \sigma^2/n). \end{aligned}$$

Similar to the selection procedure S_1 , we have the following theorem:

Theorem 1.3.10. Selection procedure S_3 satisfies the P^* -condition with $d_i = F^{-1}(1-P^*)$.

Definition 1.3.5. Selection procedure ϕ_k^* is defined as follows:

Step 1. Select x_1 , $i = 1$ and stop, if

$$x_1 \leq c_1 = d_{\frac{1}{n}},$$

otherwise reject x_1 and go to step 2.

Step 2. Select x_2 , $i = 2$ and stop, if

$$x_2 \leq c_2 = d_{\frac{2}{n}},$$

otherwise reject x_2 and go to step 3.

⋮

Step k-1. Select x_k , $i = k - 1$ and stop, if

$$x_{k-1} \leq c_{k-1} = d_{\frac{k-1}{n}},$$

otherwise reject x_k .

Here

$$d = F^{-1}(1-P^*) \\ = F^{-1}(P^*) \text{ if } F \text{ is symmetric.}$$

Theorem 1.3.11. The selection procedure ϕ_k^* satisfies the P^* condition.

Proof. For any i , $1 \leq i \leq k$,

$$\inf_{\theta \in \theta_1} P_{\theta}(\phi_k^* \leq c_i) = P(\phi_{k-1}^* \leq c_i) = P^*.$$

Theorem 1.3.12. The selection procedure ϕ_k^* and ϕ_k are identical.

Proof. The proof is simple hence it is omitted.

The following procedure γ_4 was given by Gupta and Sobel (1968), without assuming any ordering prior:

Definition 1.3.6. The selection procedure γ_4 is defined as follows:

$$\gamma_4: \text{ Select } \gamma_i \text{ if and only if } X_i \geq 0 + d \frac{1}{\sqrt{n}} \quad i = 1, \dots, k$$

where d is the smallest constant such that γ_4 satisfies the P^* -condition.

It was shown that the value d is determined by the equation

$$F(-d) = 1 - P^{*k}$$

or

$$F(d) = P^{*k} \text{ if } F \text{ is symmetric.}$$

1.3.7. A Dual Problem

We start with the same assumptions as in Section 1.3.1 Case 1, but change our goal to select a subset which contains no bad populations; the definition of a correct selection (CS) will now be changed to select a subset that contains no bad populations and the P^* -condition will be defined based on this new definition of correct selection (CS).

In location parameter case, this problem is a dual problem of the original problem, namely, "select a subset which contains all good populations under ordering prior assumption".

One method to solve this problem is that, first, change the sign of all statistics and the control to opposite sign; then use a procedure for selecting a subset which contains all "new good" populations.

where the "new good" populations are the "old bad" populations, before changing signs; finally, reject the selected subset and keep the remainders as the desired selected subset. Let $\gamma_i, i = 1, 2, 3, 4$ denote the above procedure which corresponds to $\gamma_i, i = 1, 2, 3, 4$, respectively.

Theorem 1.3.13. The selection procedure $\gamma_i, i = 1, 2, 3, 4$ satisfies the P^* -condition in which the correct selection (CS) means that it selects a subset which contains no bad population.

Proof. Given P^* and observations, for any selection procedure $\gamma_i, i = 1, 2, 3, 4$, after changing the signs of all associated statistics, the probability that the selected subset S contains all "new good" populations is not less than P^* . If we reject the selected subset S , then the complement subset S^C of S contains any "new good" populations with probability less than $1-P^*$, but the "new good" populations are the originally bad populations, so what we have is that the subset S^C contains any originally bad population with probability less than $1-P^*$, in other words, subset S^C contains no bad population with probability greater than or equal to P^* . Since this is true for all arbitrary true configurations, we have completed the proof.

Remark 1.3.1. It is easy to see that the value $d_{i+1}^{(i)}$ of γ_i which was used by γ_i in the i th step is determined by the equation

$$P(\bigcup_{k=i+1}^K d_{i+1}^{(i)}) = P^*,$$

where

$$\bigcup_{k=i+1}^K = \bigcup_{k=i+1}^K \frac{1}{1 + (k-i+1)^{1/(1-P^*)}}$$

If the distribution F is symmetric, then

$$d_{i:k}(\cdot, 1) = d_{i:k}^{(1)}.$$

1.3.8. Some Proposed Selection Procedures $\varphi_i^{(2)}$, $i = 1, 2, 3, 4$

When μ_0 is Unknown

Case II. μ_0 unknown, common σ^2 known, common sample size n .

Definition 1.3.2. We define a selection procedure $\varphi_i^{(2)}$ by replacing the inequalities

$$X_{i:k} - \mu_0 \geq d_{i:k}^{(1)} \frac{\sigma}{\sqrt{n}}, \quad i = 1, \dots, k$$

in procedure φ_1 (Definition 1.3.1) with

$$X_{i:k} - X_0 \geq d_{i:k}^{(2)} \frac{\sigma}{\sqrt{n}}, \quad i = 1, \dots, k, \text{ respectively.}$$

Here $c_0 = \sum_{i=1}^k X_{0i}/n$, $d_{i:k}^{(2)}$, $i = 1, \dots, k$ are the smallest constants such

that the selection procedure $\varphi_i^{(2)}$ satisfies the P^* -condition.

Similar to the Case I, we have the following theorem:

Theorem 1.3.14. For any i , $i = 1, \dots, k$, $\hat{c}_i = \frac{1}{k-i+1}$ is determined by the equation

$$\int_0^{\hat{c}_i} t(V_1 - t) \geq d_{k-i+1:k}^{(2)} \sigma H(t) = P^*, \quad (1.3.14)$$

It is easy to see that $d_{k-i+1:k}^{(2)} = d_{1:i}^{(2)}$. The following theorem gives us an identical form of the selection procedure $\phi_1^{(2)}$.

Theorem 1.3.15. The selection procedure $\phi_1^{(2)}$ will not be changed if the statistics $\bar{X}_{i:k}$, $i = 1, \dots, k$, are replaced by $X_{i:k}$, $i = 1, \dots, k$, respectively.

Proof. The proof is the same as that in Case I and hence it is omitted.

The values $d_{1:i}^{(2)}$, $i = 1, \dots, k$ are tabulated in Table II for $k = 1(1)6, 8, 10$, and $P^* = .99, .975, .95, .925, .90, .85, .80, .75, .70, .65$.

Similar to the Case I, we propose a selection procedure $\phi_2^{(2)}$ as follows:

Definition 1.3.8. We define a selection procedure $\phi_2^{(2)}$ by

$\phi_2^{(2)}$: Select π_i if and only if $X_{i:k} - \bar{X}_0 \geq d_{i:n}$, $i = 1, \dots, k$

where d is the smallest value such that $\phi_2^{(2)}$ satisfies the P^* -condition. Then, similar to Theorem 1.2.9 we have:

Theorem 1.3.16. Under assumptions of Case II, the selection procedure $\phi_2^{(2)}$ satisfies the P^* -condition with $d = d_{1:k}^{(2)}$.

Next, we define a selection procedure $\phi_3^{(2)}$ which is identical to $\phi_2^{(2)}$ but replace $\bar{X}_0$ by $\bar{X}_0$, the sample mean of population π_0 .

Definition 1.3.2. The selection procedure $\phi_3^{(1)}$ is defined by rejecting

$X_i = \bar{x}_0 + d_i \sqrt{\frac{1}{n_i}}$ in π_3 (Definition 1.3.4) by $X_i = \bar{x}_0 + d_i' \sqrt{\frac{1}{n_i}}$, $i = 1, \dots, k$, where $d_1', \dots, d_k'$ are the smallest values such that $\phi_3^{(1)}$ satisfies the P^* -condition.

Similar to Theorem 1.3.10 we have:

Theorem 1.3.12. The selection procedure $\phi_3^{(2)}$ satisfies the P^* -condition with $d_i = d$, $i = 1, \dots, k$ where d is determined by the equation

$$\int_{-\infty}^{\infty} [1 - F(t-d)] dF(t) = P^*, \quad (1.3.12)$$

$$\int_{-\infty}^{\infty} F(d-t) dF(t) = P^*, \text{ if } F \text{ is symmetric.}$$

And $\phi_3^{(2)}$ will not be changed if the statistics X_i is replaced by $\bar{x}_i$, the sample mean of population π_i for $i = 1, \dots, k$.

The following selection procedure $\phi_4^{(2)}$ was proposed by Gupta and Patel (1958):

Definition 1.3.10. The selection procedure $\phi_4^{(2)}$ is defined by

$$\phi_4^{(2)}: \text{ select } \pi_i \text{ if and only if } X_i = \bar{x}_0 + d \sqrt{\frac{1}{n_i}}, \quad i = 1, \dots, k$$

where d is determined by the following equation if F is normal distribution:

$$\int_{-\infty}^{\infty} \sum_{i=1}^k [F(t \sqrt{\frac{1}{n_i}} + d)]^k(u) du = P^*, \quad (1.3.13)$$

For the special case $n_i = n$ ($i = 0, 1, \dots, k$)

$$\int_{-\infty}^{\infty} F^k(t+d)f(t)dt = P^*, \quad (1.3.24)$$

If F is normal distribution $N(0,1)$, the tables of d values satisfying the Equation (1.3.25) for several values of P^* are given in Bechhofer (1954) for $k = 1, 2, 3$ and in Gupta (1986) for $k = 4, 5$.

1.3.9. Some Proposed Selection Procedures $\frac{1}{4}^{(2)}$, $i = 1, 2, 3, 4$

When Common Variance σ^2 is known

Case III: μ_0 known, common variance σ^2 unknown, $n_i = n$, $i = 1, 2, 3$

In this case, we assume that $f(x) = \phi(x)$ which is the density of $N(0,1)$.

Definition 1.3.11. We define the selection procedure $\frac{1}{4}^{(1)}$ to reject H_0 if the inequality

$$t_{ijk} - \mu_0 \leq d_{ijk}^{(1)}, \quad i = 1, 2, 3$$

in procedure $\frac{1}{4}$ (Definition 1.3.1) is

$$t_{ijk} - \mu_0 \leq d_{ijk}^{(3)}, \quad i = 1, 2, 3, \text{ respectively,}$$

where $d^{(3)}$'s are the smallest values such that $\frac{1}{4}^{(3)}$ satisfies the P^* -condition, $\hat{\sigma}^2$ denotes the pooled estimator of σ^2 based on

$k(n-1)$, that is

$$\hat{\sigma}^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^{n-1} (x_{ij} - \bar{x}_i)^2$$

Note that $\frac{S}{\sigma^2}$ has the chi-square distribution χ^2_k with k degrees of freedom.

By using similar arguments as in Case I, we have:

Theorem 1.3.18. The equation which determines the constant $d_{k-i+1:k}^{(S)}$ is

$$P(V_i \leq d_{k-i+1:k}^{(S)}) = P^* \quad (1.3.17)$$

or

$$\int_0^{d_{k-i+1:k}^{(S)}} P(V_i \leq y) q(y) dy = P^* \quad (1.3.18)$$

where

$$V_i = \min_{1 \leq j \leq i} \frac{1}{r-1} \sum_{j=1}^r Z_j$$

and $q(y)$ is the density of $\frac{S}{\sigma^2} = \chi^2_k$.

We can rewrite formula (1.3.18) as

$$\int_0^{d_{k-i+1:k}^{(S)}} P(V_i \leq y) \sqrt{\frac{2}{\pi}} e^{-y/2} y^{k/2-1} dy = P^*$$

or

$$\int_0^{d_{k-i+1:k}^{(S)}} P(V_i \leq y) \sqrt{\frac{2}{\pi}} e^{-y/2} y^{k/2-1} dy = P^* \quad (1.3.19)$$

Remark 1.3.4. The values of $d_{k-i+1:k}^{(3)}$, $i = 1, \dots, k$ depend on $\alpha \in (0, 1)$, hence $d_{k-i+1:k}^{(3)} \neq d_{1:i}^{(3)}$.

By using Rabinowitz and Weiss table (1959) (with $n = 21$, $4 \leq k \leq 6$), we have evaluated and tabulated the values of $d_{k-i+1:i}^{(3)}$, $i = 1, \dots, k$, in Table III, for $k = 2$ (1) 6, $P^* = .99, .975, .95, .925, .90, .85, .80$, and .75 with common sample size $n = 3, 5, 9$, and 21.

for $k = 6$ and $n = 21$, we can use $d_{1:1}^{(3)}$ as an approximation of $d_{k-i+1:k}^{(3)}$.

Definition 1.3.17. We define the selection procedure $\phi_{\frac{1}{2}}^{(3)}$ by

$$\phi_{\frac{1}{2}}^{(3)}: \text{Select } j \text{ if and only if } \bar{X}_{j:k} \geq d_{1:1}^{(3)} + d_{\frac{1}{2}}^{(3)} \quad j = 1, \dots, k,$$

where S is defined as in procedure $\phi_{\frac{1}{2}}^{(3)}$, and $d_{\frac{1}{2}}^{(3)}$ is the smallest constant such that $\phi_{\frac{1}{2}}^{(3)}$ satisfies the P^* -condition.

As before, it can be shown that $\phi_{\frac{1}{2}}^{(3)} = \phi_{1:k}^{(3)}$.

Remark 1.3.5. Theorem 1.1.2 still holds for Case III, i.e. the selection procedure $\phi_{\frac{1}{2}}^{(3)}$ will not be changed if we replace the test statistics $\bar{X}_{j:k}$ by $\bar{X}_{j:1}$, respectively. But this is not necessarily true for selection procedure $\phi_{\frac{1}{2}}^{(2)}$.

Definition 1.3.18. The selection procedure $\phi_{\frac{1}{2}}^{(2)}$ is defined to have the same form as procedure $\phi_{\frac{1}{2}}^{(3)}$ except that the inequality defining the i th step of procedure $\phi_{\frac{1}{2}}^{(2)}$ is replaced by

$$\bar{X}_{j:1} \geq d_{1:1}^{(2)} + d_{\frac{1}{2}}^{(2)} \quad j = 1, \dots, k,$$

The proof of the following theorem uses the same arguments as that in Case I, hence it is omitted.

Theorem 1.3.19. The equation which determines the constant d of selection procedure $\delta_3^{(3)}$ is

$$\int_0^{\infty} (yd) q_1(y) dy = P^*, \quad (1.3.11)$$

Gupta and Sobel (1958) gave a selection procedure $\delta_4^{(4)}$ in this case. It is as follows:

$\delta_4^{(4)}$: Select μ_i if and only if $X_i \geq \mu_0 + d \sum_{j=1}^k \frac{1}{n_j}$, $i = 1, \dots, k$

and the equation which determines d is

$$\int_0^{\infty} (yd)^k q_1(y) dy = P^*. \quad (1.3.12)$$

1.3.10. Some Proposed Selection Procedures $\delta_i^{(4)}$, $i = 1, 2, 3, 4$

When Both Control μ_0 and Common Variance σ^2 are Unknown.

Case IV. μ_0 unknown, common variance σ^2 unknown and common sample size n .

We assume that in this case distribution F is the c.d.f. $N(0,1)$, and denoted by ϕ . We replace μ_0 in each selection procedure $\delta_j^{(3)}$ by $\bar{x}_0$, $j = 1, 2, 3$, and get four procedures $\delta_i^{(4)}$, $i = 1, 2, 3, 4$, respectively. Let $\chi^2_k(t)$ denote the c.d.f. of the chi-square distribution with

$k(n-1)$ degrees of freedom. The constant $d_{k-i+1;k}^{(4)}$, $i = 1, \dots, k$, of procedure $\delta_1^{(4)}$ is determined by

$$\int_0^{\infty} \int_0^{\infty} P(V_1 = u + d_{k+1+1:k}^{(4)} \sqrt{1/d_{k+1+1:k}^{(4)}} (t)) d_{k+1+1:k}^{(4)}(t) dt = 1, \quad (1.3.10)$$

The constant d of procedure $\delta_{\frac{1}{2}}^{(4)}$ is

$$d = d_{1:k}^{(4)}.$$

The constants d of procedures $\delta_{\frac{1}{4}}^{(4)}$ and $\delta_{\frac{3}{4}}^{(4)}$ are determined by

$$\int_0^{\infty} \int_0^{\infty} r(u + td) d(u) d_{1:k}^{(4)}(t) dt = P^*, \quad (1.3.11)$$

with $r = 1$ and k , respectively, and their values for selected values of P^* , k and α are given in Gupta and Sobel (1957) and Duggan (1967).

1.3.11. Properties of the Selection Procedures

Under simple ordering prior, it is natural to require that an order-preserving selection procedure is order-preserving as defined below:

Definition 1.3.14. A selection procedure δ is order-preserving, if it selects τ_i with parameter α_i , and if $\alpha_i < \alpha_j$, then it also selects τ_j . Procedure δ is weak order-preserving or monotone if

$$P(\tau_i \text{ is selected} | \cdot) \leq P(\tau_j \text{ is selected} | \cdot) \text{ whenever } \alpha_i < \alpha_j.$$

It is easy to see that any order-preserving selection procedure is weak order-preserving, but the converse is not true.

Now, let $\delta_{\frac{1}{4}}^{(1)} = \delta_{\frac{1}{4}}^{(1)}$, $i = 1, 2, 3, 4$.

Theorem 1.3.20. The selection procedures $\delta_{\frac{1}{4}}^{(1)}$, $\delta_{\frac{1}{2}}^{(1)}$ and $\delta_{\frac{3}{4}}^{(1)}$ are order-preserving and procedure $\delta_{\frac{1}{4}}^{(1)}$ is monotone, for $r = 1, 2, 3, 4$.

Proof. The proof follows immediately from the definitions of the procedures.

Given observations $X = x = (x_0, \dots, x_k)$ where x_i is the sample mean of population π_i , $i = 1, \dots, k$, and $x_0 = \bar{x}_0$ if π_0 is known, otherwise x_0 is the sample mean of population π_0 . Let

$$\pi_i(x, \pi) = P(\pi_i \text{ included in the selected subset} \mid x, \pi) \quad \text{for } i = 1, \dots, k.$$

Definition 1.3.16. A selection procedure ϕ is called translation-invariant if for any $x \in \mathbb{R}^{k+1}$, $c \in \mathbb{R}$

$$\pi_i(x_0 + c, x_1 + c, \dots, x_k + c; \pi) = \pi_i(x_0, \dots, x_k; \pi) \quad i = 1, \dots, k.$$

Theorem 1.3.21. The selection procedures $\phi_1^{(i)}$, $\phi_2^{(i)}$, $\phi_3^{(i)}$ and $\phi_4^{(i)}$ are translation-invariant for $i = 1, 2, 3, 4$.

Proof. By Corollary 1.3.1 the isotonic regression is a linear operation; on the other hand,

$$\frac{\sum_{i=1}^n (x_{i1} + c)}{n} = \frac{\sum_{i=1}^n x_{i1}}{n} + c,$$

hence we have the result.

Expected Number (Size of Bad) Populations in the Selected Subset

Suppose the control π_0 is known and we have common known μ_0 and common known variance σ^2 ; without loss of generality we assume that $\mu_0 = 0$ and $\sigma^2 = 1$. Let $E(\pi^*)$ denote the expected number of bad populations in the selected subset according to the selection procedure

then for any $j, 0 \leq j \leq k$,

$$\begin{aligned} \sup_{u \in \mathcal{C}_{k-j}} E(S^{(1)}_{k-j}) &= \sup_{u \in \mathcal{C}_{k-j}} \sum_{r=1}^j P\left(\bigcup_{i=1}^r X_{i:k} = d_{1:k}^{(1)}\right) \\ &= \sum_{r=1}^j P\left(\bigcup_{i=1}^r Z_{i:j} = d_{1:j}^{(1)}\right) \quad (1.3.35) \end{aligned}$$

On the other hand, for procedure $\mathcal{P}_2$

$$\sup_{u \in \mathcal{C}_{k-j}} E(S^{(1)}_{k-j}) = \sum_{r=1}^j P\left(\bigcup_{i=1}^r Z_{i:j} = d_{1:j}^{(1)}\right) \quad (1.3.36)$$

Formula (1.3.36) is increasing in j and is greater than or equal to

Formula (1.3.35), since

$$d_{1:k}^{(1)} = d_{1:k-1+1}^{(1)} = d_{1:k}^{(1)}.$$

Therefore, we have the following theorem.

Theorem 1.3.22. For any $j, 0 \leq j \leq k$

$$\sup_{u \in \mathcal{C}_j} E(S^{(1)}_{\mathcal{P}_2}) = \sup_{u \in \mathcal{C}_j} E(S^{(1)}_{\mathcal{P}_1}),$$

$$\sup_{u \in \mathcal{C}} E(S^{(1)}_{\mathcal{P}_2}) = \sup_{u \in \mathcal{C}} E(S^{(1)}_{\mathcal{P}_1}).$$

Theorem 1.3.23. In Section 1.3.1, Case 1, for any $j, 0 \leq j \leq k$

$$\sup_{u \in \mathcal{C}_{k-j}} E(S^{(1)}_{k-j}) = (1 - q(1 - q^j)/2)P^* \quad (1.3.37)$$

where $q = 1 - P^*$.

Proof.

$$\begin{aligned}
 & \sup_{\mathcal{C}_{k-j}} E(S^{(j)}_3) \\
 &= \sup_{\mathcal{C}_{k-j}} \sum_{i=1}^j P(\text{select } x^{(j)}_i) \\
 &= \sup_{\mathcal{C}_{k-j}} \sum_{i=1}^j P(\max_{r=1}^i x^{(j)}_r \geq d) \\
 &= \sum_{i=1}^j (1 - \prod_{r=1}^i F(-d)) \\
 &= j - \sum_{i=1}^j q^i \\
 &= j - q(1-q^j)/P^*
 \end{aligned}$$

where $q = (1-P^*)$.

Theorem 1.3.24. $\sup_{\mathcal{C}_{k-j}} E(S^{(j)}_3)$ is increasing in j , hence

$$\sup_{\mathcal{C}_{k-j}} E(S^{(j)}_3) = \sup_{j=0} E(S^{(j)}_3) = k - q(1-q^k)/P^*. \quad (1.3.25)$$

Proof. Since the function

$$f(x) = x - ab^{x+1}$$

is increasing in x , for $0 < a < 1$, $0 < b < 1$, and $0 \leq x < \infty$.

In Case I of Gupta (1965) showed that

$$\sup_{\mathcal{C}} E(S^{(1)}_4) = kP^* \quad (1.3.26)$$

When the ordering prior among the unknown parameters is unknown, we can use the selection procedure of Gupta and label (1961) as follows: the ordering of the sample means as the ordering of unknown parameters and apply the selection procedure which is optimally invariant to the unknown prior. With the latter approach, the substitution of the isotonic regression of the sample means, taken to the level of the sample means, and that the selection procedures $\phi_j^{(1)}$, $j = 1, 2, 3, 4$ are the same as $\phi_j^{(i)}$, $(i = 1, 2, 3, 4)$, respectively, and the selection procedures $\phi_j^{(i)}$, $i = 1, 2, 3, 4$ become $\phi_j^{(1)}$, $j = 1, 2, 3, 4$ respectively, which are equivalent to the procedures proposed by Gupta (1975) and Broström (1977), independently.

It has been proved in some quite general situation and verified by using Monte Carlo technique in some selected cases by Gupta (1975) and Broström (1977), separately, that $\phi_j^{(1)}$ ($j = 1, 2, 3, 4$) is slightly better than ϕ_4 . The values $d_{i;k}^{(j)}$ in the i th step of the procedure $\phi_j^{(1)}$, $j = 1, 2, 3, 4$, are given by Broström (1977) as follows:

$$d_{1;k}^{(1)} = \tau^{-1} [1 - (p^*)^k], \quad (1.10)$$

$$\int_{-\infty}^{\infty} \tau(x) + d_{1;k}^{(2)} \tau(x) dx = p^*, \quad (1.11)$$

$$\int_0^{\infty} \tau(x) d_{1;k}^{(3)} dx = p^*, \quad (1.12)$$

and

$$\int_0^{\infty} \int_{-\infty}^{\infty} \tau(x) + d_{1;k}^{(4)} \tau(x) q(y) dy dx = p^* \quad (1.13)$$

where $q(\cdot)$ is the density of $\tau(\cdot)$.

1.1. Selection Procedures for the Parameters of Gamma Populations

Suppose we have $k+1$ independent populations $\pi_1, \dots, \pi_{k+1}$. The distribution π_i has a gamma density function

$$g(x; \alpha_i, \beta_i) = \frac{1}{\Gamma(\alpha_i)} \frac{\beta_i^{\alpha_i}}{x^{\alpha_i+1}} e^{-\beta_i x} \quad (1.1)$$

$$x \in (0, \infty), \quad i = 1, \dots, k+1.$$

$i = 1, \dots, k$. The ordering prior of $\pi_1, \dots, \pi_k$ and π_{k+1} is not known. For example, say, $0 < \beta_1 < \beta_2 < \dots < \beta_k < \beta_{k+1}$. Notice that the values of $\alpha_1, \dots, \alpha_k$ and $\beta_1, \dots, \beta_k$ are unknown, α_{k+1} 's are known.

In this section we define population π_i , $i = 1, \dots, k$, as the i th population if the scale parameter $\beta_i = \beta$. Let $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k$ be known, then the parameter space is denoted by Ω , where

$$\Omega = \{ \theta \in \mathbb{R}^{k+1} : \theta = (\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k) \}$$

is a subspace of $(k+1)$ -dimension Euclidean space $\mathbb{R}^{k+1}$.

Suppose we have independent observations X_{ij} , $(i = 1, \dots, n_i)$ from population π_i , $(i = 1, \dots, k)$. Let $n_i = n_i \alpha_i$, then

$$X_i = \sum_{j=1}^{n_i} X_{ij} / n_i \text{ has density } g(x; \alpha_i, \beta_i / n_i),$$

and

$$X_i / \alpha_i \text{ has density } g(x; \alpha_i, \beta_i).$$

Suppose our goal is to select a subset which contains all good populations under the ordering prior with probability greater than or equal to P^* , a preselected value between zero and one.

Let Ω_i , $i = 1, \dots, k$, be the subspace of parameter space Ω such that $\bigcap_{i=1}^k \Omega_i = \emptyset$ where

$$\begin{aligned}
 & \hat{\mu}_i = \hat{\mu}_0, \hat{\mu}_i = \hat{\mu}_0 + \hat{\mu}_{i+1}, \text{ if } i = 1, \dots, \text{or } t-1, \\
 & \hat{\mu}_i = \hat{\mu}_k, \text{ if } i = k, \\
 & \hat{\mu}_i = \hat{\mu}_0 + \hat{\mu}_1, \text{ if } i = 0.
 \end{aligned}$$

1.4.1. Proposed Selection Procedures $\pi_i, i = 6, 7, 8, 9$

Case I. Control π_0 known and common sample size n .

Definition 1.4.1. The selection procedure π_6 is defined as follow

Step 1. Select $\pi_i, i = k$ and stop, if

$$\bar{X}_{k:k} \geq c_{k:k}(0)$$

otherwise reject π_k and go to step 2.

Step 2. Select $\pi_i, i = k-1$ and stop, if

$$\bar{X}_{k-1:k} \geq c_{k-1:k}(0)$$

otherwise reject π_{k-1} and go to step 3.

⋮

Step $k-1$. Select $\pi_i, i = 2$ and stop, if

$$\bar{X}_{2:k} \geq c_{2:k}(0)$$

otherwise reject π_2 and go to step k .

Step k . Select π_1 and stop, if

$$\bar{X}_{1:k} \geq c_{1:k}(0)$$

otherwise reject π_1 .

where $c_{i;k} \in (-1)$, $i = 1, \dots, k$ are the smallest values such that the procedure ϕ_6 satisfies the P^* -condition.

Theorem 1.4.1. Assume we have common sample size n and $s_j = n$, $j = 1, \dots, k$, and the constant $c_{i;k} \in (-1)$ is determined by the condition

$$P(U_i \leq c_{i;k}) = P^*, \quad i = 1, \dots, k, \quad (1.5)$$

where

$$U_i = \max_{j \in S, i} \left(\sum_{j=s+1}^n Y_j + \dots + Y_i \right) \quad (1.6)$$

and Y_j are i.i.d. with density

$$g(x) = \frac{1}{n}, \text{ and c.d.f. } G(x) = \frac{x}{n}$$

then the procedure ϕ_6 satisfies the P^* -condition.

Proof. For any i , $1 \leq i \leq k$, if the unknown true $c_{i;k}$ is not $\leq c_{i;k}$, then there are i good populations, then, under the procedure ϕ_6

$$\begin{aligned} \inf_{c_{i;k}} P(\text{FC} | \phi_6) &= \\ &= \inf_{c_{i;k}} P \left(\bigcup_{j=1}^k \left(\bigcap_{i=1}^j Y_{i;k} \leq c_{i;k} \right) \cap \left(\bigcap_{i=j+1}^k Y_{i;k} > c_{i;k} \right) \right) \\ &= \inf_{c_{i;k}} P \left(\bigcup_{j=1}^k \left(\bigcap_{i=1}^j \left(\sum_{t=s+1}^n Y_t + \dots + Y_i \right) \leq c_{i;k} \right) \cap \left(\bigcap_{i=j+1}^k \left(\sum_{t=s+1}^n Y_t + \dots + Y_i \right) > c_{i;k} \right) \right) \\ &= \inf_{c_{i;k}} P \left(\bigcup_{j=1}^k \left(\bigcap_{i=1}^j \left(\sum_{t=s+1}^n Y_t + \dots + Y_i \right) \leq c_{i;k} \right) \cap \left(\bigcap_{i=j+1}^k \left(\sum_{t=s+1}^n Y_t + \dots + Y_i \right) > c_{i;k} \right) \right) \end{aligned}$$

$$= \inf_{c_i} P_{n,\star} \left(\bigcup_{j=i}^k \left(\max_{1 \leq s \leq j} \min_{j \leq t \leq k} \frac{Y_{s-1}^{(0)} + \dots + Y_{t-1}^{(0)}}{t-s+1} \leq c_{i:k} \right) \right)$$

$$= P_{n,\star} \left(\bigcup_{j=i}^k \left(\max_{1 \leq s \leq j} \min_{j \leq t \leq i} \frac{Y_{s-1}^{(0)} + \dots + Y_{t-1}^{(0)}}{t-s+1} \leq c_{i:k} \right) \right)$$

$$= P \left(\max_{1 \leq s \leq i} \frac{Y_{s-1} + \dots + Y_i}{i-s+1} \leq c_{i:k} \right)$$

$$= P(U_i \leq c_{i:k}),$$

where Y_i 's are i.i.d with density $q(\cdot; v, \frac{1}{n})$,

$$\star = (\underbrace{0, 0, \dots, 0}_{i+1}, \dots, 0)$$

and

$$U_i = \max_{1 \leq s \leq i} \frac{Y_{s-1} + \dots + Y_i}{i-s+1}.$$

Corollary 1.4.1. $c_{i:k} = c_{i,i}$, $i = 1, \dots, k$.

For any $x \geq 0$, let $S_n = \sum_{i=1}^n (Y_i - x)$, $n = 1, 2, \dots$, $x_0 = 0$. Since

$Q = P(Y_1 - x \geq 0) = Q(x, x, \frac{1}{n}) \rightarrow 1$, the distribution of $Y_1 - x$ is not concentrated on a half-axis. By Theorem 1.3.1 the probability percentiles function of cumulative distribution functions $P(U_k \leq c)$, $c \geq 0$, $k = 1, 2, \dots$, is given by

$$\exp \left(- \sum_{k=1}^{\infty} \frac{1}{k} \int_0^k Q(z/n, k, 1) dz \right).$$

Hence by Theorem 1.3.6, we have the following recurrence formula for all $x \geq 0$

$$P(U_{k+1} \leq x) = \frac{1}{k+1} \sum_{j=0}^k P(U_{k-j} \leq x) + G(x(j+1)n; (j+1)n, 1) \quad (1.4.4)$$

where

$$P(U_0 \leq x) = 1.$$

When $x = 0$, both sides of Equation (1.4.4) equal to zero, hence it also holds for $x = 0$.

Note that

$$P\left(\frac{1}{r} \sum_{i=1}^r Y_i \leq x\right) = G(x; rn, \frac{1}{rn}) \\ = G(xrn; r, 1). \quad (1.4.5)$$

The values $c_{k;k}(P^*, r, n)$ which satisfy Equation (1.4.1) are tabulated in Table IV for $k = 1(1)10$, $P^* = .99, .95, .90, .75$, $r = 2, 4, 6$, and $n = 4, 6, 8, 10, 15, 20$.

Lemma 1.4.1. $c_{i;k} = c_{i+1;k}$ for all $1 \leq i \leq k-1$.

Proof. The constants $c_{i;k}$ ($i = 1, \dots, k$) are determined by (1.4.1), respectively,

$$U_{i+1} = U_i + 1, \text{ implies } c_{i;k} = c_{i+1;k} \text{ for all } 1 \leq i \leq k-1.$$

Theorem 1.4.2. The selection procedure ϕ_k will not be changed if the monotonic estimators Y_{i+1} , $i = 1, \dots, k$ are replaced by Y_{i+1} , $i = 1, \dots, k$.

where

$$\begin{aligned} X_{i+1} &= \max_{1 \leq j \leq i} X_{ij} \\ &= \max_{1 \leq s \leq i-s+1} \{X_s + (i-s) \cdot \lambda_i\} \end{aligned}$$

Proof. The proof is similar to that of Lemma 1.4.1.

Next, we define a selection procedure by using an efficient estimator and a fixed constant which depends on P^* , k , sample size n and common $\lambda_1 = \lambda$.

Definition 1.4.2. The selection procedure γ is defined as

$$\gamma: \text{Select } \gamma_i \text{ if and only if } X_{i+1} \leq c_{k,k} \cdot (i+1) \cdot \lambda_i,$$

where $c_{k,k} \in (0,1)$ is the smallest value such that procedure γ satisfies the P^* -condition.

Corollary 1.4.2. The constant $c_{k,k} \in (0,1)$ of the selection procedure γ equals to $c_{k,k}$ which is determined by equation (1.4.1).

Proof. Follows immediately from Theorem 1.4.1 and Lemma 1.4.2.

Definition 1.4.3. The selection procedure γ is defined as follows

Step 1. Select γ_i if and only if $X_{i+1} \leq$

$$\begin{aligned} & \frac{1}{k} \cdot \sum_{j=1}^k X_{ij} \\ & \leq c_{k,k} \cdot (i+1) \cdot \lambda_i \end{aligned}$$

otherwise reject γ_i and go to Step 2.

Step 2. Select x_j , $j = k+1$, and stop, if

$$\frac{x_{k+1}}{x_{k-1}} - c_{k+1} = 0,$$

otherwise reject x_{k+1} and go to step 3.

3.

Step $k+1$. Select x_i , $i = 2$, and stop, if

$$\frac{x_i}{x_2} - c_2 = 0,$$

otherwise reject x_i and go to step k .

Step k . Select x_1 and stop, if

$$\frac{x_1}{x_1} - 1 = 0,$$

otherwise reject x_1 .

Since the c_j 's are the smallest real values $(j = 1)$ such that the process x_j satisfies the P^* -condition,

where $i = 1, \dots, k$, the constants c_j are determined by

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{1}{x_j} = \frac{1}{c_j}, \quad j = 1, \dots, k, \quad (1.10)$$

where x_j is arbitrary $(j = 1, \dots, k)$, if $x_j \neq 0$.

Now, let

$$x_j =$$

$$\begin{aligned}
&= \inf_{c_1, \dots, c_k} P\left(\bigcup_{i=1}^k \left(\frac{Y_i}{c_i} \leq c_{i+1}^{-1}\right)\right) \\
&= \inf_{c_1} P\left(\bigcup_{i=1}^k \left(\frac{Y_i}{c_i} \leq c_{i+1}^{-1}\right)\right) \\
&= P\left(\bigcup_{j=1}^k \left(\frac{Y_j}{c_j} \leq c_{j+1}^{-1}\right)\right) \\
&= P(Z_1 \leq c_1^{-1})
\end{aligned}$$

where $Z_i = \frac{Y_i}{c_i}$, $i = 1, \dots, k$ are i.i.d. with the gamma density.

$g(\cdot; c_j, 1), c_j^* = (\underbrace{c_0, \dots, c_0}_{i+1}, \dots, c_k)$. Hence $c_j, j = 1, \dots, k$ are

determined by (1.4.6). If $c_1 = \dots = c_k$, then $c_1 = \dots = c_k$.

The following selection procedure ϕ_0 was given by Gupta and Patel (1958).

Definition 1.4.4. The selection procedure ϕ_0 is defined by

ϕ_0 : Select π_i if and only if $\frac{Y_i}{c_i} \leq c^*$, $i = 1, \dots, k$, where c^* is determined by

$$\frac{k}{i+1} \int_0^{c^*} \frac{1}{u} \left(\int_0^u t^i e^{-t} dt \right) e^{-u} du = P^*, \quad (1.4.7)$$

for $c_i = c, i = 1, \dots, k$, it turns out

$$\frac{k}{i+1} \int_0^{c^*} \frac{1}{u} \left(\int_0^u t^i e^{-t} dt \right) e^{-u} du = (P^*)^{\frac{1}{k}} \quad (1.4.8)$$

The left hand side is the c.d.f. of $\frac{1}{\sigma_0^2} \chi_{n_0}^2$ with n_0 degrees of freedom, hence the value c^* can be easily solved with the help of a table of chi-square distribution.

Application to the Selection of Variance of Normal Population

Let $\sigma_i^2 = \sigma_{i1}^2, i = 0, 1, \dots, k$ are the scale parameters for the $k+1$ normal populations and $y_{ij} (j = 1, \dots, n; i = 1, \dots, k)$ are the n observations on the population σ_i with the mean μ_i (known). We assume that the order $\sigma_1^2 \leq \dots \leq \sigma_k^2$ is known.

In the application of selection procedure γ_6 or γ_7 , what we need to do is to evaluate the isotonic regression of S_i^2 which is the sample variance of population $\sigma_i, i = 1, \dots, k$ and denote it by $S_{i1}^2, i = 1, \dots, k$, then directly apply γ_6 or γ_7 . The constant we need is determined by equations (1.4.2) and (1.4.4) where we replace σ_0 by σ_i , the reason being that ns_i^2/σ_i^2 has χ_n^2 distribution with n degrees of freedom and $t_i = \bar{y}_i/\sigma_i$ has the c.d.f.

$$G(2nt; n, 2) = G(t; n, \frac{1}{n}),$$

hence

$$P(\frac{1}{n} \sum_{i=1}^r Y_i \leq t) = G(trn; rn, 1).$$

The application of γ_7 is similar to that of γ_6 (see Gupta et al. (1962)). What we need to do is to replace σ_0 in γ_7 and σ_i by σ_{i1} and replace n in equation (1.4.6) and (1.4.8) by $n, i = 1, \dots, k$.

Remark 1.4.1. τ_0 (Gupta and Sobel (1958)) does not depend on the unknown prior and the sample sizes for each population need not be equal.

If the means μ_i , $i = 1, \dots, k$ are unknown and common sample size $n > 1$, let $S_i^2 = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 / n-1$ and use $n-1$ in place of n in (1.4.4), (1.4.6) and (1.4.8) which determined the constants $c_{i:k}^{(1)}$, $c_{i:k}^{(2)}$, c and c' for τ_6 , τ_8 and τ_9 , respectively.

1.4.2. Selection Procedure $\tau_i^{(2)}$, $i = 6, 7, 8, 9$

Case II. τ_0 unknown.

The assumptions are the same as in Case I except that n_0 observations, viz., $x_{01}, \dots, x_{0n_0}$ are taken on τ_0 .

For selection procedure $\tau_i^{(2)}$, the inequalities defining the procedure and corresponding to $\bar{X}_{i:k} \geq c_{i:k}^{(2)} \tau_0$ ($i = 1, \dots, k$) is replaced by $\bar{X}_{i:k} \geq c_{i:k}^{(2)} \bar{X}_0$ ($i = 1, \dots, k$), respectively. The equation determining $c_{i:k}^{(2)}$ is obtained as before and is given by

$$\int_0^{\infty} P(U_1 \geq c_{i:k}^{(2)} x) f(x) dx = P^* \quad (1.4.9)$$

where $P(U_1 \geq x)$ is the same as that in theorem 1.4.1, and $f(x)$ is the p.d.f. of $\bar{X}_0$ of population τ_0 .

If population τ_0 has gamma distribution with density $g(x)$, $x > 0$ (τ_0 known and τ_0 unknown), then $g(x) = \frac{1}{\Gamma(n_0)} x^{n_0-1} e^{-x}$.

For selection procedure $\tau_i^{(1)}$, the inequalities defining the procedure is

$$\bar{x}_{k+1} = c^* \bar{x}_0$$

and it can be shown $c^* = \frac{\binom{r}{k}}{\binom{r}{k+1}}$.

For selection procedure $\frac{\binom{r}{k}}{\binom{r}{k+1}}$, the inequality defining the procedure and corresponding the $(k+1)$ th step is

$$\frac{\bar{x}_i}{\bar{x}_0} \leq c \frac{\bar{x}_0}{\bar{x}_0} \text{ where } c = \frac{\binom{r}{k}}{\binom{r}{k+1}} \quad (1.1.1)$$

The equation determining c of $\frac{\binom{r}{k}}{\binom{r}{k+1}}$ is given by

$$\int_0^1 \int_0^1 \frac{c^{i+j} u^2 - 1}{\binom{r}{i} \binom{r}{j}} e^{-u} du \frac{t^2 - 1}{\binom{r}{i} \binom{r}{j}} e^{-t} dt = P^* \quad (1.1.2)$$

For selection procedure $\frac{\binom{r}{k}}{\binom{r}{k+1}}$, the inequality defining the procedure is

$$\frac{\bar{x}_i}{\bar{x}_0} \leq c \frac{\bar{x}_0}{\bar{x}_0} \quad (1.1.3)$$

and the equation determining c is given by Gupta and Sorai (1966) as follows:

$$\int_0^1 \int_0^1 \frac{c^{i+j} u^2 - 1}{\binom{r}{i} \binom{r}{j}} e^{-u} du \frac{t^2 - 1}{\binom{r}{i} \binom{r}{j}} e^{-t} dt = P^* \quad (1.1.4)$$

1.1. Selection Rules for the Location Parameter under Partial Ordering Prior Assumption

Assume that we have only a partial ordering prior of k unknown location parameters, that is, the parameter space

of $\mathbb{R}^k$ and there is a partial ordering on them

for an arbitrary partition for set $\{x_1, \dots, x_k\}$ of $\mathbb{R}^k$

sets, say $B_1, \dots, B_k$, so that $x_i \in B_i$. We let $\mathcal{B} = \{B_1, \dots, B_k\}$

and for each B_i , $i = 1, \dots, k$ form a linear ordering on B_i . There is no order relation among the elements of distinct B_i .

Let $b_i = |B_i|$, the number of elements contained in B_i , so we have

$$\sum_{i=1}^k b_i = k.$$

If we denote the new induced partial ordering by $\leq_{\mathcal{B}}$, then $\mathcal{B}$ is a parameter space $\mathcal{B} \in \mathcal{B}$. We can now proceed to form a new ordering on $\mathcal{B}$ as induced partial order.

Example. Suppose $k = 9$, and we have a partial ordering on $\mathcal{B}$ as $\mathcal{B} = \{x_1, \dots, x_9\}$ and $\mathcal{B} = \{B_1, \dots, B_9\}$ and $\mathcal{B} = \{B_1, \dots, B_9\}$ denote a linear ordering. represent this partial ordering as in figure 1.



Figure 1. A partial partial ordering

Then we have an induced partial ordering $\leq_1$ on $\mathcal{P}_1$ as shown in Figure 2.

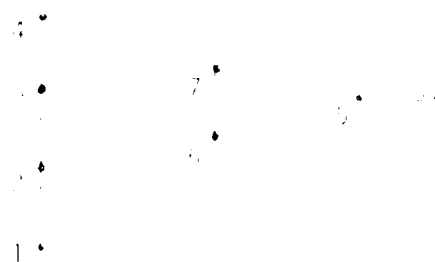


Figure 2. Induced partial ordering.

And

$$P_0^1 = \{5, 6\}$$

$$P_1^1 = \{1, 2, 3, 4\}$$

$$P_2^1 = \{6, 7\}$$

It is clear that the induced partial order is not unique, for example, we can partition $\{1, \dots, 7\}$ into three other subsets P_0^2, P_1^2, P_2^2 where

$$P_0^2 = \{5, 6\}$$

$$P_1^2 = \{1, 2, 3, 4\}$$

$$P_2^2 = \{6, 7\}$$

For the location characterizing a selection procedure $\hat{f}$, we define as follows:

Definition 1.1.1. We define a selection procedure $\hat{f}$ as follows:

Suppose $\beta_0, \dots, \beta_r$ are the induced subsets and that for each $i = 0, \dots, r$, $\beta_i = \{1, \dots, n_i\}$, there is a partial order on β_i . We choose a proper

selection procedure ϕ_{α} for each subset A_{α} such that the probability of a correct selection is not less than $1 - \alpha$.

probability of a correct selection is not less than $1 - \alpha$.

subset A_{α} we may use selection procedure ϕ_{α} or ϕ_{α} with $\alpha_{\alpha} = \alpha$.

theorem 1.5.1. The selection procedure $\bar{\phi}$ satisfies the following conditions:

$$\begin{aligned} \inf_{\theta} P_{\theta}(\text{CS}) &\geq 1 - \alpha \\ \inf_{\theta} P_{\theta}(\text{PS} = P) &\geq 1 - \alpha \\ \inf_{\theta} \text{var}_{\theta}(\bar{P}) &\leq \frac{1}{n} \\ \bar{P} &\leq \bar{h}_k \\ P(\bar{P} \leq \bar{h}_k) &= 1 - \alpha \end{aligned}$$

where $\bar{h}_k$ is the parameter value associated with the subset A_{α} .

Remark 1.5.1. For the selective procedure the only parameter of the gamma distributions, which must be known in advance, is the parameter of the selection procedure can be properly selected, hence it is not critical.

1.6. Comparison of the Performance of Selection Procedures

1.6.1. The location parameters of gamma distribution

In this section we use Monte Carlo technique to compare the performance of selection procedures ϕ_{α} , ϕ_{α} and $\bar{\phi}$ with $\alpha = 0.05$.

k independent populations, each population with distribution $N(\mu_i, \sigma^2)$, with common known variance σ^2 and common sample size n . Assume that the mean μ_0 of the control is known; without loss of generality we assume that $\mu_0 = 0$ and $\sigma^2/n = 1$.

In the simulation, we use Rubin and Hinkle's RVP-Random Variable Package, Purdue University Computing Center, to generate random numbers. For each k , we generated one random number (variable) for each population, then applied each selection procedure separately and repeated it ten thousand times; we used the relative frequency as an approximation of the exact values of the associated performance characteristics for each procedure. In Table V we use the following notations:

$\mu = (\mu_1, \dots, \mu_k)$, μ_i is the parameter of population π_i .

PS = $P(CS)$

PI = $P(\text{correctly rejecting all bad populations})$

PC = $P(\text{correct classification of all population})$

where the correct classification means that we select all good populations and reject all bad populations.

EL = Expected number (size) of bad populations contained in the selected subset.

ED = $\sum_{i=1}^k (\mu_i - 0)^2 \cdot P(\pi_i \text{ is selected})$

ES = Expected size of the selected subset.

Table V.1 consists of four parts, namely, the first, $k = 1$, $k = 2$, $k = 3$, $k = 4$, for each value of k we assume that the best population is the one and only one bad population with respect to μ that is less than the control ($\mu = 0$). At glance at the table it is clear that the performance can roughly be ordered as follows:

$$r_2 \approx r_3 > r_4 > r_1$$

i.e. procedure r_2 is the best one and, finally, better than r_3 , r_2 and r_3 are very close and both are better than r_4 . The conclusion is based on the characteristics $\bar{t}_1$, $\bar{t}_2$, $\bar{t}_3$, $\bar{t}_4$ and $\bar{t}_5$ and $\bar{t}_6$ given in Section 1.1.1. As the number of populations k increases from 1 to five and the three additional populations are good with respect to parameter 1, 2, and 3, respectively, we find that $\bar{t}_i$ ($i = 1, 2, 3, 4$) is 0.6124, 0.919, 0.9951, and 0.9999, respectively. This means that when k increases and the additional populations are good, then r_2 is the best sensitive procedure with respect to μ .

Table V.2 shows the same situation as Table V.1 but with respect to σ . Assume that we have the best population among k populations with respect to the performance characteristics $\bar{t}_1$, $\bar{t}_2$, $\bar{t}_3$, $\bar{t}_4$ and $\bar{t}_5$ and that the remaining populations are bad.

$$r_1 > r_2 > r_3 > r_4$$

In Table V.2 we observe that we have the best procedure for $k = 1$, $k = 2$, and that for the remaining values of k the best procedure is r_1 . The reason for this is that in Table V.2 $\bar{t}_1$ is 0.9999, $\bar{t}_2$ is 0.9999, $\bar{t}_3$ is 0.9999, $\bar{t}_4$ is 0.9999, and $\bar{t}_5$ is 0.9999, by using strategy r_1 we can find the best population with respect to σ .

populations has exponential distribution (mean θ to be estimated) with $\theta \in [0, 1]$ with unknown scale parameters $\theta_1, \theta_2, \dots, \theta_n$.

In the simulation study, the parameter θ is chosen as $\theta = 0.6$ (we used Rubin and Hinkley's PVP algorithm for generating random variable) with specified exponential distribution. The simulation results are based on two thousand random observations per population. In Table VI.1 and Table VI.2, we show the results of the test T_{α} (1.6.1) but the number of a bad population is chosen to be 1000 (i.e., a bad population has a parameter which is greater than 0.6 (assumed one)).

A good population means that its parameter is less than 0.6 (less than one). The results of Table VI.1 and Table VI.2 indicate that we have the performance $\beta_6 < \beta_7 < \beta_8 < \beta_9$ and it implies that β_8 is slightly better than β_7 if there is only one bad population.

Table of $d_{13}^{(1)}$ values calculated for $\alpha = 0$ and $\beta = 1$ according to the procedure (1) for the normal distribution problem under the uniform ordering prior.

$d_{13}^{(1)}$	p^*				
r	.39	.375	.36	.345	.33
1	2.3264	1.9655	1.6449	1.4361	1.3118
2	2.3337	1.9775	1.6756	1.4861	1.3431
3	2.3379	1.9837	1.6847	1.4961	1.3531
4	--	--	1.6873	1.4991	1.3561
5	--	--	1.6896	1.4991	1.3571
6	--	--	--	--	1.3571
	2.3340	1.9787	1.6824	1.4961	1.3531
$d_{13}^{(1)}$	p^*				
r	.39	.38	.36	.34	.33
1	1.0964	1.0416	.6735	.5631	.5161
2	1.1739	1.0937	1.0396	.7976	.7361
3	1.1428	1.0416	1.0371	.7831	.7361
4	1.1498	1.0417	1.0346	.7831	.7361
5	1.1506	1.0417	1.0346	.7831	.7361
6	1.1514	1.0417	1.0346	.7831	.7361
7	1.1517	1.0417	1.0346	.7831	.7361
8	1.1516	1.0416	1.0346	.7831	.7361
9	1.1516	1.0416	1.0346	.7831	.7361

Note: For $p^* = 1$ and $p^* = 0$, the values of $d_{13}^{(1)}$ are the same as the values of $d_{13}^{(0)}$ and $d_{13}^{(2)}$, respectively.

Table of $d_{1k}^{(2)}$ values: satisfactory for use in the procedure $d_{1k}^{(2)}$ for the normal mean and variance.

$d_{1k}^{(2)}$	$d_{1k}^{(2)}$	$d_{1k}^{(2)}$	$d_{1k}^{(2)}$	$d_{1k}^{(2)}$	$d_{1k}^{(2)}$
1	1.100	1.171	1.242	1.313	1.384
2	1.400	1.500	1.600	1.700	1.800
3	1.900	2.000	2.100	2.200	2.300
4	2.400	2.500	2.600	2.700	2.800
5	2.900	3.000	3.100	3.200	3.300
6	3.400	3.500	3.600	3.700	3.800
7	3.900	4.000	4.100	4.200	4.300
8	4.400	4.500	4.600	4.700	4.800
9	4.900	5.000	5.100	5.200	5.300
10	5.400	5.500	5.600	5.700	5.800
11	5.900	6.000	6.100	6.200	6.300
12	6.400	6.500	6.600	6.700	6.800
13	6.900	7.000	7.100	7.200	7.300
14	7.400	7.500	7.600	7.700	7.800
15	7.900	8.000	8.100	8.200	8.300
16	8.400	8.500	8.600	8.700	8.800
17	8.900	9.000	9.100	9.200	9.300
18	9.400	9.500	9.600	9.700	9.800
19	9.900	10.000	10.100	10.200	10.300
20	10.400	10.500	10.600	10.700	10.800
21	10.900	11.000	11.100	11.200	11.300
22	11.400	11.500	11.600	11.700	11.800
23	11.900	12.000	12.100	12.200	12.300
24	12.400	12.500	12.600	12.700	12.800
25	12.900	13.000	13.100	13.200	13.300
26	13.400	13.500	13.600	13.700	13.800
27	13.900	14.000	14.100	14.200	14.300
28	14.400	14.500	14.600	14.700	14.800
29	14.900	15.000	15.100	15.200	15.300
30	15.400	15.500	15.600	15.700	15.800
31	15.900	16.000	16.100	16.200	16.300
32	16.400	16.500	16.600	16.700	16.800
33	16.900	17.000	17.100	17.200	17.300
34	17.400	17.500	17.600	17.700	17.800
35	17.900	18.000	18.100	18.200	18.300
36	18.400	18.500	18.600	18.700	18.800
37	18.900	19.000	19.100	19.200	19.300
38	19.400	19.500	19.600	19.700	19.800
39	19.900	20.000	20.100	20.200	20.300
40	20.400	20.500	20.600	20.700	20.800
41	20.900	21.000	21.100	21.200	21.300
42	21.400	21.500	21.600	21.700	21.800
43	21.900	22.000	22.100	22.200	22.300
44	22.400	22.500	22.600	22.700	22.800
45	22.900	23.000	23.100	23.200	23.300
46	23.400	23.500	23.600	23.700	23.800
47	23.900	24.000	24.100	24.200	24.300
48	24.400	24.500	24.600	24.700	24.800
49	24.900	25.000	25.100	25.200	25.300
50	25.400	25.500	25.600	25.700	25.800
51	25.900	26.000	26.100	26.200	26.300
52	26.400	26.500	26.600	26.700	26.800
53	26.900	27.000	27.100	27.200	27.300
54	27.400	27.500	27.600	27.700	27.800
55	27.900	28.000	28.100	28.200	28.300
56	28.400	28.500	28.600	28.700	28.800
57	28.900	29.000	29.100	29.200	29.300
58	29.400	29.500	29.600	29.700	29.800
59	29.900	30.000	30.100	30.200	30.300
60	30.400	30.500	30.600	30.700	30.800
61	30.900	31.000	31.100	31.200	31.300
62	31.400	31.500	31.600	31.700	31.800
63	31.900	32.000	32.100	32.200	32.300
64	32.400	32.500	32.600	32.700	32.800
65	32.900	33.000	33.100	33.200	33.300
66	33.400	33.500	33.600	33.700	33.800
67	33.900	34.000	34.100	34.200	34.300
68	34.400	34.500	34.600	34.700	34.800
69	34.900	35.000	35.100	35.200	35.300
70	35.400	35.500	35.600	35.700	35.800
71	35.900	36.000	36.100	36.200	36.300
72	36.400	36.500	36.600	36.700	36.800
73	36.900	37.000	37.100	37.200	37.300
74	37.400	37.500	37.600	37.700	37.800
75	37.900	38.000	38.100	38.200	38.300
76	38.400	38.500	38.600	38.700	38.800
77	38.900	39.000	39.100	39.200	39.300
78	39.400	39.500	39.600	39.700	39.800
79	39.900	40.000	40.100	40.200	40.300
80	40.400	40.500	40.600	40.700	40.800
81	40.900	41.000	41.100	41.200	41.300
82	41.400	41.500	41.600	41.700	41.800
83	41.900	42.000	42.100	42.200	42.300
84	42.400	42.500	42.600	42.700	42.800
85	42.900	43.000	43.100	43.200	43.300
86	43.400	43.500	43.600	43.700	43.800
87	43.900	44.000	44.100	44.200	44.300
88	44.400	44.500	44.600	44.700	44.800
89	44.900	45.000	45.100	45.200	45.300
90	45.400	45.500	45.600	45.700	45.800
91	45.900	46.000	46.100	46.200	46.300
92	46.400	46.500	46.600	46.700	46.800
93	46.900	47.000	47.100	47.200	47.300
94	47.400	47.500	47.600	47.700	47.800
95	47.900	48.000	48.100	48.200	48.300
96	48.400	48.500	48.600	48.700	48.800
97	48.900	49.000	49.100	49.200	49.300
98	49.400	49.500	49.600	49.700	49.800
99	49.900	50.000	50.100	50.200	50.300
100	50.400	50.500	50.600	50.700	50.800
101	50.900	51.000	51.100	51.200	51.300
102	51.400	51.500	51.600	51.700	51.800
103	51.900	52.000	52.100	52.200	52.300
104	52.400	52.500	52.600	52.700	52.800
105	52.900	53.000	53.100	53.200	53.300
106	53.400	53.500	53.600	53.700	53.800
107	53.900	54.000	54.100	54.200	54.300
108	54.400	54.500	54.600	54.700	54.800
109	54.900	55.000	55.100	55.200	55.300
110	55.400	55.500	55.600	55.700	55.800
111	55.900	56.000	56.100	56.200	56.300
112	56.400	56.500	56.600	56.700	56.800
113	56.900	57.000	57.100	57.200	57.300
114	57.400	57.500	57.600	57.700	57.800
115	57.900	58.000	58.100	58.200	58.300
116	58.400	58.500	58.600	58.700	58.800
117	58.900	59.000	59.100	59.200	59.300
118	59.400	59.500	59.600	59.700	59.800
119	59.900	60.000	60.100	60.200	60.300
120	60.400	60.500	60.600	60.700	60.800
121	60.900	61.000	61.100	61.200	61.300
122	61.400	61.500	61.600	61.700	61.800
123	61.900	62.000	62.100	62.200	62.300
124	62.400	62.500	62.600	62.700	62.800
125	62.900	63.000	63.100	63.200	63.300
126	63.400	63.500	63.600	63.700	63.800
127	63.900	64.000	64.100	64.200	64.300
128	64.400	64.500	64.600	64.700	64.800
129	64.900	65.000	65.100	65.200	65.300
130	65.400	65.500	65.600	65.700	65.800
131	65.900	66.000	66.100	66.200	66.300
132	66.400	66.500	66.600	66.700	66.800
133	66.900	67.000	67.100	67.200	67.300
134	67.400	67.500	67.600	67.700	67.800
135	67.900	68.000	68.100	68.200	68.300
136	68.400	68.500	68.600	68.700	68.800
137	68.900	69.000	69.100	69.200	69.300
138	69.400	69.500	69.600	69.700	69.800
139	69.900	70.000	70.100	70.200	70.300
140	70.400	70.500	70.600	70.700	70.800
141	70.900	71.000	71.100	71.200	71.300
142	71.400	71.500	71.600	71.700	71.800
143	71.900	72.000	72.100	72.200	72.300
144	72.400	72.500	72.600	72.700	72.800
145	72.900	73.000	73.100	73.200	73.300
146	73.400	73.500	73.600	73.700	73.800
147	73.900	74.000	74.100	74.200	74.300
148	74.400	74.500	74.600	74.700	74.800
149	74.900	75.000	75.100	75.200	75.300
150	75.400	75.500	75.600	75.700	75.800
151	75.900	76.000	76.100	76.200	76.300
152	76.400	76.500	76.600	76.700	76.800
153	76.900	77.000	77.100	77.200	77.300
154	77.400	77.500	77.600	77.700	77.800
155	77.900	78.000	78.100	78.200	78.300
156	78.400	78.500	78.600	78.700	78.800
157	78.900	79.000	79.100	79.200	79.300
158	79.400	79.500	79.600	79.700	79.800
159	79.900	80.000	80.100	80.200	80.300
160	80.400	80.500	80.600	80.700	80.800
161	80.900	81.000	81.100	81.200	81.300
162	81.400	81.500	81.600	81.700	81.800
163	81.900	82.000	82.100	82.200	82.300
164	82.400	82.500	82.600	82.700	82.800
165	82.900	83.000	83.100	83.200	83.300
166	83.400	83.500	83.600	83.700	83.800
167	83.900	84.000	84.100	84.200	84.300
168	84.400	84.500	84.600	84.700	84.800
169	84.900	85.000	85.100	85.200	85.300
170	85.400	85.500	85.600	85.700	85.800
171	85.900	86.000	86.100	86.200	86.300
172	86.400	86.500	86.600	86.700	86.800
173	86.900	87.000	87.100	87.200	87.300
174	87.400	87.500	87.600	87.700	87.800
175	87.900	88.000	88.100	88.200	88.300
176	88.400	88.500	88.600	88.700	88.800
177	88.900	89.000	89.100	89.200	89.300
178	89.400	89.500	89.600	89.700	89.800
179	89.900	90.000	90.100		

1. The first part of the report deals with the general situation of the country and the progress of the work during the year. It is a summary of the work done and the results obtained. It is a general statement of the work done and the results obtained. It is a general statement of the work done and the results obtained.

2. The second part of the report deals with the specific work done during the year. It is a detailed statement of the work done and the results obtained. It is a detailed statement of the work done and the results obtained. It is a detailed statement of the work done and the results obtained.

3. The third part of the report deals with the financial statement of the work done during the year. It is a statement of the money received and the money spent. It is a statement of the money received and the money spent. It is a statement of the money received and the money spent.

4. The fourth part of the report deals with the conclusions of the work done during the year. It is a statement of the results of the work done and the results obtained. It is a statement of the results of the work done and the results obtained. It is a statement of the results of the work done and the results obtained.

5. The fifth part of the report deals with the recommendations of the work done during the year. It is a statement of the suggestions for the future work and the results obtained. It is a statement of the suggestions for the future work and the results obtained. It is a statement of the suggestions for the future work and the results obtained.

TABLE 1.

Summary of the results of the investigation of the effects of the various factors on the rate of the reaction between the various substances and the various conditions.

Substance	Condition	Rate of Reaction	Temperature	Concentration	Time
Hydrogen	Normal	1.0	25°C	1.0	1.0
Hydrogen	High	1.5	25°C	1.0	1.0
Hydrogen	Low	0.5	25°C	1.0	1.0
Hydrogen	Normal	1.0	35°C	1.0	1.0
Hydrogen	Normal	1.0	45°C	1.0	1.0
Hydrogen	Normal	1.0	55°C	1.0	1.0
Hydrogen	Normal	1.0	65°C	1.0	1.0
Hydrogen	Normal	1.0	75°C	1.0	1.0
Hydrogen	Normal	1.0	85°C	1.0	1.0
Hydrogen	Normal	1.0	95°C	1.0	1.0
Hydrogen	Normal	1.0	105°C	1.0	1.0
Hydrogen	Normal	1.0	115°C	1.0	1.0
Hydrogen	Normal	1.0	125°C	1.0	1.0
Hydrogen	Normal	1.0	135°C	1.0	1.0
Hydrogen	Normal	1.0	145°C	1.0	1.0
Hydrogen	Normal	1.0	155°C	1.0	1.0
Hydrogen	Normal	1.0	165°C	1.0	1.0
Hydrogen	Normal	1.0	175°C	1.0	1.0
Hydrogen	Normal	1.0	185°C	1.0	1.0
Hydrogen	Normal	1.0	195°C	1.0	1.0
Hydrogen	Normal	1.0	205°C	1.0	1.0
Hydrogen	Normal	1.0	215°C	1.0	1.0
Hydrogen	Normal	1.0	225°C	1.0	1.0
Hydrogen	Normal	1.0	235°C	1.0	1.0
Hydrogen	Normal	1.0	245°C	1.0	1.0
Hydrogen	Normal	1.0	255°C	1.0	1.0
Hydrogen	Normal	1.0	265°C	1.0	1.0
Hydrogen	Normal	1.0	275°C	1.0	1.0
Hydrogen	Normal	1.0	285°C	1.0	1.0
Hydrogen	Normal	1.0	295°C	1.0	1.0
Hydrogen	Normal	1.0	305°C	1.0	1.0
Hydrogen	Normal	1.0	315°C	1.0	1.0
Hydrogen	Normal	1.0	325°C	1.0	1.0
Hydrogen	Normal	1.0	335°C	1.0	1.0
Hydrogen	Normal	1.0	345°C	1.0	1.0
Hydrogen	Normal	1.0	355°C	1.0	1.0
Hydrogen	Normal	1.0	365°C	1.0	1.0
Hydrogen	Normal	1.0	375°C	1.0	1.0
Hydrogen	Normal	1.0	385°C	1.0	1.0
Hydrogen	Normal	1.0	395°C	1.0	1.0
Hydrogen	Normal	1.0	405°C	1.0	1.0
Hydrogen	Normal	1.0	415°C	1.0	1.0
Hydrogen	Normal	1.0	425°C	1.0	1.0
Hydrogen	Normal	1.0	435°C	1.0	1.0
Hydrogen	Normal	1.0	445°C	1.0	1.0
Hydrogen	Normal	1.0	455°C	1.0	1.0
Hydrogen	Normal	1.0	465°C	1.0	1.0
Hydrogen	Normal	1.0	475°C	1.0	1.0
Hydrogen	Normal	1.0	485°C	1.0	1.0
Hydrogen	Normal	1.0	495°C	1.0	1.0
Hydrogen	Normal	1.0	505°C	1.0	1.0
Hydrogen	Normal	1.0	515°C	1.0	1.0
Hydrogen	Normal	1.0	525°C	1.0	1.0
Hydrogen	Normal	1.0	535°C	1.0	1.0
Hydrogen	Normal	1.0	545°C	1.0	1.0
Hydrogen	Normal	1.0	555°C	1.0	1.0
Hydrogen	Normal	1.0	565°C	1.0	1.0
Hydrogen	Normal	1.0	575°C	1.0	1.0
Hydrogen	Normal	1.0	585°C	1.0	1.0
Hydrogen	Normal	1.0	595°C	1.0	1.0
Hydrogen	Normal	1.0	605°C	1.0	1.0
Hydrogen	Normal	1.0	615°C	1.0	1.0
Hydrogen	Normal	1.0	625°C	1.0	1.0
Hydrogen	Normal	1.0	635°C	1.0	1.0
Hydrogen	Normal	1.0	645°C	1.0	1.0
Hydrogen	Normal	1.0	655°C	1.0	1.0
Hydrogen	Normal	1.0	665°C	1.0	1.0
Hydrogen	Normal	1.0	675°C	1.0	1.0
Hydrogen	Normal	1.0	685°C	1.0	1.0
Hydrogen	Normal	1.0	695°C	1.0	1.0
Hydrogen	Normal	1.0	705°C	1.0	1.0
Hydrogen	Normal	1.0	715°C	1.0	1.0
Hydrogen	Normal	1.0	725°C	1.0	1.0
Hydrogen	Normal	1.0	735°C	1.0	1.0
Hydrogen	Normal	1.0	745°C	1.0	1.0
Hydrogen	Normal	1.0	755°C	1.0	1.0
Hydrogen	Normal	1.0	765°C	1.0	1.0
Hydrogen	Normal	1.0	775°C	1.0	1.0
Hydrogen	Normal	1.0	785°C	1.0	1.0
Hydrogen	Normal	1.0	795°C	1.0	1.0
Hydrogen	Normal	1.0	805°C	1.0	1.0
Hydrogen	Normal	1.0	815°C	1.0	1.0
Hydrogen	Normal	1.0	825°C	1.0	1.0
Hydrogen	Normal	1.0	835°C	1.0	1.0
Hydrogen	Normal	1.0	845°C	1.0	1.0
Hydrogen	Normal	1.0	855°C	1.0	1.0
Hydrogen	Normal	1.0	865°C	1.0	1.0
Hydrogen	Normal	1.0	875°C	1.0	1.0
Hydrogen	Normal	1.0	885°C	1.0	1.0
Hydrogen	Normal	1.0	895°C	1.0	1.0
Hydrogen	Normal	1.0	905°C	1.0	1.0
Hydrogen	Normal	1.0	915°C	1.0	1.0
Hydrogen	Normal	1.0	925°C	1.0	1.0
Hydrogen	Normal	1.0	935°C	1.0	1.0
Hydrogen	Normal	1.0	945°C	1.0	1.0
Hydrogen	Normal	1.0	955°C	1.0	1.0
Hydrogen	Normal	1.0	965°C	1.0	1.0
Hydrogen	Normal	1.0	975°C	1.0	1.0
Hydrogen	Normal	1.0	985°C	1.0	1.0
Hydrogen	Normal	1.0	995°C	1.0	1.0
Hydrogen	Normal	1.0	1005°C	1.0	1.0

Table 17

Use of the χ^2_{min} algorithm for quadrupole fitting. The χ^2_{min} value is given for each fit, the procedure χ^2_{min} for the quadrupole is produced with χ^2_{min} for the quadrupole and the error $\sigma_{\text{quadrupole}}$ for the quadrupole period.

χ^2_{min}		$\sigma_{\text{quadrupole}}$		χ^2_{min}	
1	2	3	4	5	6
1	1.5113	1.4896	2.1535	1.5281	1.5064
2	1.5217	2.1107	2.1556	1.5629	1.5064
3	1.5127	1.4924	1.5089	1.5131	1.5064
4	--	--	--	--	--
5	1.5063	1.5317	1.5439	1.5548	1.5064
6	1.5113	1.5099	1.5294	1.5156	1.5064
7	1.5117	1.5032	1.5387	1.5107	1.5064
8	1.5116	--	--	--	--
9	--	--	--	--	--
10	1.5078	1.5061	1.5111	1.5116	1.5064
11	1.5036	1.5033	1.5111	1.5111	1.5064
12	1.5031	1.5040	1.5116	1.5116	1.5064
13	--	--	--	--	--
14	--	--	--	--	--

Use of the χ^2_{min} algorithm for quadrupole fitting. The χ^2_{min} value is given for each fit, the procedure χ^2_{min} for the quadrupole is produced with χ^2_{min} for the quadrupole and the error $\sigma_{\text{quadrupole}}$ for the quadrupole period.

J	B					
	4	5	6	7	8	9
1	1.6732	1.5836	1.5719	1.4736	1.441	1.417
2	1.7238	1.5558	1.5357	1.4565	1.4236	1.4004
3	1.7406	1.5404	1.5137	1.4553	1.4229	1.3997
4	1.7440	1.5976	1.5157	1.4579	1.4231	1.3994
5	1.7451	1.5365	1.5156	1.4584	1.4234	1.399
6	1.7455	1.6038	1.5156	1.4585	1.4235	1.399
7	1.7456	1.6039	1.5156	1.4586	--	--
8	1.7457	--	--	--	--	--
10	--	--	--	--	--	--
1	2.7317	2.7567	2.6615	2.5905	2.54799	2.5414
2	3.0126	2.8191	2.7351	2.6079	2.5373	2.54004
3	3.0264	2.8233	2.7133	2.6350	2.5144	2.54334
4	3.0301	2.8320	2.7155	2.6361	2.5158	2.5445
5	3.0313	2.8328	2.7157	2.6374	2.5162	2.5446
6	3.0317	2.8331	2.7157	2.6376	2.5164	2.5446
7	3.0318	2.8332	2.7158	2.6377	--	2.5446
8	3.0319	--	--	--	--	--
10	--	--	--	--	--	--
1	3.1495	3.1634	3.1609	3.1719	3.16858	3.16814
2	3.1736	3.19678	3.17633	3.17635	3.16496	3.16814
3	3.17439	3.19667	3.16663	3.17717	3.16459	3.16814
4	3.17420	3.19992	3.16666	3.17737	3.16474	3.16814
5	3.17497	3.19101	3.16666	3.17734	3.16477	3.16814
6	3.17495	3.19104	3.16667	3.17736	3.16477	3.16814
7	3.17493	3.19105	3.16666	3.17737	--	3.16814
10	--	--	--	--	--	--

and the mean values of the average percentage of larvae in the first four periods of the life cycle of the parasite, calculated according to the formula: $\bar{P} = (P_1 + P_2 + P_3 + P_4) / 4$.

$P_1 = 1000$

$K = 1, \quad (-1, 1)$

$\bar{P} = 1$

P_1	1.9453	1.9453	1.9453	1.9453
P_2	1.3394	1.3394	1.3394	1.3394
P_3	1.3507	1.3507	1.3507	1.3507
P_4	1.6146	1.6146	1.6146	1.6146
P_5	1.6146	1.6146	1.6146	1.6146
P_6	1.5599	1.5599	1.5599	1.5599

$K = 3, \quad (-1, 1, 1)$

$\bar{P} = 1$

P_1	1.9531	1.9531	1.9531	1.9531
P_2	1.3741	1.3741	1.3741	1.3741
P_3	1.3777	1.3777	1.3777	1.3777
P_4	1.6259	1.6259	1.6259	1.6259
P_5	1.6259	1.6259	1.6259	1.6259
P_6	1.5771	1.5771	1.5771	1.5771

$K = 4, \quad (-1, 0, 1, 1)$

$\bar{P} = 1$

P_1	1.9580	1.9580	1.9580	1.9580
P_2	1.3664	1.3664	1.3664	1.3664
P_3	1.3744	1.3744	1.3744	1.3744
P_4	1.6136	1.6136	1.6136	1.6136
P_5	1.6136	1.6136	1.6136	1.6136
P_6	1.5901	1.5901	1.5901	1.5901

$K = 5, \quad (-1, 0, 1, 1, 1)$

$\bar{P} = 1$

P_1	1.9581	1.9581	1.9581	1.9581
P_2	1.3735	1.3735	1.3735	1.3735
P_3	1.3754	1.3754	1.3754	1.3754
P_4	1.6175	1.6175	1.6175	1.6175
P_5	1.6175	1.6175	1.6175	1.6175
P_6	1.5711	1.5711	1.5711	1.5711

1. *Journal of the American Medical Association*, 1997; 277: 1039-1043.

1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 26

1. *Chlorophyll a* (Chl *a*)

The following are the results of the analysis of the data obtained from the
 procedure described in the preceding section. The results are presented in
 Table 1. The data are given for the following cases:

(a) $\alpha = 0.1$

(b) $\alpha = 0.2$

(c) $\alpha = 0.3$

(d) $\alpha = 0.4$

(e) $\alpha = 0.5$

(f) $\alpha = 0.6$

(g) $\alpha = 0.7$

(h) $\alpha = 0.8$

(i) $\alpha = 0.9$

(j) $\alpha = 1.0$

(k) $\alpha = 1.1$

(l) $\alpha = 1.2$

(m) $\alpha = 1.3$

(n) $\alpha = 1.4$

(o) $\alpha = 1.5$

(p) $\alpha = 1.6$

(q) $\alpha = 1.7$

(r) $\alpha = 1.8$

(s) $\alpha = 1.9$

(t) $\alpha = 2.0$

(u) $\alpha = 2.1$

(v) $\alpha = 2.2$

(w) $\alpha = 2.3$

(x) $\alpha = 2.4$

(y) $\alpha = 2.5$

(z) $\alpha = 2.6$

(aa) $\alpha = 2.7$

(ab) $\alpha = 2.8$

(ac) $\alpha = 2.9$

(ad) $\alpha = 3.0$

(ae) $\alpha = 3.1$

(af) $\alpha = 3.2$

(ag) $\alpha = 3.3$

(ah) $\alpha = 3.4$

(ai) $\alpha = 3.5$

(aj) $\alpha = 3.6$

(ak) $\alpha = 3.7$

(al) $\alpha = 3.8$

(am) $\alpha = 3.9$

(an) $\alpha = 4.0$

(ao) $\alpha = 4.1$

(ap) $\alpha = 4.2$

(aq) $\alpha = 4.3$

(ar) $\alpha = 4.4$

(as) $\alpha = 4.5$

(at) $\alpha = 4.6$

(au) $\alpha = 4.7$

(av) $\alpha = 4.8$

(aw) $\alpha = 4.9$

(ax) $\alpha = 5.0$

Simulation results for the comparative performance of various selection procedures for the normal mean, *variance* (notation explained in section 1.6.1) under simple random sampling.

* $\sigma^2 = 1000$					
K = 10					
1					
P0	1.0000	1.0000	1.0000	1.0000	1.0000
P1	.9613	.9760	.9600	.9600	.9600
P2	.9613	.9760	.9600	.9600	.9600
P3	.9613	.9760	.9600	.9600	.9600
P4	.9613	.9760	.9600	.9600	.9600
P5	.9613	.9760	.9600	.9600	.9600
P6	.9613	.9760	.9600	.9600	.9600
K = 10					
2					
P0	1.0000	1.0000	1.0000	1.0000	1.0000
P1	.9613	.9760	.9600	.9600	.9600
P2	.9613	.9760	.9600	.9600	.9600
P3	.9613	.9760	.9600	.9600	.9600
P4	.9613	.9760	.9600	.9600	.9600
P5	.9613	.9760	.9600	.9600	.9600
P6	.9613	.9760	.9600	.9600	.9600
K = 10					
3					
P0	1.0000	1.0000	1.0000	1.0000	1.0000
P1	.9613	.9760	.9600	.9600	.9600
P2	.9613	.9760	.9600	.9600	.9600
P3	.9613	.9760	.9600	.9600	.9600
P4	.9613	.9760	.9600	.9600	.9600
P5	.9613	.9760	.9600	.9600	.9600
P6	.9613	.9760	.9600	.9600	.9600
K = 10					
4					
P0	1.0000	1.0000	1.0000	1.0000	1.0000
P1	.9613	.9760	.9600	.9600	.9600
P2	.9613	.9760	.9600	.9600	.9600
P3	.9613	.9760	.9600	.9600	.9600
P4	.9613	.9760	.9600	.9600	.9600
P5	.9613	.9760	.9600	.9600	.9600
P6	.9613	.9760	.9600	.9600	.9600
K = 10					
5					
P0	1.0000	1.0000	1.0000	1.0000	1.0000
P1	.9613	.9760	.9600	.9600	.9600
P2	.9613	.9760	.9600	.9600	.9600
P3	.9613	.9760	.9600	.9600	.9600
P4	.9613	.9760	.9600	.9600	.9600
P5	.9613	.9760	.9600	.9600	.9600
P6	.9613	.9760	.9600	.9600	.9600

TABLE V. 5

Simulation results for the comparative performance of various selection procedures for the normal means problem (notation explained in Section 1.6.1) under simple ordering setup.

$p^* = .900$				
$k = 2, \theta = (-1, +1)$				
	1	2	3	4
PS	1.0000	1.0000	1.0000	1.0000
PI	.5405	.5347	.2937	.1722
PC	.5405	.5349	.2937	.1737
LI	.8331	.8337	1.3151	1.0004
LS	2.2116	2.2340	3.4732	2.1578
FS	.8331	.8337	1.3151	1.0964
$k = 3, \theta = (-1, +1, +1)$				
	1	2	3	4
PS	.9937	.9947	.9957	.9976
PI	.5365	.5347	.2937	.1197
PC	.5297	.5347	.2944	.1139
LI	.8347	.8365	1.3116	1.1154
LS	2.2257	2.2510	3.4155	2.1491
FS	1.8279	1.8306	2.3075	2.1196
$k = 4, \theta = (-1, +1, +1, +1)$				
	1	2	3	4
PS	.9921	.9937	.9975	.9976
PI	.5271	.5365	.2924	.1077
PC	.5197	.5365	.2967	.1047
LI	.8398	.8365	1.3135	1.1077
LS	2.2655	2.2597	3.4734	2.2375
FS	1.8396	1.8365	2.3202	2.1377
$k = 5, \theta = (-1, +2, 1, 1, +1)$				
	1	2	3	4
PS	.9906	.9937	.9993	.9993
PI	.5317	.5365	.2937	.1077
PC	.5275	.5365	.2937	.1077
LI	.8461	.8365	1.3135	1.1077
LS	2.2551	2.2597	3.4734	2.2375
FS	1.8341	1.8365	2.3135	2.1377

TABLE 11.1

Simulation results for the comparative performance of various selection procedures for the gamma means problem (notation explained in Section 1.6.1) under simple ordering prior.

$P^* = .500$

$k = 4, \theta = (.3, .6, .9, 1.2)$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$
PS	.0840	.0645	.0000	.0000
PI	.2029	.1745	.1989	.0000
PC	.1866	.1687	.0000	.0000
LI	.7980	.6655	.6645	.0000
LD	.0319	.0000	.0000	.0000
ES	1.7720	3.8100	1.5417	.0000

$k = 6, \theta = (.5, .6, .9, 1.2, 1.5)$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$
PS	.0905	.0910	.0965	.0000	.0000
PI	.0990	.0950	.0645	.0000	.0000
PC	.0695	.0650	.0615	.0000	.0000
LI	1.5040	1.5560	1.5890	.0000	.0000
LD	.1868	.1909	.0000	.0000	.0000
ES	4.4945	4.5470	4.5855	4.7700	.0000

$k = 4, \theta = (.5, 1, 1.5, 2)$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$
PS	.0685	.0770	.0700	.0000
PI	.2685	.2560	.2100	.0000
PC	.2370	.2260	.1920	.0000
LI	1.0790	1.0000	1.1100	.0000
LD	.5304	.5000	.4900	.0000
ES	2.0475	1.9900	2.0000	.0000

$k = 6, \theta = (.5, 1, 1.5, 2, 2.5)$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$
PS	.0600	.0600	.0675	.0000	.0000
PI	.2500	.2500	.2000	.0000	.0000
PC	.2000	.2000	.1600	.0000	.0000
LI	1.2000	1.2000	1.2000	.0000	.0000
LD	.5000	.5000	.4600	.0000	.0000
ES	2.0000	2.0000	2.0000	2.0000	.0000

TABLE VI.

Simulation results for the comparative performance of various selection procedures for the gamma mean problem (notation explained in Section 1.6.1) under simple ordering prior.

$P^* = .900$

$k = 4, \quad \gamma = (.1, .6, 1.1, 1.6)$

	6	7	8	9
PS	.9995	.9995	1.0000	.9995
PI	.0690	.0645	.0545	.0110
PC	.0685	.0640	.0545	.0110
LI	1.4940	1.5345	1.5315	1.7500
LI	.2120	.2250	.2204	.2940
LS	3.4935	3.5340	3.5315	3.7495

$k = 5, \quad \gamma = (.1, .6, 1.1, 1.6, 2.1)$

	6	7	8	9
PS	.9990	.9995	1.0000	.9995
PI	.0540	.0530	.0795	.0130
PC	.0530	.0520	.0295	.0030
LI	1.9205	1.9700	2.1385	2.4385
LI	.6140	.6646	.7727	1.0577
LS	3.9195	3.9640	4.1325	4.4000

$k = 4, \quad \gamma = (1.1, 1.4, 1.7, 1.9)$

	6	7	8	9
PS	1.0000	1.0000	1.0000	1.0000
PI	.0375	.0375	.0095	.0000
PC	.0375	.0375	.0095	.0000
LI	.6890	.7575	.8115	.8600
LI	.6796	.7710	.8000	.8600
LS	2.6890	2.7475	3.1155	3.3600

$k = 5, \quad \gamma = (1.1, 1.4, 1.7, 1.9, 2.2)$

	6	7	8	9
PS	1.0000	1.0000	1.0000	1.0000
PI	.0315	.0375	.0000	.0000
PC	.0315	.0375	.0050	.0000
LI	.9750	1.1000	1.2600	1.4750
LI	1.0000	1.1000	1.2600	1.9500
LS	3.9750	4.1000	4.4600	4.8750

CHAPTER 11 BAYES + P^{*} SELECTION RULES FOR SELECTING A SUBSET CONTAINING THE BEST POPULATION

11.1. Introduction

Suppose we have $k(k \geq 2)$ independent populations $\pi_1, \pi_2, \dots, \pi_k$ and that the random variable X_i associated with π_i has a distribution with r_i known parameters $\theta_i, i = 1, \dots, k$. First, we give some definitions.

Definition 11.1.1. The population π_j is the best population of π_i for all $i \neq j$. If there are more than one populations with this condition we arbitrarily tag one of them and call it the best population which is not the best is called a non-best population.

Assume that we have n_i independent observations $X_{i1}, X_{i2}, \dots, X_{in_i}$ for population $\pi_i, i = 1, \dots, k$, and that $\bar{x}_i = (x_{i1}, \dots, x_{in_i})'$ is an estimator of $\theta_i, i = 1, \dots, k$, so that $\bar{x}_i$'s are independent. If a vector $\bar{x}$ is obtained from $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ then $\bar{x}$ is called a sample. For each π_i we have two mutually exclusive events, π_i is selected or π_i is not selected. Let π_i be selected if $\bar{x}_i \in R_i$ and π_i is not selected if $\bar{x}_i \notin R_i$ where

$$R_i = \{ \bar{x}_i : \bar{x}_i \in R_i \} \text{ where } R_i = \{ \bar{x}_i : \bar{x}_i \in R_i \}$$

The set R_i is called a selection region for π_i and is denoted by R_i .

$1, 2, \dots, k$. An action $A \in \mathcal{A}$ is the selection of some subset of the k populations. $i \in A$ means that π_i is included in the selected subset, or action $A \in \mathcal{A}$ is called a correct selection (CS) if the best population is included in the selected subset A .

Definition 2.1.2. A measurable function s defined on $\mathcal{Z} \times \mathcal{A}$ called a selection procedure provided that for each $x \in \mathcal{Z}$, we have

$$0 \leq s(x, A)$$

and

$$\sum_{A \in \mathcal{A}} s(x, A) = 1$$

where $s(x, A)$ denotes the probability that the subset A is selected when x is observed.

The individual selection probability $\pi_i(x)$ for the population π_i is then given by

$$\pi_i(x) = \sum_{A \in \mathcal{A}} s(x, A)$$

where the summation is over all subsets A which contain i . If the selection probability $\pi_i(x)$ takes on only value 0 or 1, then we say that the selection procedure $s(x, A)$ is completely nonrandomized and is called a non-randomized procedure.

Definition 2.1.3. Two selection procedures s and $\tilde{s}$ are equivalent if they have the same individual selection probabilities $\pi_i(x)$ for all x , $i = 1, \dots, k$.

Hence we can use the following definition to replace (2.1.1).

Definition 2.1.4. A subset selection rule r is a measurable function from $\mathcal{Z}$ to $\mathbb{R}^k$,

$$r(x) = (r_1(x), \dots, r_k(x))$$

with

$$0 \leq r_i(x) \leq 1, \quad i = 1, \dots, k.$$

If r_i 's are 0 or 1, the rule is non-randomized. Note that according to Definition 2.1.1, we have $\sum_{i=1}^k r_i(x) = 1$.

Suppose our goal is to find a nontrivial subset which contains the best population. A large body of literature exists in the area of subset selection procedures (see Gupta and Vanchapakanan (1976)). Gupta (1956, 1968) gave maximum-type subset selection procedures. Hsu (1978) studied the performance of Gupta-type maximum procedure, Soal-type average procedure (Soal (1958, 1957)) and the Bayes rule. Berger (1978) and Berger and Gupta (1980) proved that Gupta-type maximum procedure minimizes under certain loss function. In the second, the metric approach to the subset selection problem, Soal and Pitterman (1962), Chernoff and Sahay (1977), Bickel and Ghosh (1977), Gupta and Ghosh (1978) and Kim (1979) gave different formulations and different loss functions. The loss functions proposed by Ghosh and Ghosh consist of two components: L_1 and L_2 and the weight assigned to L_1 depends on the ratio of the maximum of the two components. Ghosh, whenever the ratio is small, may wish to try some other method of attack.

On the other hand, an important difficulty arises when the X_i 's are correlated. Banerjee (1980) discussed the testing of non-independence of X_i 's. In this case, an alternative attack is to find out the correlation structure.

relationship. In fact, the χ^2 test is conservative in the sense that it tends not to reject the null hypothesis, even if it is not true, and is more rational, usually, than t -tests when the distributions are identical, even for the extremes of the noncentral t -distribution, and are more powerful than χ^2 for noncentrality. It is rejected at a small significance level, α , if we may wish to relax (modify) the so-called P^* -condition.

In this chapter we define the posterior- P^* -condition in Section 2.1, the Bayes- P^* -selection procedure $\hat{P}_{\pi}^*$ and $\hat{P}_{\pi}^{\dagger}$ are presented in Section 2.2 and Section 2.3 separately, and their properties are discussed in Section 2.4. In Section 2.5 we discuss their applications to selection problems. In Section 2.6 procedure $\hat{P}_{\pi}^{\dagger}$ is compared with the data-driven procedure. An application for the selection of selection procedures $\pi_1, \dots, \pi_k$ of the normal distributions $N(\mu_j, \sigma_j^2)$, $j = 1, \dots, k$, is given in Section 2.7. In Section 2.8 we discuss their application to the selection problems for Pareto distributions and their connection with their relation to the selection of gamma distributions. In Section 2.9 we deal with comparisons of the performance of selection procedures $\hat{P}_{\pi}^*$ and $\hat{P}_{\pi}^{\dagger}$ with $\hat{P}_{\text{app}}^*$ and $\hat{P}_{\text{app}}^{\dagger}$, where $\hat{P}_{\text{app}}^*$ and $\hat{P}_{\text{app}}^{\dagger}$ are the data-driven selection procedures based on sample means and sample moments, respectively, Gupta (1986, 1987) and Gupta and Singh (1980). The comparison is done on Monte Carlo simulation. Estimates of the relative efficiency, expressed in terms of the expected one-sided size and the probability of correct selection in Section 2.9, where the relative efficiency is defined as the ratio of the expected size of the selected procedure to the expected size of the data-driven procedure, are given in Table 2.1. The relative efficiency of the data-driven procedure is also given in Table 2.1. The relative efficiency of the data-driven procedure is also given in Table 2.1.

2.2. Definitions of the Posterior- P^* Condition and the Generalized Bayes- P^* Procedure

Let $\pi_1, \dots, \pi_k$ be the ordered unknown π_j 's. Suppose that the prior distribution π_j for $j = 1, \dots, k$ is given by $p_j(x)$ for $x \in \mathcal{X}$. Then the posterior probability of a correct selection under selection procedure δ , given $x \in \mathcal{X}$, is

$$P(\text{CS} | \delta, x) = \sum_{j=1}^k \pi_j(x) p_j(x)$$

where

$$p_j(x) = P(\pi_j \text{ is the best} | X = x).$$

It is clear

$$\sum_{j=1}^k p_j(x) = 1.$$

Definition 2.2.1. Given a number $P^* (0 < P^* < 1)$ and the prior π_j , we say a selection procedure δ satisfies the posterior- P^* condition if

$$P(\text{CS} | \delta, x) \geq P^* \quad \text{for all } x.$$

Remark 2.2.1. The posterior- P^* condition is based on the prior distribution π_j and is different from the usual so-called P^* -condition.

Definition 2.2.2. The loss function L_{π_j} is defined by $L_{\pi_j}(x)$

where $|A|$ is the size (number) of populations associated with the select set A . The loss function L_{π_j} is defined by $L_{\pi_j}(x) = \frac{1}{|A|} \sum_{i \in A} L_{\pi_j}(x, i)$

which is the number of the mis-best populations selected by selection δ .

Note that the indicator function

$$I_{\pi_j}(x) = \begin{cases} 1 & \text{if } x = \pi_j \\ 0 & \text{otherwise} \end{cases}$$

Definition 2.2.3. Given a number $P^*(\frac{1}{k} - P^* + 1)$ and the prior π , we define the class $\mathcal{A}_{NR}(\cdot, P^*)$ as follows.

$$\mathcal{A}_{NR}(\cdot, P^*) = \{ \cdot \}_{\pi} \text{ is any non-randomized rule which} \\ \text{satisfies the posterior-} P^* \text{ condition (1).}$$

For the sake of convenience sometimes we will use $\mathcal{A}_{NR}$ instead of $\mathcal{A}_{NR}(\cdot, P^*)$.

Definition 2.2.4. Given a number $P^*(\frac{1}{k} - P^* + 1)$, a prior π , and the loss function L , a selection procedure $\delta \in \mathcal{A}_{NR}(\cdot, P^*)$ is called a non-randomized Bayes- P^* procedure (rule) if δ is a Bayes rule in the class $\mathcal{A}_{NR}(\cdot, P^*)$.

Let $p_{[1]}(x) \leq \dots \leq p_{[k]}(x)$ be the ordered $p_i(x)$'s and $\pi_{[i]}$ be the population associated with $p_{[i]}(x)$, $i = 1, \dots, k$, then a subset selection rule δ is completely specified by $\gamma_{(1)}, \dots, \gamma_{(k)}$ where $\gamma_{(i)}$ is defined by

$$\gamma_{(i)}(x) = P(\pi_{[i]} \text{ is selected} | x, \pi = \pi_{[i]}), \quad i = 1, \dots, k.$$

Next, we propose a non-randomized selection rule which belongs to $\mathcal{A}_{NR}(\cdot, P^*)$.

Definition 2.2.5. Given a number $P^*(\frac{1}{k} - P^* + 1)$, $\lambda > 0$, and a prior distribution π , the selection rule δ_{λ}^* is defined by $\gamma_{(i)}^*(x) = \gamma_{(i)}(x)$, where

$$\gamma_{(i)}^*(x) = \begin{cases} 1, & \text{if } \pi_{[i]} = \pi(x) \\ 0, & \text{otherwise} \end{cases}$$

and $j(x)$ is the maximum integer such that

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

Lemma 2.2.1. $\frac{B}{NR} \in \mathcal{A}_{NR}$.

Proof. follows from the definition of $\frac{B}{NR}$.

Theorem 2.2.1. Given a number $P^* (\frac{1}{k} \leq P^* \leq 1)$, the prior π , and the function L_1 , the selection procedure $\frac{B}{NR}$ is a non-randomized selection rule.

Proof. It is sufficient to show that the selection procedure $\frac{B}{NR}$ selects the smallest posterior risk in the class $\mathcal{A}_{NR}(\cdot, P^*)$. Given the observation $X = x$. Let the posterior risk of $\frac{B}{NR}(\cdot, P^*)$ be $r(x, \frac{B}{NR})$. Then

$$r(x, \frac{B}{NR}) = k - j + 1$$

and

$$\sum_{i=j+1}^k p_{[i]}(x) \leq P^*$$

for some j , $1 \leq j \leq k$.

Hence the inequality

$$r(x, \cdot) \leq r(x, \frac{B}{NR})$$

is not true for any $\cdot \in \mathcal{A}_{NR}(\cdot, P^*)$, $\cdot \neq \frac{B}{NR}$. Therefore, the result follows.

Theorem 2.2.2. Theorem 2.2.1 also holds when we replace the loss L_1 by L_2 .

Proof. Under the loss function L_2 , the posterior risk of the selection procedure $\delta \in \Delta_{NR}(i, P^*)$ is

$$r(x, \delta) = \sum_{i=1}^k \pi(i)(x) [1 - p_{[i]}(x)], \text{ given } \delta \in \Delta_{NR}(i, P^*)$$

By Theorem 2.2.1, we have

$$\sum_{i=1}^k \pi_{NR(i)}^B(x) = \sum_{i=1}^k \pi(i)(x).$$

If

$$\sum_{i=1}^k \pi_{NR(i)}^B(x) = \sum_{i=1}^k \pi(i)(x)$$

then by definition of π_{NR}^B , we have

$$\sum_{i=1}^k \pi_{NR(i)}^B(x) p_{[i]}(x) = \sum_{i=1}^k \pi(i)(x) p_{[i]}(x).$$

On the other hand, if

$$\sum_{i=1}^k \pi_{NR(i)}^B(x) < \sum_{i=1}^k \pi(i)(x)$$

then

$$\begin{aligned} \sum_{i=1}^k \pi_{NR(i)}^B(x) &= \sum_{i=1}^k \pi(i)(x) - 1 \\ &= \sum_{i=1}^k \pi(i)(x) (1 - p_{[1]}(x)). \end{aligned}$$

Therefore, we have

$$r(x, \pi_{NR}^B) = r(x, \delta) \text{ for all } \delta \in \Delta_{NR}(i, P^*).$$

Corollary 2.2.1. For a given prior π and the loss function

$L = L_1 L_1 + L_2 L_2$ where $L_1, L_2 \geq 0$ then π_{NR}^B is a non-randomized Bayes

rule with the loss function l for all $x \in X, y \in Y$.

Proof. For the given prior π and the loss function l , the Bayes risk of any procedure $\delta \in \mathcal{N}_B$ is, given $X = x, Y = y$,

$$\begin{aligned} r(x, y) &= E_{\pi} \left[\sum_{i=1}^k \pi_i(x) l_i(y, \delta_i(x)) + \sum_{j=1}^k \pi_j(y) l_j(\delta_j(x), y) \right] \\ &= E_{\pi} \left[\sum_{i=1}^k \pi_i(x) l_i(x, \delta_i(x)) + \sum_{j=1}^k \pi_j(y) l_j(\delta_j(y), y) \right] \\ &= r(x, \frac{b}{N_R}) \end{aligned}$$

with the loss function l .

Hence $\delta_{\frac{b}{N_R}}^B$ is a Bayes- P^* rule with the loss function l for all $x \in X$.

2.3. Proposed Bayes- P^* Procedure $\delta_{\frac{b}{N_R}}^B$ in general

Suppose we are interested in the randomized subset selection rule and we would like to find such a rule which also satisfies the uniform P^* condition and has the minimum risk with the loss function l for all $x \in X$ and the prior distribution π .

Definition 2.3.1. Given a prior π , we define a family of randomized selection rules, in which all rules satisfy the uniform P^* condition, for any given observation $X = x$, that are

$$\delta_i(x, P^*) = \pi_i(x) \delta_i(x, P^*) + \pi_j(y) \delta_j(x, P^*) \quad \text{for all } x \in X.$$

Definition 2.3.2. Given a number $\frac{b}{N_R} \in (0, 1)$, a prior π and the loss function l , a selection procedure $\delta \in \mathcal{N}_B$ is called a Bayes- P^* rule if this procedure is a Bayes rule in the class $\mathcal{N}_B(P^*)$.

For the sake of convenience, sometimes we will use λ instead of $\lambda(\cdot, P^*)$.

Definition 2.3.2. We define a subset selection procedure δ^B as follows.

Given a prior π and observation $X = x$, δ^B is defined by

$$[\delta^B(1), \delta^B(2), \dots, \delta^B(k)]^T,$$

where

$$\delta^B(k)(x) = 1$$

and

$$\delta^B(i)(x) = 1, \quad \text{if} \quad \begin{cases} \sum_{i=1}^k p_{[i]}(x) = P^*, & i \neq k \\ \sum_{i=1}^k p_{[i]}(x) < P^* \\ \sum_{i=1}^k p_{[i]}(x) > P^* \end{cases}$$

$$= \begin{cases} \frac{\sum_{i=1}^k p_{[i]}(x) - P^*}{k - i + 1} p_{[i]}(x) < P^* \\ \frac{\sum_{i=1}^k p_{[i]}(x) - P^*}{k - i + 1} p_{[i]}(x) > P^* \\ \frac{P^* - \sum_{i=1}^k p_{[i]}(x)}{k - i + 1} p_{[i]}(x) < P^* \end{cases}$$

$$0, \quad \text{otherwise.}$$

Example. If $k = 3$, $P^* = .90$ and the posterior probabilities are $p_{[1]}(x) = .3$, $p_{[2]}(x) = .50$, $p_{[3]}(x) = .15$, then we select the population π_1 corresponding to $p_{[1]}(x)$ with probability 1. And we select π_2 with probability 0, which is given by

$$[\delta^B(1), \delta^B(2), \delta^B(3)]^T =$$

$$[1, 0, 0]^T.$$

By Definition 2.3.2 we have

$$B_{\pi}(x) = B_{\pi}(x, \pi) = \sum_{i=1}^k \pi_i(x) p_{\pi_i}(x) = P_{\pi}^{\star} \text{ if } P_{\pi} \in \mathcal{P}^{\star}(x),$$

hence we have the following lemma.

Lemma 2.3.2. $\sum_{i=1}^k \pi_i(x, P^{\star}) \in \mathcal{P}^{\star}(x, P^{\star})$.

Definition 2.3.3. We define a subclass $\mathcal{L}^{\star}(x, P^{\star})$ of $\mathcal{P}^{\star}(x, P^{\star})$ by

$$\mathcal{L}^{\star}(x, P^{\star}) = \{ \pi \in \mathcal{P}^{\star}(x, P^{\star}) : \pi_i(x) = \pi_{[i]}(x) \text{ for all } i \},$$

where $\pi_{[1]}(x) \geq \pi_{[2]}(x) \geq \dots \geq \pi_{[k]}(x)$ are the ordered $\pi_i(x)$'s.

By the definition of $\mathcal{L}^{\star}(x, P^{\star})$ we have the following lemma.

Lemma 2.3.4. $\sum_{i=1}^k \pi_i \in \mathcal{L}^{\star}(x, P^{\star})$.

Lemma 2.3.5. For all $\pi \in \mathcal{L}^{\star}(x, P^{\star})$ there exists $\pi' \in \mathcal{L}^{\star}(x, P^{\star})$ such that $\pi(x, \pi') = \pi(x, \pi)$ w.r.t the loss function L_1 , for all x .

Theorem 2.3.6. Selection procedure $\pi^{\star}$ is a Bayes- $P^{\star}$ procedure in $\mathcal{L}^{\star}(x, P^{\star})$ w.r.t the loss function L_1 .

Proof. Given the observation $x = x_1$ and any selection procedure $\pi \in \mathcal{L}^{\star}(x, P^{\star})$ s.t. $\pi_{[1]}(x) < \pi_{[2]}^{\star}(x)$, then $\pi_{[1]}(x) < 1$ and $\pi_{[1]}^{\star}(x) = 1$ for $i \neq k$. Hence we have

$$\pi_{[1]}(x, \pi) < \pi_{[1]}^{\star}(x, \pi^{\star}) \text{ for all } x \in \mathcal{X}.$$

If $\pi_{[1]}^{\star}(x) = 1$, then we have $\pi_{[1]}(x) = 1$ and $\pi_{[1]}^{\star}(x) = 1$ for all $x \in \mathcal{X}$.

We will now show that for any x ,

$$\pi_{[1]}(x, \pi) \leq \pi_{[1]}^{\star}(x, \pi^{\star})$$

implies

$$P(CS_{i+1}, x) = P(CS_{i+1}^B, x) = c^*,$$

that is

$$\sum_{i=1}^k \rho_i(x) = \sum_{i=1}^k \rho_i^B(x)$$

implies

$$\sum_{i=1}^k \rho_i(x) (1) \rho_{i+1}(x) = \sum_{i=1}^k \rho_i^B(x) (1) \rho_{i+1}^B(x).$$

For any ϵ , $1 > \epsilon > \sum_{i=1}^k \rho_i(x)$, we have $\epsilon = a + \epsilon_1$ where a is a positive integer and $0 < \epsilon_1 < 1$.

It is easy to see that the maximum posterior probability of correct selection of procedure δ with $\sum_{i=1}^k \rho_i(x) = \epsilon$ is

$$\sum_{i=k-a+1}^k \rho_i(x) + \epsilon P_{[k-a]}(x),$$

and it is less than $\sum_{i=1}^k \rho_i^B(x) P_{[1]}(x)$, since $\sum_{i=1}^k \rho_i^B(x) = \epsilon > \epsilon$, therefore δ^B is Bayes- P^* procedure in $\mathcal{A} \cap \mathcal{A}^*$.

Lemma 3.3. Given the loss function l_{δ} for all $\delta \in \mathcal{A}^*$, there exists $\delta^B \in \mathcal{A} \cap \mathcal{A}^*$ such that, given $\theta = \theta_1, \dots, \theta_k, \theta_{k+1} = \theta_1$, where $\theta_1, \dots, \theta_k > 0$, the posterior risk with the prior θ and the loss function l_{δ} is

Proof. Given $\theta = \theta_1, \dots, \theta_k, \theta_{k+1} = \theta_1$, then

$$\sum_{i=1}^k \theta_i \rho_i(x) = \sum_{i=1}^k \theta_i \rho_i^B(x) = \sum_{i=1}^k \theta_i (1) \rho_{i+1}^B(x) = \sum_{i=1}^k \theta_i \rho_i(x).$$

$$= \sum_{i=1}^k (1-p_{1i})^{t(x)} p_{1i}^{t(x)} \\ = p^{\star}$$

hence $\delta^{\star} \in \omega^{\star}(\cdot, p^{\star})$.

Now,

$$\begin{aligned} \delta^{\star}(x, t(x)) &= \sum_{i=1}^k \delta_i^{\star}(x) (1-p_{1i})^{t(x)} p_{1i}^{t(x)} \\ &= \sum_{i=1}^k \delta_i^{\star}(x) (1-p_{1i})^{t(x)} \\ &= \sum_{i=1}^k \delta_i^{\star}(x) (1-p_{1i})^{t(x)} \\ &= \delta^{\star}(x, t(x)). \end{aligned}$$

Hence the proof is complete.

Theorem 2.3.2. Given the prior π and the observation $t = x$, the procedure δ^b is a Bayes- $p^{\star}$ procedure in the class $\omega(\cdot, p^{\star})$ when the loss function is l_{π} .

Proof. By Lemma 2.2.4, it is sufficient to show that

$$\delta^b(x, t(x)) = \min_{\delta \in \omega(\cdot, p^{\star})} \int \delta(x, t(x)) d\pi(x)$$

where

$$\begin{aligned} \delta^b(x, t(x)) &= \sum_{i=1}^k \delta_i^b(x) (1-p_{1i})^{t(x)} p_{1i}^{t(x)} \\ &= \sum_{i=1}^k \frac{(1-p_{1i})^{t(x)} p_{1i}^{t(x)}}{(1-p_{1i})^{t(x)} + (1-p_{1i})^{t(x)} p_{1i}^{t(x)}} \\ &= \sum_{i=1}^k \frac{(1-p_{1i})^{t(x)} p_{1i}^{t(x)}}{(1-p_{1i})^{t(x)} + (1-p_{1i})^{t(x)} p_{1i}^{t(x)}} \end{aligned}$$

□

$$i_0 = \frac{rH}{1+i_k} + i - \frac{P}{1+i}(x) = 0$$

$$A_1 = -i - \frac{1}{1+i}(x) + \frac{B}{1+i}(x) = 0$$

$$A_2 = -i - \frac{1}{1+i}(x) + \frac{P}{1+i}(x) = 0, x$$

then $A_1 = A_2 = 0$.

And we have

$$a_1 = \max A_1 = i_0 = 1, \quad \text{if} \quad -\frac{B}{1+i_0}(x) = -\frac{1}{1+i_0}(x)$$

$$i_0 = 1, \quad \text{if} \quad -\frac{B}{1+i_0}(x) = -\frac{1}{1+i_0}(x)$$

$$a_2 = \min A_2 = i_0 = 1, \quad \text{if} \quad -\frac{P}{1+i_0}(x) = -\frac{1}{1+i_0}(x)$$

$$i_0 = 1, \quad \text{if} \quad -\frac{B}{1+i_0}(x) = -\frac{1}{1+i_0}(x)$$

hence $a_1 = a_2 = 1$.

Therefore, we have

$$f(x, x') = f(x, x'') = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right) = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right)$$

$$= \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right) = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right)$$

$$= \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right) = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right)$$

$$= \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right) = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right)$$

$$= \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right) = \frac{1}{1+i} \left(\frac{1}{1+i}(x) + \frac{P}{1+i}(x) - 1 \right)$$

$$\begin{aligned}
& \leq \frac{1}{\epsilon} \sum_{i \in A_1} \left(\frac{1}{\epsilon} \frac{1}{\epsilon} (x_i)^{\frac{1}{\epsilon}} \frac{1}{\epsilon} (x_i)^{\frac{1}{\epsilon}} \right) \frac{1}{\epsilon} \sum_{i \in A_1} \left(\frac{1}{\epsilon} \frac{1}{\epsilon} (x_i)^{\frac{1}{\epsilon}} \right)^{\frac{1}{\epsilon}} \\
& \leq (1 + \epsilon) \sum_{i \in A_1} (x_i)^{\frac{1}{\epsilon}}.
\end{aligned}$$

$$\sum_{i \in I} \left(\frac{1}{\epsilon} \frac{1}{\epsilon} (x_i)^{\frac{1}{\epsilon}} \right) \leq \frac{B}{\epsilon} (x)^{\frac{1}{\epsilon}} (1 + \epsilon) \sum_{i \in A_1} (x_i)^{\frac{1}{\epsilon}}.$$

and by Theorem 2.3.1,

Corollary 2.3.1. Procedure $\bar{\pi}^B$ is a $\text{Save-}d^B$ rule in $(\mathcal{X}, \mathcal{Y}, \bar{L})$ with the loss function $\bar{L} = c_1 l_1 + c_2 l_2$, $c_1, c_2 > 0$.

Proof. Similar to Corollary 2.2.1, hence it is omitted.

2.4. Properties of $\bar{\pi}^B$ and $\bar{\pi}_{RP}^B$

In this section we discuss some properties of selective procedures $\bar{\pi}^B$ and $\bar{\pi}_{RP}^B$. The following definition of the ordering of distributions was introduced by Lehmann (1960) and further discussed by Lehmann and Alam (1973).

Definition 2.4.1. A subset $K \subseteq \mathcal{P}^k$ is a system of ordered distributions if $\lambda_i \leq \lambda_j$ for all $i = 1, \dots, k$, implies $\lambda_i \leq \lambda_j$.

Definition 2.4.2. A family of probability distributions $\{\lambda_i\}_{i=1}^k$ is said to be $\mathcal{P}^k$ -free if the procedure $\bar{\pi}^B$ is invariant to the project $\pi: \mathcal{X} \rightarrow \mathcal{Y}$, $\pi(x_i) = x_i$ for all $i = 1, \dots, k$, implies

$$\bar{\pi}^B(\lambda_i) = \bar{\pi}^B(\lambda_j)$$

for all distributions λ_i, λ_j .

Let $f(\cdot, \cdot, \cdot)$ be the p.d.f. of population π_1 . Let $\pi_1(\cdot)$ be the given prior where π_1^i 's are mutually independent. Suppose for $x = x_1$ we have absolutely continuous posterior c.d.f. $h(\cdot, x)$. Hence we can write the p.d.f. as

$$h(\cdot, x) = \prod_{j=1}^k g_j(\cdot, x) = \prod_{j=1}^k q_j(\cdot, x_j).$$

Let $q_i(\cdot, x_i)$ be the posterior c.d.f. associated with $\pi_1^i = \pi_1^i(x_i)$, $i = 1, \dots, k$.

Definition 2.4.3. The absolutely continuous posterior c.d.f. $q_i(\cdot, x_i)$, $i = 1, \dots, k$, have the generalized (strictly) stochastic increasing property (G(SIP)) if for any i, j , $i \neq j$, $z = k$, $x_i(\cdot) \leq x_j(\cdot)$ then

$$q_i^z(\cdot, x_i^z(\cdot)) \leq q_j^z(\cdot, x_j^z(\cdot)).$$

Note that if $q_1(\cdot, x_1) = q_2(\cdot, x_2) = h(\cdot, x)$ for all x_1, x_2 , $1 \leq x_1, x_2 \leq x$, then the GSIP is the usual SIP.

Definition 2.4.4. A selection procedure ϕ is monotone (generalized) if and only if for every $x \in \mathbb{R}^k$, $x_j \leq x_i$ implies $\phi_j(x) \leq \phi_i(x)$ for all i, j and ϕ is monotone with the exception of x above that deviation being of probability zero.

Theorem 2.4.1. If the prior π_1^i is such that we have a selection procedure with independent posterior distributions $q_i(\cdot, x_i)$, $i = 1, \dots, k$, and $q_i^z(\cdot, x_i^z(\cdot))$ is SIP, then for every $x \in \mathbb{R}^k$, $x_j \leq x_i$ implies

$$\phi_j(x) \leq \phi_i(x)$$

and hence a selection procedure $\phi^k = \prod_{i=1}^k \phi_i^k$ is monotone (generalized).

Proof.

$$\begin{aligned}
 p_1(x) &= P(\cdot_1 \in \cdot_1 | \cdot_1 \in x) = \int_{\mathcal{F}_1} G_1(t, x) d\mu_1(t, x_1) \\
 &= \int_{\mathcal{F}_1} G_1(t, x) M_1(t, x_1) d\mu_1(t, x_1) \\
 &= \int_{\mathcal{F}_1} G_1(t, x_1) B_1(t, x_1) d\mu_1(t, x_1) \\
 &= \int_{\mathcal{F}_1} \frac{d}{dt} \left[\int_{\mathcal{F}_1} G_1(t, x_1) B_1(t, x_1) dt \right] \\
 &= \int_{\mathcal{F}_1} \frac{d}{dt} \left[\int_{\mathcal{F}_1} G_1(t, x_1) B_1(t, x_1) dt \right] \\
 &= \int_{\mathcal{F}_1} G_1(t, x_1) d\mu_1(t, x_1) = p_1(x).
 \end{aligned}$$

Since

$$B_1(x) = \frac{b_1(x)}{p_1(x)} = \frac{b_1(x)}{p_1(x)} = \frac{b_1(x)}{p_1(x)}$$

and

$$B_{NR_1}(x) = \frac{b_{NR_1}(x)}{p_{NR_1}(x)} = \frac{b_{NR_1}(x)}{p_{NR_1}(x)} = \frac{b_{NR_1}(x)}{p_{NR_1}(x)}$$

Therefore, the procedures $\frac{P}{NR_1}$ and $\frac{B}{NR_1}$ are ordered too.

Under IVP assumptions, we can relabel the population such that $x_1, \dots, x_k$, hence we have $p_1(x) = p_{NR_1}(x)$ and $\frac{P}{NR_1}(x) = \frac{B}{NR_1}(x)$. In fact, target (1970) defined the "first" property of a selection rate as follows: definition 2.4.5. A selection rate r is said to be "first" if and only if $r(x) = r(x_1)$ for all $x \in \mathcal{F}_1$, $x_1 \in \mathcal{F}_1$ where $x_1 = x_1$ and $x_1 = x_1$.

we add $\frac{P}{NR_1}$ and $\frac{B}{NR_1}$ almost everywhere if $r(x) = r(x_1)$ and $r(x) = r(x_1)$ for all $x \in \mathcal{F}_1$, $x_1 \in \mathcal{F}_1$ where $x_1 = x_1$ and $x_1 = x_1$.

Definition 2.4.6. A selection rule δ is called translation invariant if for all $x \in \mathbb{R}^k$, and for all $c \in \mathbb{R}$

$$\delta_j(x_1 + c, \dots, x_k + c) = \delta_j(x_1, \dots, x_k) \quad (j = 1, \dots, r).$$

Definition 2.4.7. A selection rule δ is called scale-invariant if for all $x \in \mathbb{R}^k$, and for all $c > 0$

$$\delta_j(x \cdot c, \dots, x_k \cdot c) = \delta_j(x_1, \dots, x_k) \quad (j = 1, \dots, r).$$

Theorem 2.4.2. If the posterior distributions $\pi_j(x_j | y_j) \quad (j = 1, \dots, r)$ have the dSIP property, then both selection procedures δ_{LW}^j and δ_{AB}^j are translation-invariant.

Proof. It is sufficient to show that

$$p_j(x) = p_j(x') \quad \text{whenever } x_1' = x_1 \\ \text{and } x_2' = x_2 + \lambda(x_1 - x_2) \quad \lambda \in \mathbb{R} \setminus \{0\}$$

for any fixed j ,

$$\begin{aligned} p_j(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_j(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_j(t) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_j(t) e^{it\lambda(x_1 - x_2)} dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx'} \phi_j(t) dt \\ &= p_j(x') \end{aligned}$$

and for all $\lambda \in \mathbb{R} \setminus \{0\}$, since the proof follows.

AD-A100 948

PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS

F/6 12/1

SOME RESULTS IN THE THEORY OF SUBSET SELECTION PROCEDURES.(U)

AUG 80 H YANG

N00014-75-C-0455

UNCLASSIFIED

MMS-80-22

2 OF 2

AD-A
100 948

END

DATE

7-81

DTIC

Definition 2.4.8. Given a number $P^* (\frac{1}{k} < P^* < 1)$, $X = x$ and a prior τ ; for any selection procedure $\psi \in \mathcal{A}(\tau, P^*)$ the ratio of the posterior probability $P(CS|\psi, \underline{x})$ and the posterior expected selected size $E(S|\psi, \underline{x})$ is called the posterior-efficiency of ψ and is denoted by $EFF(\psi|\underline{x})$.

$$EFF(\psi|\underline{x}) = \frac{P(CS|\psi, \underline{x})}{E(S|\psi, \underline{x})}.$$

If $EFF(\psi|\underline{x}) \geq EFF(\psi'|\underline{x})$ for all $\psi' \in \mathcal{A}$ and all x , then the selection procedure ψ is called "posterior most efficient" (PME) selection procedure in $\mathcal{A}(\tau, P^*)$.

Theorem 2.4.3. The non-randomized posterior- P^* selection procedure ψ_{NR}^B is the PME selection procedure in $\mathcal{A}_{NR}(\tau, P^*) = \mathcal{A}_{NR}$, given τ, P^* .

Proof. By Lemma 2.2.1, for all $\psi \in \mathcal{A}_{NR}$

$$\exists \psi' \in \mathcal{A}_{NR}' \supset EFF(\psi'|\underline{x}) \geq EFF(\psi|\underline{x}) \quad \forall x,$$

hence it is sufficient to show that:

Given $\tau(\underline{g}), P^*, \underline{x}$, $EFF(\psi_{NR}^B|\underline{x}) \geq EFF(\psi'|\underline{x})$ for all $\psi' \in \mathcal{A}_{NR}'(\tau, P^*)$.

We know that, in $\mathcal{A}_{NR}(\tau, P^*)$ hence in $\mathcal{A}_{NR}'(\tau, P^*)$, ψ_{NR}^B always has minimum selected size, i.e. $\forall \underline{x}, \sum_{i=1}^k \psi_{NR}^B(i)(\underline{x}) + c = \sum_{i=1}^k \psi'_i(\underline{x})$ for some integer c , $0 \leq c \leq k-1$.

$$\begin{aligned} EFF(\psi'|\underline{x}) &= \frac{\sum_{i=1}^k \psi'_i(\underline{x}) p_{[i]}(\underline{x})}{\sum_{i=1}^k \psi'_i(\underline{x})} \\ &< \frac{\sum_{i=1}^k \psi_{NR}^B(i)(\underline{x}) p_{[i]}(\underline{x}) + p_{[k-s-c+1]}(\underline{x}) + \dots + p_{[k-s]}(\underline{x})}{\sum_{i=1}^k \psi_{NR}^B(i) + c} \end{aligned}$$

if $\psi_{NR}^B(x) = (0, \dots, \underbrace{0, 1, \dots, 1}_{s \text{ terms}})$.

$$\begin{aligned} \text{EFF}(\psi' | x) &\leq \frac{\sum_{i=k-s+1}^k \psi_{NR(i)}^B(x) p_{[i]}(x) + c p_{[k-s]}(x)}{\sum_{i=1}^k \psi_{NR(i)}^B(x) + c} \\ &\leq \frac{\sum_{i=1}^k \psi_{NR(i)}^B(x) p_{[i]}(x)}{\sum_{i=1}^k \psi_{NR(i)}^B(x)} \\ &= \text{EFF}(\psi_{NR}^B | x) . \end{aligned}$$

The last inequality is obtained by

$$\begin{aligned} \sum_{i=1}^k \psi_{NR(i)}^B(x) p_{[i]}(x) &= \sum_{i=k-s+1}^k \psi_{NR(i)}^B(x) p_{[i]}(x) \\ &\geq \left(\sum_{i=k-s+1}^k \psi_{NR(i)}^B(x) \right) p_{[k-s]}(x) . \end{aligned}$$

Theorem 2.4.4. The randomized selection procedure ψ^B is the PME procedure in $\mathcal{A}(\tau, P^*) = \mathcal{A}$ for given τ, P^* .

Proof. It suffices to show that, given $\tau, P^*, \underline{x} = x$,

$$\text{EFF}(\psi^B | x) = \text{EFF}(\psi' | x), \quad \forall \psi' \in \mathcal{A}^* .$$

Suppose $\psi^B(x) = (0, \dots, \underbrace{v, 1, \dots, 1}_{s \text{ terms}})$ $0 \leq v < 1, 1 \leq s < k-1$.

By theorem 2.3.1 there exists $c > 0$ such that

$$\sum_{i=1}^k \psi_{NR(i)}^B(x) + c = \sum_{i=1}^k \psi'_{NR(i)}(x) .$$

If $0 \leq c < 1$, then

$$\begin{aligned}
 \text{EFF}(\psi' | \underline{x}) &= \frac{\sum_{i=1}^k \psi'(i)(\underline{x}) p_{[i]}(\underline{x})}{\sum_{i=1}^k \psi'(i)(\underline{x})} \\
 &\leq \frac{\sum_{i=1}^k \psi^B(i)(\underline{x}) p_{[i]}(\underline{x}) + c p_{[k-s]}(\underline{x})}{\sum_{i=1}^k \psi^B(i)(\underline{x}) + c} \\
 &\leq \frac{\sum_{i=1}^k \psi^B(i)(\underline{x}) p_{[i]}(\underline{x})}{\sum_{i=1}^k \psi^B(i)(\underline{x})} \\
 &= \text{EFF}(\psi^B | \underline{x}) .
 \end{aligned}$$

If $1 \leq c = v' + t + (1-v)$, $t \geq 0$ integer, $0 < v' < 1$ then

$$\begin{aligned}
 \sum_{i=1}^k \psi'(i)(\underline{x}) p_{[i]}(\underline{x}) &= \sum_{i=k-s+1}^k \psi^B(i)(\underline{x}) p_{[i]}(\underline{x}) + v' p_{[k-s-t+1]}(\underline{x}) \\
 &\quad + p_{[k-s-t]}(\underline{x}) + \dots + (1-v) p_{[k-s]}(\underline{x}) \\
 &\leq \sum_{i=1}^k \psi^B(i)(\underline{x}) p_{[i]}(\underline{x}) + c p_{[k-s]}(\underline{x})
 \end{aligned}$$

hence by the same argument as above we have

$$\text{EFF}(\psi' | \underline{x}) \leq \text{EFF}(\psi^B | \underline{x}) .$$

Since $\underline{x}$ is arbitrary, the result holds for all $\underline{x}$.

2.5. Applications to Normal Model

Suppose we have k populations $\pi_1, \dots, \pi_k$; population π_i has distribution $N(\mu_i, \sigma_i^2)$, where σ_i^2 's are known and μ_i 's are unknown. Assume that we

have independent observations $X_{i1}, \dots, X_{in_i}$, $i = 1, \dots, k$. Let

$$X_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \text{ and let } \underline{X} = (X_1, \dots, X_k).$$

Suppose we are interested in selecting a subset containing the best (the population having the largest mean) under the posterior- P^* condition, wrt some prior $\tau = \tau(\underline{\mu})$. Then to find a Bayes- P^* selection procedure is equivalent, in some sense, to finding $p_i(\underline{x})$, which is the posterior probability of the event $\{\mu_i \text{ is the best}\}$, given observations $\underline{X} = \underline{x}$, wrt a given prior τ , for all $i = 1, \dots, k$.

Case I. Assume that we have a common sample size n and a common known variance σ^2 .

Ia. Suppose we have no prior information about the unknown parameters, and use the "non-informative" (Box and Tiao (1973)) or "locally uniform" prior $p(\mu_i) \propto c$ for each population.

The posterior density function g_i of μ_i , given x_i is the normal density with mean x_i and variance σ^2/n , i.e.,

$$g_i(\mu_i | x) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n(\mu_i - x_i)^2}{2\sigma^2}\right).$$

Hence

$$p_{[i]}(x) = P(\mu_{(i)} = \mu_{[k]} | \underline{X} = \underline{x}) \\ = \int_{-\infty}^{\infty} \frac{1}{n} \sum_{j \neq i} \exp\left(-\frac{\sqrt{n}}{\sigma} (x_{[i]} - x_{[j]})\right) d\left(\frac{t}{\sigma}\right) \\ i = 1, \dots, k.$$

Here $\mu_{(i)}$ is the quantity corresponding to the i^{th} largest observation $X_{[i]}$.

Ib. If μ_j 's are independent and have the identical prior distribution $N(\theta_0, \sigma_0^2)$ and $X_i | \mu_i \sim N(\mu_i, \sigma_1^2/n)$, then it is well known that the posterior density function g_i of μ_i , given $\underline{x} = \underline{x}$ is

$$g_i(\mu_i | \underline{x}) \sim N(\bar{\theta}_{x_i}, \xi^2) \text{ with SIP property}$$

where

$$\bar{\theta}_{x_i} = \xi^2(\sigma_0^{-2}\theta_0 + n\sigma_1^{-2}x_i)$$

$$\xi^2 = (\sigma_0^{-2} + n\sigma_1^{-2})^{-1}.$$

Hence

$$p_{[i]}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi(t + \xi n \sigma_1^{-2}(x_{[i]} - x_{[j]})) d\psi(t).$$

The last expression for $p_{[i]}(\underline{x})$ is the same as that for the non-informative prior whenever $\sigma_0 \rightarrow \infty$.

Since $p_{[i]}(\underline{x}) = p_{[i]}(\underline{x} + \underline{b})$ and since the normal distribution has the strictly SIP, it follows that ψ^B and ψ_{NR}^B are "just" a.e. and translation-invariant in both case Ia and Ib.

Case II. Variance σ_i 's are known but σ_i 's and n_i 's are not all equal.

IIa. Using the non-informative prior $p(\mu_i) = c, i=1, \dots, k$, we have

$$p_{(i)}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi\left(t \frac{v(i)}{v(j)} + \frac{x_{[i]} - x_{[j]}}{v(j)}\right) d\phi(t)$$

where $v(i) = \frac{\sigma(i)}{n(i)}$ $i = 1, \dots, k$. $p(i)$, $\psi(i)$ and $n(i)$ are corresponding to $x_{[i]}$ and we have the following theorem.

Theorem 2.5.1. $p_{(i)}(\underline{x})$ is non-decreasing in i , i.e., $p_{(i)}(\underline{x}) = p_{[i]}(\underline{x})$.

Remark 2.5.1. From the above formula of $p_{(i)}(\underline{x})$, it is easy to see, increasing the sample size of the non-best populations will increase the probability that the best population to be selected, however, before doing this, we don't know which one is the best one.

In this case ψ^B and ψ_{NR}^B are "just" a.e. and translation-invariant.

Case III. Assume that priors are independent but not identical normal distributions, namely, $\mu_i \sim N(\theta_i, \sigma_{0i}^2)$, where θ_i 's are not all equal; if the conditional distribution of X_i , given μ_i , is $N(\mu_i, \frac{\sigma_{1i}^2}{n_i})$, then the posterior density of μ_i , given $X_i = x_i$ is $g_i(\mu_i | x_i)$, which is the probability density function of normal distribution $N(\bar{\theta}_{x_i}, \epsilon_i^2)$ where

$$\begin{aligned}\bar{\theta}_{x_i} &= \epsilon_i^2 (\sigma_{0i}^{-2} \theta_i + n_i \sigma_{1i}^{-2} x_i) \\ \epsilon_i^2 &= (\sigma_{0i}^{-2} + \sigma_{1i}^{-2} n_i)^{-1}.\end{aligned}$$

Hence we have

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi\left[t \frac{\epsilon_j}{\epsilon_j} + \frac{1}{\epsilon_j} (\bar{\theta}_{x_i} - \bar{\theta}_{x_j})\right] d\phi(t).$$

If $\sigma_{0i} = \sigma_0$, $\sigma_{1i} = \sigma_1$ and $n_i = n$, $i = 1, \dots, k$, then

$$\epsilon_i = \epsilon = (\sigma_0^{-2} + \sigma_1^{-2} n)^{-1} \quad i = 1, \dots, k$$

and

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi\left[t + \left(\frac{\sigma_1^{-2} n}{\sigma_0^{-2} + \sigma_1^{-2} n} + \frac{n(x_i - x_j)}{\sigma_1^2}\right)\right] d\phi(t).$$

Case IV. The General Normal Model

Here we consider a more general prior. Suppose we have k populations, common sample size n for each population, and common known variance $\sigma^2 > 0$. The observation can reduce to $X = (X_1, \dots, X_k)$ where

$$X_i = \sum_{j=1}^n X_{ij}/n, \text{ by sufficiency.}$$

The "Normal Model" is defined as follows:

$$(a) \quad X|\underline{\mu} \sim N(\underline{\mu}, qI), \quad q = \frac{\sigma^2}{n}$$

where I is the $k \times k$ identity matrix.

So the X 's are (conditionally) independent when $\underline{\mu}$ is given.

$$(b) \quad \underline{\mu} \sim N(\theta_0 \underline{1}, \gamma I + tU)$$

where $\theta_0 \in \mathbb{R}$, $\gamma > 0$, $t > -\frac{\gamma}{k}$,

$$\underline{1} = (1, \dots, 1) \quad \text{and} \quad U = \underline{1}' \underline{1}.$$

Here $\gamma > 0$ and $t > -\frac{\gamma}{k}$ are necessary and sufficient for $\gamma I + tU$ to be positive definite. This model was chosen by Chernoff and Yahav (1977) ($t > 0$), Gupta and Hsu (1978) and Miescke (1979).

By (a) and (b) we get the posterior distribution of $\underline{\mu}$, given $X = \underline{x}$, and the distribution of X as follows:

$$\underline{\mu}|\underline{x} \sim N(\underline{\eta}, aI + bU)$$

where

$$\underline{\eta} = \gamma(q+\gamma)^{-1} \underline{x} + qt((q+\gamma)(q+\gamma+kt))^{-1} \underline{x}' U + q(q+\gamma+kt)^{-1} \underline{1}$$

$$a = \gamma q(q + \gamma)^{-1}$$

$$b = q^2 t(q + \gamma)^{-1} (q + \gamma + kt)^{-1}$$

$$\underline{X} \sim N(\underline{m}, (q + \gamma)I + tU)$$

Lemma 2.5.1. Let $\underline{Y} \sim N(\underline{\mu} + \rho \underline{1}, aI + bU)$ with $\underline{\mu} \in \mathbb{R}^k$, $\rho \in \mathbb{R}$, $a > 0$ and $b > -a/k$. Then there exists a random vector $\underline{Z} \sim N(\underline{\mu}, aI)$ such that $h(\underline{Y}) = h(\underline{Z})$ everywhere for every translation-invariant $h: \mathbb{R}^k \rightarrow \mathbb{R}^k$.

Proof. (See Miescke (1979)).

With this lemma, it is easy to get

$$\begin{aligned} p_i(\underline{x}) &= P(\mu_i = \mu[k] | \underline{x}) \\ &= \int I_{\{\mu_i = \mu[k]\}} d\phi_{\left(\left(\frac{\gamma}{q+\gamma}\right)\underline{x}, \frac{\gamma q}{q+\gamma} I\right)}(\underline{\mu}) \end{aligned}$$

where $\phi_{(\underline{\mu}, V)}$ is the normal distribution with mean $\underline{\mu}$ and variance-covariance matrix V .

We can rewrite $p_i(\underline{x})$ as

$$p_i(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi\left(t + \left(\frac{\gamma}{q(q+\gamma)}\right)^{\frac{1}{2}}(x_i - x_j)\right) d\lambda(t).$$

Let $\gamma = \sigma_0^2$, $q = \sigma^2/n$, we have

$$p_i(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j \neq i} \phi\left(t + \left(\frac{\sigma_0^2}{\frac{\sigma^2}{n} \left(\frac{\sigma^2}{n} + \sigma_0^2\right)}\right)^{\frac{1}{2}}(x_i - x_j)\right) d\lambda(t).$$

The above expression is exactly the same as that of the independent prior Case I, Ib.

Case V. Under normal assumption as before, but suppose μ_i 's are unknown and that neither μ_i 's nor n_i 's are all equal.

Suppose we have no prior information about (μ, σ) , for each individual population π_i assign prior $p(\mu_i, \sigma_i) \propto \sigma_i^{-1}$ then we have (See Box and Tiao (1973)) that the posterior density of μ_i , given $X_i = x_i = (x_{i1}, \dots, x_{in_i})$ is

$$p(\mu_i | x_i) = \frac{(s_i/\sqrt{n_i})^{-1}}{B(\frac{1}{2}v_i, \frac{1}{2})\sqrt{v_i}} \left[1 + \frac{n_i(\mu_i - \bar{x}_i)^2}{v_i s_i^2} \right]^{-\frac{1}{2}(v_i+1)}$$

where s_i^2 is the sample variance, $B(\cdot, \cdot)$ is a Beta function and $v_i = n_i - 1$. Hence

$$p(t_i = \frac{\mu_i - \bar{x}_i}{s/\sqrt{n_i}} | x_i) = \frac{1}{B(\frac{1}{2}v_i, \frac{1}{2})\sqrt{v_i}} \left(1 + \frac{t_i^2}{v_i} \right)^{-\frac{1}{2}(v_i+1)},$$

which is the density of the student's t distribution with $v_i = (n_i - 1)$ degrees of freedom.

Using this result we can write the formula of $p_i(x)$ by

$$\begin{aligned} p_i(x) &= P(\mu_i > \mu_j, \quad \forall j \neq i | x) \\ &= \int \prod_{j \neq i} T_{v_j} \left(t \frac{s_j/\sqrt{n_j}}{s_i/\sqrt{n_i}} + \frac{\bar{x}_i - \bar{x}_j}{s_j/\sqrt{n_j}} \right) dT_{v_i}(t) \end{aligned}$$

where $v_i = n_i - 1, \quad i = 1, \dots, k$

$$v_i s_i^2 = \sum_{r=1}^{n_i} (x_{ir} - \bar{x}_i)^2$$

T_{v_i} is the c.d.f. of t distribution with v_i degrees of freedom. When v_i 's are large, t distribution approaches normal distribution, hence, for large $n_i, i = 1, \dots, k$, we can replace T by Φ .

Case VI. Suppose we are interested in finding a subset which contains the population with the smallest variance; i.e., we define the best population as the one with the smallest variance, and suppose that we have no prior information about σ . In this case, it is reasonable to assume that

$$p(\mu, \sigma) \propto \sigma^{-1}, \text{ if } \mu \text{ is unknown}$$

$$p(\sigma) \propto \sigma^{-1}, \text{ if } \mu \text{ is known.}$$

Let

$$v_j = n_j, v_j S_j^2 = \sum_{r=1}^k (X_{jr} - \bar{x}_j)^2 \text{ if } \mu \text{ is known}$$

$$v_j = n_j - 1, v_j S_j^2 = \sum_{r=1}^k (X_{jr} - \bar{x}_j)^2 \text{ if } \mu \text{ is unknown, } n_j \geq 1 \quad j = 1, 2, \dots, k$$

$$S^2 = (S_1^2, \dots, S_k^2), X = (X_{11}, \dots, X_{1n_1}, \dots, X_{kn_k})$$

and Y_{ij} be the random variable with c.d.f. $\chi^2_{v_j}$ which is the χ^2 distribution with v_j degrees of freedom.

Then for either case (μ known or unknown), we have

$$\begin{aligned} p_i(x) &= P(\sigma_i^2 = \sigma_{[1]}^2 | X = x) \\ &= P(\sigma_i^2 \leq \sigma_j^2, \forall j \neq i | X = x) \\ &= P\left(\frac{v_j S_j^2}{\sigma_j^2} \leq \frac{v_i S_i^2}{\sigma_i^2} \left(\frac{v_j S_j^2}{v_i S_i^2}\right), \forall j \neq i | S^2 = s^2\right) \\ &= P\left(Y_{ij} \leq Y_{vi} \left(\frac{v_j S_j^2}{v_i S_i^2}\right), \forall j \neq i | S^2 = s^2\right) \\ &= \int_0^\infty \dots \int_0^\infty \mathbb{I}_{\left\{u_j \leq u_i \left(\frac{v_j S_j^2}{v_i S_i^2}\right)\right\}} dF_{v_j}^2(u_j) \dots dF_{v_i}^2(u_i) \end{aligned}$$

$$= \int_0^\infty \prod_{j \neq i} \left(\frac{s_j^2}{s_i^2} \right) d\psi_i^2(u) \quad \text{if } n_1 = \dots = n_k = 1.$$

With these $p_1(x), \dots, p_k(x)$ we can apply Bayes- P^* rules ψ^B and ψ_{NR}^B easily.

Lemma 2.5.2. In Case VI, ψ^B and ψ_{NR}^B are just P^* a.e. and (scale) translation invariant.

* Here the definition of the "just" property for a selection rule is

$$\psi_i(s_i^2) \leq \psi_i(s_i'^2) \text{ if } s_i^2 \geq s_i'^2, s_j^2 \leq s_j'^2, \forall j \neq i.$$

2.6. Comparison of Selection Rules ψ^B and ψ^M in the Normal Location

Parameter Case

We have k normal populations with a common known variance σ^2 and common sample size n . For this case Gupta (1956) proposed and studied the procedure ψ^M .

ψ^M : Select μ_i iff $X_i > X_{[k]} - d \frac{\sigma}{\sqrt{n}}$ $i = 1, \dots, k$ where $d = d(P^*, k)$ is to be determined by

$$\inf_{\theta \in \Omega} P(CS | \psi^M) = P^*$$

and Ω is the parameter space.

We will show that $\psi^M \in \Delta_{NR}(\cdot, P^*)$ where π is the locally uniform prior distribution. For fixed P^* and k , let d be determined by

$$\int_{-\infty}^{\infty} t^{k-1} (t + d) d\pi(t) = P^*. \quad (2.6.1)$$

Let

$\mathcal{X} = \{\text{all possible observed values}\} = \mathbb{R}^k$

$$\mathcal{X}_1 = \{x \in \mathcal{X} | x_{[k]} = d \frac{1}{\sqrt{n}} + x_{[1]}\}$$

$$\mathcal{X}_i = \{x \in \mathcal{X} | x_{[i-1]} + x_{[k]} = d \frac{1}{\sqrt{n}} + x_{[i]}\}, \quad 2 \leq i \leq k$$

$$\mathcal{X}_i^{(1)} = \{x \in \mathcal{X} | x_{[1]} = x_{[i-1]} + x_{[k]} = d \frac{1}{\sqrt{n}} + x_{[i]}\} \subset \mathcal{X}_i$$

$$\mathcal{X}_i^{(2)} = \{x \in \mathcal{X} | x_{[1]} = x_{[i-1]} + x_{[k]} = d \frac{1}{\sqrt{n}} + x_{[i]} = x_{[k-1]}\} \subset \mathcal{X}_i^{(1)}$$

then we have the following theorem.

Theorem 2.6.1. Given a number $P^*(\frac{1}{k} \leq P^* \leq 1)$ and locally uniform prior for each population π_i , $X = \underline{x} \in \mathcal{X}_i$, then

$$P(CS, \hat{\mu}_i^M, \underline{x} = \underline{x}) \geq q^*(i)$$

where

$$q^*(i) = \frac{k-i}{k-1} (1 - P^*) + P^*.$$

Hence

$$\hat{\mu}_i^M \in \mathcal{A}_{NR}(i, P^*).$$

Proof. It is sufficient to show that

$$\inf_{x \in \mathcal{X}_i} \sum_{j=1}^k p_{i,j}(x) = q^*(i) = \frac{k-i}{k-1} (1 - P^*) + P^*.$$

Since $x \in \mathcal{X}_i$,

$$P(CS, \hat{\mu}_i^M, x) = \inf_{x \in \mathcal{X}_i} P(CS, \hat{\mu}_i^M, x)$$

$$= \inf_{x \in \mathcal{X}_i} \sum_{j=1}^k p_{i,j}(x).$$

Without loss of generality we can assume $\frac{\sigma}{\sqrt{n}} = 1$.

Since

$$\sum_{i=1}^k p_{[i]}(\underline{x}) = 1 \quad \forall \underline{x} \in \mathcal{X}, \text{ and } \forall i \leq k$$

$p_{[i]}(\underline{x})$ is nonincreasing for all $x_{[j]}$, $j \leq i-1$, we have

$$\begin{aligned} \inf_{\underline{x} \in \mathcal{X}_i} \sum_{i=1}^k p_{[i]}(\underline{x}) &= \inf_{\underline{x} \in \mathcal{X}_i} (1) - \sum_{i=1}^k p_{[i]}(\underline{x}) \\ &= 1 - \sup_{\underline{x} \in \mathcal{X}_i} \sum_{i=1}^{i-1} p_{[i]}(\underline{x}) \\ &= 1 - \sup_{\underline{x} \in \mathcal{X}_i} (1) - \sum_{i=1}^{i-1} \int_{-\infty}^{\infty} \frac{1}{j!} \phi(t + x_{[i]} - x_{[j]}) d\phi(t) \\ &= 1 - \sup_{\underline{x} \in \mathcal{X}_i} (1) - \sum_{i=1}^{i-1} \int_{-\infty}^{\infty} \frac{1}{j!} \phi(t + x_{[i]} - x_{[j]}) \\ &\quad \cdot \sum_{\substack{j \neq i \\ j < i}} \phi(t + x_{[i]} - x_{[j]}) d\phi(t) \\ &= 1 - \sup_{\underline{x} \in \mathcal{X}_i} (1) - \sum_{i=1}^{i-1} \int_{-\infty}^{\infty} \frac{1}{j!} \phi(t + x_{[i]} - x_{[j]}) \\ &\quad \cdot \phi^{i-2}(t) d\phi(t) \\ &= 1 - \sum_{i=1}^{i-1} \int_{-\infty}^{\infty} \phi(t-d) \phi^{k-2}(t) d\phi(t) \quad (2.6.7) \end{aligned}$$

$$\begin{aligned}
&= 1 - (i - 1) \int_{-\infty}^{\infty} \phi(t-d) \phi^{k-2}(t) d\phi(t) \\
&= (k - i) \int_{-\infty}^{\infty} \phi^{k-2}(t) \phi(t - d) d\phi(t) \\
&\quad + \int_{-\infty}^{\infty} \phi^{k-1}(t + d) d\phi(t) \quad (2.6.3)
\end{aligned}$$

The supremum of (2.6.3) occurs when $\underline{x} \in \mathcal{A}_i^{(2)}$. The last equality follows from the identity

$$\begin{aligned}
&(k - 1) \int \phi^{k-2}(t) \phi(t - d) d\phi(t) \\
&= 1 - \int \phi^{k-1}(t + d) d\phi(t),
\end{aligned}$$

which can be shown by the integration by parts. By (2.6.1), the second term of (2.6.3) equals P^* ; then use the integration by parts to the first term of (2.6.2), we get

$$\begin{aligned}
\inf_{\underline{x} \in \mathcal{A}_i} \sum_{j=i}^k P_{[j]}(\underline{x}) &= \frac{k-i}{k-1} [1 - P^*] + P^* \quad (2.6.4) \\
&= q^*(i).
\end{aligned}$$

Remark 2.6.1. If the procedure ψ^M selects $\pi_{(k)}$ only, i.e. $\underline{X} = \underline{x} \in \mathcal{A}_k$, then by Theorem 2.6.1 we have $p_{[k]}(\underline{x}) \geq P^*$ so that ψ^B or ψ_{NR}^B selects $\pi_{(k)}$ only. But the converse is not necessarily true.

Remark 2.6.2. For the case $k = 2$, $\psi_{NR}^B = \psi^M$ a.e. For any given $\underline{X} = \underline{x}$: if $\underline{x} \in \mathcal{A}_2$, then $p_{[2]}(\underline{x}) \geq P^*$, hence ψ^M and ψ_{NR}^B select the population $\pi_{(2)}$ associated $\underline{x}_{[2]}$. If $\underline{x} \in \mathcal{A}_1$, and $\underline{x}_{[2]} = d \frac{\sigma_1}{\sigma_2} < \underline{x}_{[1]}$ then ψ^M and

ψ_{NR}^B select both populations π_1 and π_2 . Since

$$P(X_{[2]} - d \frac{\sigma}{\sqrt{n}} = X_{[1]}) = 0,$$

we have $\psi_{NR}^B = \psi^M$ a.e. .

Remark 2.6.3. The above Theorem and Remark 2.6.1 gives us a lower bound on the value of $\sum_{i=1}^k P_{[\psi]}(\underline{x})$, over all $\underline{x} \in \mathcal{L}_j$. The exact value of the difference of the selected sizes between ψ^M and ψ^B depends on the observations.

2.7. Applications to Select $\max_{1 \leq i \leq k} \mu_i$, $\theta_i = \frac{\mu_i - a}{\sigma_i}$ for Normal

Distribution $N(\mu_i, \sigma_i^2)$, $i = 1, \dots, k$

Let $\pi_1, \dots, \pi_k$ be k independent normal populations with mean μ_i and variance σ_i^2 , both μ_i and σ_i are unknown. For the goal of finding a random subset which contains the population with maximum $\theta_i = \frac{\mu_i - a}{\sigma_i}$ for some given constant a , we assume that apriori (μ_i, σ_i) , $i = 1, \dots, k$ are independent. Suppose we have n_i independent observations $X_{i1}, \dots, X_{in_i}$ from π_i , and let $\bar{X}_i$ be their sample mean, $i = 1, \dots, k$.

Let $Y_1, \dots, Y_n$ be i.i.d. $\sim N(\mu, \sigma^2)$. If no prior information is available to (μ, σ) , we could assign a locally uniform prior $p(\mu, \sigma) \propto \sigma^{-1}$ to (μ, σ) , (see Box and Tiao (1973)). And the posterior joint distribution of $\mu' = \mu - a$ and σ , given observations $Y = y = (y_1, \dots, y_n)$ is given by

$$P(\mu', \sigma | \underline{y}) = k \sigma^{-(n+1)} \exp \left\{ -\frac{1}{2\sigma^2} [vs^2 + n(y' - \mu')^2] \right\}$$

where

$$\left. \begin{aligned} y' &= y - a, \quad y = \frac{1}{n} \sum_{i=1}^n y_i \\ vs^2 &= \sum_{i=1}^n (y_i - y)^2, \quad v = n - 1 \\ k &= \sqrt{\frac{n}{2\pi}} \left[\frac{1}{2} \Gamma\left(\frac{v}{2}\right) \right]^{-1} \left(\frac{vs}{2} \right)^{\frac{v}{2}} \end{aligned} \right\} \quad (2.7.1)$$

Let $t = \sqrt{n} (\mu - a)/\sigma$, with (2.7.1) the posterior distribution of t , given $Y = y$ is

$$p(t | Y = y) = p(t | t) \\ = \{2^{\frac{v}{2}-1} \Gamma\left(\frac{v}{2}\right)\}^{-1} \left(\frac{vs}{2} \right)^{\frac{v}{2}} \exp \left\{ -\frac{1}{2} \left(\frac{vs}{v+t^2} \right) \right\} \Gamma(v) I_{v-1} \left(\frac{t \sqrt{vs}}{(v+t^2)^{1/2}} \right)$$

where

$$t = \sqrt{n}(y - a)/s, \quad v = n - 1$$

$$I_v(x) = \int_0^x (\sqrt{2\pi} \Gamma(v))^{-1} u^v \exp \left\{ -\frac{1}{2} (u + x)^2 \right\} du.$$

Now, let $p(\alpha_i, \sigma_i) = \sigma_i^{-1}$ be the assigned locally uniform prior to (α_i, σ_i) . Then let $x = (x_{11}, \dots, x_{1n_1}, \dots, x_{kr_k})$, we have

$$p_i(x) = P(\alpha_i = \alpha_{[k]} | x) \\ = P\left(\alpha_i = \max_{1 \leq j \leq k} (\alpha_j) | x \right)$$

$$\begin{aligned}
&= P\left(\sqrt{\frac{n_j}{n_i}} \varepsilon_i > \varepsilon_j \quad \forall j \neq i \mid t\right) \\
&= \int \prod_{j \neq i} G_{\varepsilon_j} \left(\sqrt{\frac{n_j}{n_i}} z \mid t \right) d G_{\varepsilon_i}(z \mid t) \quad (2.7.1) \\
&= \int \prod_{j \neq i} G_{\varepsilon_j}(z \mid t) d G_{\varepsilon_i}(z \mid t) \quad \text{if } n_1 = \dots = n_k = n,
\end{aligned}$$

where G_{ε_i} is the posterior c.d.f. of ε_i , given x or t .

By (2.7.2), the Bayes-P* procedure is completely specified.

If the prior distribution for (μ, σ) is the conjugate distribution (see Raiffa and Schlaifer (1960)), then

$$\begin{aligned}
p(\mu, \sigma) &\propto \exp \left\{ -\frac{1}{2\sigma^2} n'(\mu - m')^2 \right\} \cdot \exp \left\{ -\frac{1}{2\sigma^2} (v') \right\} \\
&= p(\mu \mid \sigma) p(\sigma)
\end{aligned}$$

that is

$$p(\mu \mid \sigma) \sim N(m', \sigma^2/n'), \quad n' > 0$$

$$p(\sigma) \sim \frac{v' v'^{1/2}}{\sigma^2} \exp \left\{ -\frac{v'}{\sigma^2} \right\}, \quad v', v'^{1/2} > 0.$$

Let

$$x' = \frac{nx + n'm'}{n + n'}, \quad x \text{ is the sample mean.}$$

$$u^2 = \{(n-1)s^2 + v'v' + [(nn')/(n+n')](x-m')^2\} / *$$

$$v^* = (n-1) + v' + 1$$

$$t^* = (n+n')^{1/2}(\mu - a)/\sigma$$

$$t^* = (n+n')^{1/2}(x' - a)/u,$$

the posterior distribution of ε^* , given x is $p(\varepsilon^* \mid x) = p(\varepsilon^* \mid t^*)$ which has the same form as $p(\varepsilon \mid t)$, but replace ε, t, n by ε^*, t^*, n^* .

Thus, for the conjugate prior case, we get

$$p_i(x) = p(\lambda_i \in [\lambda]) \cdot x \\ = \int_{\lambda \neq 1} G_{\lambda} \cdot \int_{\lambda \neq 1} z \cdot t^* dG_{\lambda}(z \cdot t^*) \quad (2.7.2) \\ \int_{\lambda \neq 1} G_{\lambda}(z \cdot t^*) dG_{\lambda}(z \cdot t^*) \quad \text{if } n_1 = 0$$

where G_{λ} is the posterior c.d.f. of λ given x or t .

Note that (2.7.3) has the same form as (2.7.2), but replace λ , t by λ^* , t^* .

2.8. Applications to Poisson Distributions and Poisson Processes

2.8.1. Poisson Distributions Case

Suppose that $\pi_1, \dots, \pi_k$ are k independent Poisson populations, where the independent observations $X_{i1}, \dots, X_{in_i}$ from π_i have the Poisson density with parameter λ_i ; denoted by $P(\cdot | \lambda_i)$, $i = 1, \dots, k$.

Let $Y_1, \dots, Y_n$ be i.i.d. with $p(\cdot | \lambda)$. If we use non-informative prior $p(\cdot) \propto \lambda^{-1/2}$ (Box and Tiao (1973)), then given $Y = y = (y_1, \dots, y_k)$, we have the posterior density as follows:

$$p(\cdot | y) = c \cdot n y^{-\frac{1}{2}} \exp(-n \cdot)$$

where

$$y = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{and} \quad c = n^{ny + \frac{1}{2}} [\sum_{i=1}^n y_i + \frac{1}{2}]^{-1}$$

We see that $2n_i | y \sim \chi^2_{2ny+1}$, the chi-square distribution with $2ny+1$ degrees of freedom. Hence by using non-informative prior

$p(\lambda_i) \propto \lambda_i^{-1/2}$ for each population π_i , we have

$$\begin{aligned} p_i(x) &= p(\lambda_i = \lambda[k] | x) \\ &= \int_{-\infty}^{\infty} \prod_{j \neq i} \chi^2_{n_j} \left(z \frac{n_j}{n_i} \right) d \chi^2_{n_i} (z) \end{aligned}$$

where

$$\lambda_i = 2n_i x_i + 1, \quad x_i = \sum_{j=1}^{n_i} x_{ij} / n_i.$$

If $n_1 = \dots = n_k$, then

$$p_i(x) = \int_0^{\infty} \prod_{j \neq i} \chi^2_{n_j} (z) d \chi^2_{n_i} (z).$$

With $p_i(x)$, $i = 1, \dots, k$, we can apply Bayes-P* selection rules.

J^B and J^B_{NR} easily to select a subset which contains the population with the largest parameter λ . On the other hand, if we are interested in selecting the population with the smallest parameter λ , then

$$\begin{aligned} p_i(x) &= \int_0^{\infty} \prod_{j \neq i} [1 - \chi^2_{n_j} (z \frac{n_j}{n_i})] d \chi^2_{n_i} (z) \\ &= \int_0^{\infty} \prod_{j \neq i} [1 - \chi^2_{n_j} (z)] d \chi^2_{n_i} (z) \quad \text{if } n_1 = \dots = n_k. \end{aligned}$$

In this case, the simulation results for selection procedures J^B and J^B_{NR} are tabulated on Table VII.

2.8.2. Poisson Processes Case

Suppose we have k independent Poisson processes

$\{X^{(1)}(t)\}, \dots, \{X^{(k)}(t)\}$ with expected arrival times equal to

$\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_k}$, respectively. Hence for the processes $\{X^{(i)}(t)\}$, the

probability that there are m_i arrivals until time t_i is

$$p(X^{(i)}(t_i) = m_i | \lambda_i, t_i) = \frac{(t_i \lambda_i)^{m_i}}{m_i!} e^{-t_i \lambda_i}$$

$$i = 1, 2, \dots, k, \quad m_i = 0, 1, 2, \dots$$

If there exists no prior information, then we use the non-informative

prior $p(\lambda_i) \propto \lambda_i^{-1/2}$ for all processes. Therefore, we get the posterior density function of λ_i , given (m_i, t_i) as follows:

$$\begin{aligned} p(\lambda_i | X^{(i)}(t_i) = m_i, t_i) &= p(\lambda_i | m_i, t_i) \\ &= \frac{(t_i \lambda_i)^{m_i + \frac{1}{2} - 1}}{\Gamma(m_i + \frac{1}{2})} t_i e^{-t_i \lambda_i}. \end{aligned}$$

Thus $2t_i \lambda_i$ has χ^2 distribution with $2m_i + 1$ degrees of freedom, given the number m_i of arrivals before time t_i .

Let $m = (m_1, \dots, m_k)$ and $t = (t_1, \dots, t_k)$, then it can be shown that

the Poisson process $\{X^{(i)}(t)\}$ has the maximum parameter (or minimum mean waiting time), given (m, t) is

$$p_i(\underline{m}, t) = \int_0^\infty \prod_{j \neq i} x_{2m_j+1}^2(y \frac{t_j}{t_i}) d x_{2m_i+1}^2(y) \quad i = 1, \dots, k. \quad (2.3.1)$$

Here we list two special cases which are of interest.

- (a) Observations of all processes are obtained in a common time interval $[s_i, t + s_i]$. Since Poisson process is stationary, we can assume that $s_i = 0$, and $t_1 = \dots = t_k = t$. In this case

$$p_i(\underline{m}, t) = \int_0^\infty \prod_{j \neq i} x_{2m_j+1}^2(y) d x_{2m_i+1}^2(y)$$

which is independent of t .

- (b) All m_i 's are equal, i.e. we fix m first, then get observations t

Hence

$$p_i(m, t) = \int_0^\infty \prod_{j \neq i} x_{2m+1}^2(y \frac{t_j}{t_i}) d x_{2m+1}^2(y).$$

There is an alternative way to approach the cases (a) and (b).

Let T_i be the waiting time of the n th arrival in the i th process,

then T_i has a gamma distribution with density given by

$$p(t) = \frac{\lambda_i^{m_i}}{\Gamma(m_i)} (\lambda_i t)^{m_i-1} e^{-\lambda_i t} \quad t > 0.$$

If we have only non-informative prior $p(\lambda) = \lambda^{-1/2}$; then, given m_i and

t_i , $2 t_i \lambda_i$ has posterior distribution $x_{2m_i+1}^2$, therefore the formula

of $p_i(\underline{m}, t)$ we get here is exactly the same as before.

Remark 2.8.1. Under non-informative prior, in comparing the subset selection problem in k Poisson distributions with the problem in k Poisson processes, it is easily seen that Poisson distributions model is a special case of Poisson processes model, namely, $t_i = n_i$ where n_i denotes the sample size of the i th Poisson population.

2.8.3. Relation Between Selection from Poisson Processes and Selection from Populations with Gamma or Exponential Distribution

Suppose we have k independent populations, the i th population having the gamma distribution with parameters $\alpha = m_i$ (known), $\lambda = 1/\tau_i$ (unknown). Since the random variable T_i , the waiting time until m_i arrivals in a Poisson process with parameter λ_i , has a gamma distribution with parameters $\alpha = m_i$, $\lambda = 1/\tau_i$. If the m_i 's are given and if the goals for both selection problems are the same, namely, to select a subset containing the population (process) with the largest parameter λ , then it is easily seen that these are identical problems. Note that in the selection problem of Poisson processes, m_i 's might not be the preassigned values but are given random observations whenever t_i 's are preassigned values. In this case, the selection problem of Poisson processes is different from that of the gamma distributions.

If the process associated with the minimum parameter λ (or the maximum waiting time) is the best, then the posterior probability of process $X^{(i)}(t)$ to be the best is analogous to the one obtained

before with the modifications that the integrand function

$$\prod_{j \neq i} \exp\left(-\frac{1}{2} \left(\frac{t_j}{t_i}\right)^2\right)$$

of (2.8.1) is replaced by

$$\prod_{j \neq i} \left[1 - \exp\left(-\frac{1}{2} \left(\frac{t_j}{t_i}\right)^2\right)\right].$$

2.9. Comparison of the Performance of ψ^B , ψ_{NR}^B , ψ^M and ψ^{MED}

Let π_i , $i = 1, \dots, k$ be k independent populations, where π_i has the associated c.d.f. $F(x, \theta_i) = F(x - \theta_i)$ with unknown location parameter θ_i . Let $f(x, \theta_i) = f(x - \theta_i)$ be the p.d.f. 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 $\tilde{X}_i \leq \tilde{X}_{[k]} - d$

where $\tilde{X}_i$ is the median of the $2m+1$ random observations from population π_i and $\tilde{X}_{[k]} = \max_i \tilde{X}_i$. The value d is determined by the following equation so that the P^* -condition is met.

$$\int_{-\infty}^{\infty} G(u + d)^{k-1} q(u) du = P^*$$

where

$$q(u) = \frac{(2m+1)!}{(m!)^2} [F(u)]^m [1 - F(u)]^m f(u)$$

$$G(u) = I_{F(u)}(m+1, m+1)$$

$I_y(p, q)$ is the incomplete beta function.

In this section we use Monte Carlo simulation techniques to compare the performance of selection procedures δ^B , δ_{NR}^B , δ^M and δ^{MED} in the normal means problem. Because both rules δ^M and δ^{MED} are not based on any prior information about the unknown parameters, we assume that the prior distribution π for both δ^B and δ_{NR}^B is locally uniformly distributed. Since the selection procedure δ^M satisfies both the P^* -condition and the posterior- P^* condition wrt the locally uniform priors, it makes sense to compare the Bayes- P^* procedures δ^B and δ_{NR}^B with δ^M and compare δ^M with δ^{MED} in terms of efficiency which is the ratio of the probability of a correct selection to the expected selected size. For studying the robustness of these four rules, δ^B , δ_{NR}^B , δ^M and δ^{MED} , we change the true distribution to 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). 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 1. Each time we generate five random variables with the given distribution of each population, then apply the selection procedures. The simulation process is repeated 100 times for each random variable. The relative frequency of selecting the population π_j is

used as an approximation to the probability of selecting the population π_i . The sum of relative frequency of selecting each population π_i , $i = 1, \dots, k$ is treated as an approximation of the expected selected size. The efficiency EFF of each selection procedure is approximated by the ratio of relative frequency of selecting the best one to the expected size. The simulation results indicate that in all cases we have the performance

$$\phi^B > \phi_{NR}^B > \phi^M.$$

It should be noted that in the above comparison of the performance, we restrict attention to these rules which satisfy the posterior- P^* condition. For small sample size, the efficiency of rule ϕ^M tends to be larger than ϕ^{MED} under P^* -condition.

Remark 2.9.1. The Laplace distribution has the density function

$$f(x - \mu) = \frac{1}{2} e^{-|x - \mu|} \quad x \in \mathbb{R}$$

for which the variance is 2.

The logistic distribution has the density function

$$f(x - \mu) = \frac{e^{-(x - \mu)}}{(1 + e^{-(x - \mu)})^2}$$

for which the variance $\text{Var}(X) = \frac{\pi^2}{3}$.

The gross error model we used has the density function

$$f(x - \mu) = (1 - \alpha) \phi(x - \mu) + \frac{\alpha}{4} \phi\left(\frac{x - \mu}{4}\right) \quad \alpha \in [0, 1] \quad (2.15)$$

for which ϕ is the p.d.f. of $N(0, 1)$ and the variance $\text{Var}(X) = (1 - \alpha) + \alpha \cdot 4^2 = 3.25$.

The efficiency of a selection procedure δ is defined by

$$EFF_{\delta}(\delta) = \frac{P_{\delta}(CS'_{\delta})}{E_{\delta}(S'_{\delta})}$$

where $E_{\delta}(S'_{\delta})$ is the expected selected size.

Discussion and Conclusion

For Table VIII.1 and Table VIII.2 (equal distances case) the P^* is .99 and .90 respectively, the common sample size $n = 5$, $k = 5$. If the k populations have normal distributions with the unknown parameter configuration $(\mu_1, \dots, \mu_k + (k-1)\delta)$, common variance 1. From both tables the performance based on either the efficiency or the expected selected size is

$$\frac{B}{M} > \frac{P}{NR} > \frac{M}{M}$$

if the posterior- P^* condition is considered, and

$$M > \frac{M}{M}$$

under the P^* -condition.

When the true distributions are not normal, but the logistic, the Laplace or the Gauss error model, the results are very close to the normal case, hence the two rules are robust. From Table VIII.3 all entries are larger than the corresponding ones in Table VIII.1. The P^* value expected to be smaller than P^* is smaller in the robust tables.

For Table VIII.1 and Table VIII.2 (unequal distances case) the P^* is .99 and .90 respectively, the common sample size $n = 5$, $k = 5$. If the k populations

have normal distributions with unknown parameter configurations $(\mu_1, \dots, \mu+k)$, common variance 1. From both tables the performance based on either the efficiency or the expected selected size is

$$\psi^B > \psi_{NR}^B > \psi^M$$

if the posterior- P^* condition is considered, and if $\sqrt{n} = 1$

$$\psi^M > \psi^{MED}$$

under the P^* -condition.

Note that in both equal distances and slippage cases when $\sqrt{n} = 1$, that is the population means are not very close, the procedures ψ^B and ψ_{NR}^B , wrt the locally uniform priors, always satisfy not only the posterior- P^* condition but also $P_{\underline{\mu}}(CS|\psi^B \text{ or } \psi_{NR}^B) \geq P^*$, and the expected selected size of the selection procedure ψ^B or ψ_{NR}^B is much less than the selection procedures ψ^M and ψ^{MED} . For example, in the normal equal distances case, $P^* = .99$, $k = 5$, $\sqrt{n} = 4$,

$$E(S|\psi^{MED}) - E(S|\psi_{NR}^B) = 0.382;$$

in the normal slippage case, $P^* = .99$, $k = 5$, $\sqrt{n} = 4$

$$E(S|\psi^{MED}) - E(S|\psi_{NR}^B) = 1.560.$$

TABLE III

For procedures E and NR and the parameter configurations $(.5, \dots, .5, \dots)$ of k Poisson populations, this table gives the values (based on simulation) of the probability of selecting the population with parameter $(.5i, i=1, \dots, k)$ and the expected selected size ES . The prior distribution for each population is $p(i) = \frac{1}{k}$.

$n = 10$

k	p					
	0.99		0.95		0.90	
	B	B-NR	B	B-NR	B	B-NR
2						
1	.993	1.000	.966	.990	.950	.990
2	.670	.820	.450	.670	.324	.490
ES	1.663	1.820	1.415	1.660	1.274	1.480
3						
1	.998	1.000	.990	1.000	.967	1.000
2	.741	.850	.388	.630	.321	.510
3	.205	.330	.120	.130	.065	.110
ES	1.944	2.180	1.498	1.760	1.354	1.620
4						
1	.993	1.000	.981	1.000	.966	.990
2	.745	.820	.560	.730	.259	.400
3	.369	.490	.070	.140	.067	.100
4	.075	.110	.041	.060	.024	.040
ES	2.187	2.420	1.653	1.930	1.316	1.530
5						
1	.996	1.000	.981	.990	.985	1.000
2	.737	.850	.503	.650	.367	.550
3	.355	.470	.102	.160	.133	.210
4	.067	.090	.015	.030	.013	.020
5	.000	.010	.006	.010	.000	.010
ES	2.154	2.410	1.607	1.840	1.438	1.780

TABLE VIII. 1

Efficiency (EFF) and Expected Selected Size (ES) (based on simulation) of $\hat{\mu}_B$, $\hat{\mu}_M$ and $\hat{\mu}_{MED}$ when the unknown means of the k populations are $\theta_1, \dots, \theta_k + (k-1)/2$; the common variance is 1, common sample size $n = 5$ and the prior for $\hat{\mu}_B$ and $\hat{\mu}_{NR}$ is locally uniformly distributed.

TABLE 1.1.1
K = 5, $\beta^* = .99$

	normal		logistic		Laplace		gross error	
	EFF	ES	EFF	ES	EFF	ES	EFF	ES
$\beta^* = .5$								
B	.254	3.809	.250	3.963	.259	3.774	.248	3.966
B	.238	4.110	.233	4.290	.238	4.120	.230	4.303
NR								
M	.208	4.810	.207	4.840	.210	4.720	.207	4.843
MED	.208	4.810	.202	4.940	.202	4.940	.201	4.920
$\beta^* = 1$								
B	.333	2.977	.332	3.005	.336	2.941	.329	3.032
B	.305	3.280	.304	3.290	.302	3.280	.304	3.291
NR								
M	.250	4.000	.234	4.280	.246	4.030	.233	4.293
MED	.232	4.310	.224	4.460	.217	4.600	.214	4.630
$\beta^* = 2$								
B	.541	1.847	.541	1.839	.541	1.884	.559	1.773
B	.482	2.050	.481	2.020	.498	2.010	.510	1.966
NR								
M	.417	2.400	.417	2.400	.515	2.410	.437	2.290
MED	.368	2.720	.351	2.950	.362	2.761	.362	2.720
$\beta^* = 4$								
B	.625	1.212	.655	1.164	.621	1.217	.665	1.156
B	.730	1.370	.606	1.240	.746	1.340	.803	1.250
NR								
M	.676	1.420	.694	1.440	.630	1.470	.671	1.490
MED	.575	1.750	.576	1.730	.588	1.700	.571	1.750

TABLE VIII. 2

Efficiency (EFF) and Expected Selected Size (ES) (based on simulation) of $\hat{\mu}$, $\hat{\sigma}_{NR}$, $\hat{M}$ and $\hat{MED}$ when the unknown means of the k populations are $\mu_1, \dots, \mu_k + (k-1)/2$; the common variance is 1, common sample size $n = 5$ and the prior for $\hat{\sigma}_{NR}$ and $\hat{M}$ is locally uniformly distributed.

TABLE IX. I

Efficiency (E) and expected Selected Size (ES) (based on simulation) of $\hat{\beta}$, $\hat{B}$, $\hat{M}$ and $\hat{MED}$ when the parameter means of the k populations are $\mu_1, \dots, \mu_k$; the common variance is 1, common sample size $n = 5$ and the prior for β and $\hat{B}_0$ is locally uniformly distributed.

TABLE 11.1
 $\alpha = .05$, $p^* = .99$

	S.E.	normal		logistic		Laplace		gross error	
		EFF	ES	EFF	ES	EFF	ES	EFF	ES
$\rho = .5$	.07	.217	4.192	.218	4.253	.226	4.203	.220	4.361
	.14	.214	4.529	.212	4.580	.214	4.530	.212	4.670
	.21	.201	4.970	.202	4.950	.200	4.900	.202	4.890
	.28	.203	4.930	.202	4.940	.202	4.960	.200	5.000
$\rho = .7$	.07	.204	4.101	.240	4.133	.240	4.062	.231	4.264
	.14	.220	4.400	.224	4.460	.223	4.430	.216	4.580
	.21	.207	4.900	.203	4.930	.203	4.870	.204	4.900
	.28	.204	4.990	.201	4.970	.200	4.980	.200	5.000
$\rho = .9$	.07	.200	3.590	.274	3.645	.263	3.701	.295	3.381
	.14	.205	3.970	.253	3.960	.242	4.090	.269	3.720
	.21	.214	4.600	.211	4.750	.214	4.680	.212	4.510
	.28	.209	4.800	.207	4.830	.203	4.900	.200	4.900
$\rho = .95$	.07	.213	1.95	.500	3.900	.500	1.900	.400	1.950
	.14	.240	3.10	.400	3.10	.400	3.10	.400	3.10
	.21	.200	3.10	.200	3.10	.200	3.10	.200	3.10
	.28	.200	3.10	.200	3.10	.200	3.10	.200	3.10

TABLE IX. 2

the mean, σ^2 , and selected selected size (ES) (based on simulation) of B , B , M and MED when the parameters of the population are $\mu = 0$, $\sigma^2 = 1$, the common variance is 1, common sample size $n = 10$ and the parameters B and M are locally uniformly distributed.

7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530

Case	Logistic EFF	Logistic ES	Ladance		gross error	
			EFF	ES	EFF	ES
1	1.137	1.137	.238	2.734	.256	2.853
2	1.137	1.137	.228	3.160	.240	3.290
3	1.137	1.137	.211	4.406	.212	4.490
4	1.137	1.137	.202	4.840	.204	4.900
5	1.137	1.137	.134	2.653	.302	2.671
6	1.137	1.137	.105	3.650	.275	3.733
7	1.137	1.137	.098	4.388	.219	4.520
8	1.137	1.137	.207	4.736	.209	4.790
9	1.137	1.137	.133	2.127	.500	1.220
10	1.137	1.137	.535	2.473	.441	2.220
11	1.137	1.137	.171	3.650	.170	3.270
12	1.137	1.137	.138	4.163	.149	4.010
13	1.137	1.137	.145	1.113	.146	1.143
14	1.137	1.137	.141	1.113	.135	1.113
15	1.137	1.137	.133	1.113	.133	1.113
16	1.137	1.137	.133	1.113	.133	1.113

BIBLIOGRAPHY

BIBLIOGRAPHY

- Alan, E. (1973). On a multiple decision rule. *Ann. Statist.*, **1**, 750-755.
- Andersen, E. S. (1953). On the fluctuation of sums of random variables. *Math. Scand.*, **1**, 263-275.
- Ayer, M., Brunk, H. D., Eain, G. M., Reid, W. T. and Silverman, B. W. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.*, **6**, 611-621.
- Bahadur, R. R. (1950). On the power in the theory of estimation. *Ann. Math. Statist.*, **21**, 362-379.
- Bahadur, R. R. and Robbins, H. (1950). The problem of the greater mean. *Ann. Math. Statist.*, **21**, 469-487. *Correction*, **22**, 1497, 310.
- Parlow, R. E., Bartholomew, D. J., Bremner, J. M. and Fox, R. J. (1972). *Statistical Inference under Order Restrictions*. John Wiley, New York.
- Baron, A. M. and Gupta, S. S. (1972). A class of non-terminating sequential multiple decision procedures. *Operations Research Verfahren* (Ed. Henn, Künzi and Schubert). Verlag Anton Hain, Meisenheim am Glan, Germany, pp. 11-17.
- Lechhotev, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.*, **5**, 16-30.
- Lechhotev, R. E. (1967). A sequential multiple decision procedure for selecting the best mean of several normal populations with a common unknown variance and its use with various experimental designs. *Biometrika*, **54**, 405-419.
- Lechhotev, R. E., Banerji, C. W. and Sobel, M. (1964). A two-sample multiple-hypothesis procedure for ranking means of normal populations with a common unknown variance. *Biometrika*, **51**, 129-136.
- Lechhotev, R. E., Banerji, C. W. and Sobel, M. (1965). Sequential identification and ranking procedures. *Journal of the American Statistical Association*, **60**, 131-140.

- Bechhofer, R. E. and Sobel, M. (1961). A single-sample multiple decision procedure for ranking variances of normal population (prelim. report), Abstract. *Ann. Math. Statist.*, 24, 100-101.
- Berger, R. L. (1971). Minimax, admissible and gamma-minimax multiple decision rules. Ph.D. Thesis (Mimeo. Ser. No. 490), *Ann. Math. Statist.*, Purdue Univ., West Lafayette, Indiana.
- Berger, R. L. (1979). Minimax subset selection for fixed number of subset size. *Ann. Statist.*, 7, 112-119.
- Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with application to ordered variance problem. *Ann. J. Statist.*, 7, 21-26.
- Bickel, P. J. and Yahav, J. A. (1977). On selecting a subset of populations. *Statistical Decision Theory and Related Topics* (Eds. S. S. Gupta and P. J. Moore), Academic Press, New York.
- Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, Massachusetts.
- Broström, G. (1977). An improved procedure for selection of a subset of populations. *Letter to the Editor, Ann. Math. Statist.*, 1977, 1978, *Ann. Math. and Statist.*, 1978, 1979, Sweden.
- Carroll, R. J. (1971). Asymptotically nonparametric sequential selection procedures. Mimeo. Ser. No. 944, Inst. Statist. Univ. of North Carolina, Chapel Hill.
- Chengshu, H. and Yahav, J. (1977). On subset selection under generalized a new criterion. *Statistical Decision Theory and Related Topics - II* (Eds. S. S. Gupta and P. J. Moore), Academic Press, New York, pp. 103-110.
- Deely, J. J. (1965). Multiple decision procedure: from a Bayesian Bayes approach. Ph.D. Thesis (Mimeo. Ser. No. 481), *Ann. Math. Statist.*, Purdue Univ., West Lafayette, Indiana.
- Deely, J. J. and Gupta, S. S. (1969). On the properties of subset selection procedures. *Biometrika*, 56, 37-56.
- Desu, M. M. and Sobel, M. (1968). A fixed subset size approach to a selection problem. *Biometrika*, 55, 401-410. Corrections and amendments: 63 (1976), 65.
- Desu, M. M. and Sobel, M. (1971). Nonparametric procedures for selecting fixed-size subsets. *Statistical Decision Theory and Related Topics* (Eds. S. S. Gupta and J. Yackel), Academic Press, New York, pp. 255-274.

10. Bennett, C. W. (1961). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.*, 56, 1096-1111.
11. Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd Edition. J. Wiley & Sons, New York.
12. Gibbons, J. D., Olkin, I. and Soudi, W. (1977). *Selection and Ranking Populations*. Wiley, New York.
13. Guadanesikan, M. (1966). Some selection and ranking procedures for multivariate normal population. Ph.D. Thesis. Inst. of Statist., Purdue Univ., West Lafayette, Indiana.
14. Guadanesikan, M. and Gupta, S. S. (1970). Selection Procedures for multivariate normal distributions in terms of measures of dispersion. *Technometrics*, 12, 105-117.
15. Good, P. J. and Robin, R. (1971). On selecting a subset certainly the best population - A Bayesian approach. *Ann. Statist.*, 1, 689-693.
16. Gupta, S. S. (1966). On a decision rule for a problem in multiple comparisons. Ph.D. Thesis (Mimeo. Ser. No. 153). Inst. of Statist., Univ. of North Carolina, Chapel Hill.
17. Gupta, S. S. (1966). On a selection and ranking procedure for k populations. *Ann. Inst. Statist. Math.*, 14, 189-216.
18. Gupta, S. S. (1968). On some multiple decision (selection) problems. *Technometrics*, 10, 225-245.
19. Gupta, S. S. (1967). On selection and ranking procedure for k populations. *Ann. Inst. Qual. Control.*, pp. 141-165.
20. Gupta, S. S. and Dasgupta, D. (1969). On r -minimax selection and ranking procedures for selecting populations close to a best one. *Ann. Inst. Statist.*, 1, 181-191.
21. Gupta, S. S. and Prasad, R. V. (1968). On subset selection procedure for Poisson populations and some applications to two sample selection problems. *Applied Statistics (Proc. R. Soc., Ser. B)*, 17, 199-204. Holland, Amsterdam, pp. 275-284.
22. Gupta, S. S. and Prasad, R. V. (1973). On some sequential and non sequential parametric sequential subset selection procedures. *Statistical Inference and Related Topics*, Vol. 2, 113-124. Marcel Dekker, New York, pp. 101-111.

- Gupta, S. S. and Huang, D. Y. (1976). Selection procedures for the means and variances of normal populations - unequal sample size case. *Sankhyā Ser. B*, 38, 112-128.
- Gupta, S. S. and Kim, W. C. (1980). d -minimax and minimax decision rules for comparison of treatments with a control. To appear in *Recent Developments in Statistical Inference and Data Analysis*, North Holland Publishing Company, 1980, 56-72.
- Gupta, S. S. and McDonald, C. C. (1972). On some classes of selection procedures based on ranks. *Nonparametric Techniques in Statistical Inference* (Ed. M. L. Puri). Cambridge University Press, London, pp. 491-511.
- Gupta, S. S., and Miescke, R. J. (1978a). On subset selection procedures for ranking means of three normal distributions. *Mem. Ser. No. 78-19*. Dept. of Statist., Purdue Univ., West Lafayette, Indiana. To appear in *Sankhyā Ser. A*.
- Gupta, S. S. and Miescke, R. J. (1978b). On the d -minimax configurations in certain two-stage selection procedures. *Sankhyā Ser. No. 79-6*. Dept. of Statist., Purdue Univ., West Lafayette, Indiana.
- Gupta, S. S. and Nagel, K. (1971). On some contributions to the theory of decision theory. *Statistical Decision Theory and Related Topics* (Eds. S. S. Gupta and J. Yackel). Academic, New York, pp. 1-10.
- Gupta, S. S., Nagel, K. and Panchapakesan, S. (1973). On the d -minimax statistics from equally correlated normal random variables. *Biometrika*, 60, 403-413.
- Gupta, S. S. and Panchapakesan, S. (1973). On a class of subset selection procedures. *Ann. Math. Statist.*, 43, 914-921.
- Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Problems: 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 means for selection of the largest loss for parameters. *Journal of Statistical Theory, Meth.*, 49(12), 1973-1993.
- Gupta, S. S. and Sobel, M. (1961). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.*, 32, 235-244.
- Gupta, S. S. and Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. *Biometrika*, 49, 149-156.
- Gupta, S. S. and Sobel, M. (1962b). On the smallest of d -correlated t -statistics. *Biometrika*, 49, 309-311.

Gupta, S. S. and Statler, M. S. (1971). On some selection and ranking procedures with applications to multivariate populations. *Trans. in Probability and Statistics*, Eds. R. C. Bose et. al., Wiley Interscience of North Carolina Press, Chapel Hill, pp. 27-38.

Hoch, S. (1974). A simple sequentially selective multiple test procedure. *Ann. Inst. Statist.*, 6, 45-50.

Hsiao, P. (1979). Some contributions to gamma-minimax and empirical Bayes selection procedures. Ph.D. thesis (Mimeo. Ser. No. 74-B-1). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Huang, D. Y. (1975). Some contributions to fixed sample and sequential multiple decision (selection and ranking) theory. Ph.D. thesis (Mimeo. Ser. No. 363). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Huang, W. I. (1972). Some contributions to sequential selection and ranking procedures. Ph.D. thesis (Mimeo. Ser. No. 298). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a new method. *Psychometrika*, 29, 115-129.

Lehmann, E. L. (1952). Testing multiparameter hypotheses. *Ann. Math. Statist.*, 23, 541-552.

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

Lehmann, E. L. (1961). Some Model I problems of selection. *Ann. Math. Statist.*, 32, 990-1012.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *Ann. Math. Statist.*, 38, 1079-1091.

McDonald, R. C. (1969). On some distribution-free ranking and selection procedures. Ph.D. thesis (Mimeo. Ser. No. 174). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Misnick, E. J. (1979). Bayesian subset selection for additive and linear loss functions. *Commun. Statist. - Theor. Meth.*, 8(12), 1205-1226.

Postollec, L. (1948). A k-sample shapage test for an extreme population. *Ann. Math. Statist.*, 19, 58-65.

Rapel, E. (1970). On subset selection rules with certain optimality properties. Ph.D. thesis (Mimeo. Ser. No. 272). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Shankar, V. (1963). Some selection rules for comparing treatments with a standard. *Comm. Statist.*, 4, 179-186.

- Naik, C. D. (1977). Some subset selection problems. *Indian Statist. Theor. Meth.*, 46(10), 955-966.
- Naik, C. D. (1978). On selection procedures based on the exact test probability. *Calcutta Statist. Assoc. Bull.*, 29, 40.
- Paulson, E. (1949). A multiple decision procedure for certain problems in the analysis of variance. *Ann. Math. Statist.*, 20, 28-41.
- Paulson, E. (1962). A sequential procedure for comparing several experimental categories with a standard or control. *Ann. Math. Statist.*, 33, 438-443.
- Paulson, E. (1963). A sequential decision procedure for testing one of k hypothesis concerning the unknown mean of a normal distribution. *Ann. Math. Statist.*, 34, 549-554.
- Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. *Ann. Math. Statist.*, 35, 174-180.
- Paulson, E. (1967). Sequential procedures for selection of the best one of several binomial populations. *Ann. Math. Statist.*, 38, 11-14.
- Rabinowitz, P. and Weiss, G. (1959). Tables of absolute and relative for numerical evaluation of integrals of the form $\int_0^1 e^{-x^2} dx$. *Math. Tables and other Aids to Comput.*, Vol. XIII, 66, 100-101.
- Rizvi, M. H. and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest standard deviation. *Ann. Math. Statist.*, 38, 705-711.
- Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problems. *Ann. Statist.*, 3, 334-349.
- Seal, E. C. (1955). On a class of decision procedures for ranking means of normal populations. *Ann. Math. Statist.*, 26, 261-270.
- Seal, E. C. (1957). An optimum decision rule for ranking means of normal populations. *Calcutta Statist. Assoc. Bull.*, 8, 141-145.
- Sobel, M. (1967). Nonparametric procedures for selecting the best one of k populations with the largest standard deviation. *Ann. Math. Statist.*, 38, 1504-1516.
- Sobel, M. and Pryett, R. L. (1967). Selecting the best one of several binomial populations. *Bell. System Tech. J.*, 46, 867-876.
- Studden, W. J. (1967). On selecting a subset of k populations containing the best one. *Ann. Math. Statist.*, 38, 1071-1080.

- Tamhane, A. C. and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. *Commun. Statist. - Theor. Meth.*, A6(11), 1003-1033.
- Tamhane, A. C. and Bechhofer, R. E. (1979). A two-stage minimax procedure with screening for selecting the largest mean (II): An improved PCS lower bound and associated tables. *Commun. Statist. - Theor. Meth.*, A8(4), 337-358.
- Tukey, J. W. (1960). A survey of sampling from contaminated distributions. *Contributions to Probability and Statistics* (Eds. L. B. Brown and others). Stanford Univ. Press, 39, 448-458.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
GPO : 1962 O 14-5601

INDEX AND REFERENCES

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

an ordering prior. Here, by an ordering prior we mean that there exists a known simple or partial order relationship among the unknown parameters of the treatments (excluding the control). Three new selection procedures are proposed and studied. These procedures do meet the usual requirement that the probability of a correct selection is greater than or equal to a pre-specified number P^* . Two of the three procedures use the isotonic regression over the sample means of the k treatments with respect to the given order. The tables which are necessary to carry out the selection procedure with this approach for the selection of unknown means of normal populations and of t populations are given. Monte Carlo comparisons on the performance of various procedures for the normal or gamma mean problem were carried out in several selected cases. The results of this study seem to indicate that the procedure based on isotonic estimators always have superior performance, even in cases where there are more than one bad population. (In comparison with the control.)

Chapter II deals with a new 'Bayes- P^* ' approach about the problem of selecting a subset which contains the 'best' of k populations. Here, by 'best' we mean the (unknown) population with the largest unknown mean. The (non-randomized) Bayes- P^* rule refers to a rule with minimum risk in the class of (non-randomized) rules which satisfy the condition that the probability of selecting the best is at least equal to P^* . Given the priors of the unknown parameters, two 'Bayes- P^* ' subset selection procedures

δ^B and $\delta^{B_{NR}}$ (randomized and non-randomized, respectively) are constructed. Functions are obtained and compared with one classical subset selection procedure δ^M . The comparisons of the performance of δ^B with δ^M and $\delta^{B_{NR}}$ in Monte Carlo studies, indicate that the procedure δ^B has a higher 'efficiency' and smaller expected size of the selected subset. The studies also indicate that δ^B is robust when the true distributions are not normal but are some other symmetric distribution, such as, the logistic, the double exponential (Laplace) and the gross error model (the contaminated distribution).

DATE
FILMED
-18-