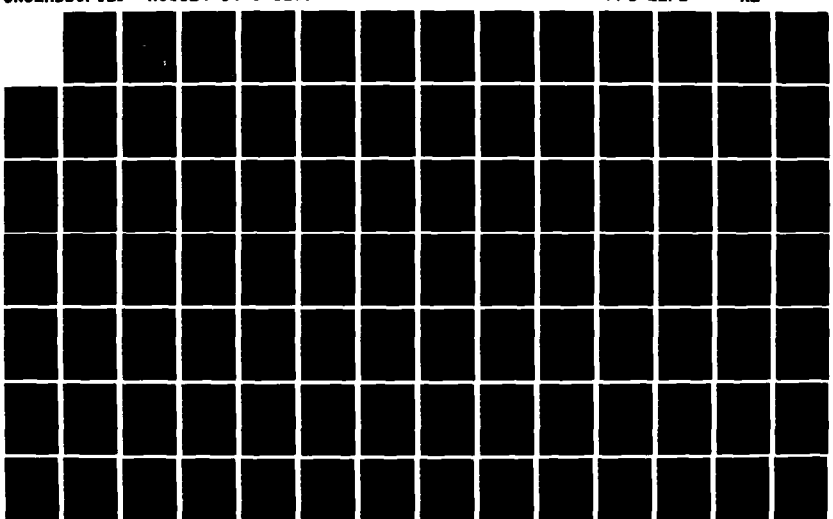
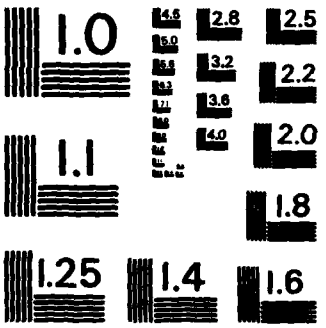
AD-A159 153 MULTIPLE DECISION PROCEDURES FOR TUKEY'S GENERALIZED 1/2
LAMBDA DISTRIBUTIONS(U) PURDUE UNIV LAFAYETTE IN DEPT
OF STATISTICS J K SOHN AUG 85 TR-85-20
UNCLASSIFIED N00014-84-C-0167 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

AD-A159 153

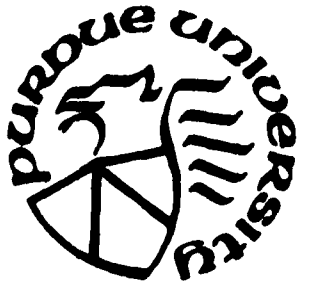
MULTIPLE DECISION PROCEDURES FOR TUKEY'S GENERALIZED LAMBDA DISTRIBUTIONS

by

Joong K. Sohn
Purdue University

Technical Report #85-20

PURDUE UNIVERSITY



DEPARTMENT OF STATISTICS

DTIC
ELECTE
SEP 1 3 1985
A

This document has been approved for public release and sale; its distribution is unlimited.

DTIC FILE COPY

85 9 11 011

MULTIPLE DECISION PROCEDURES FOR TUKEY'S GENERALIZED LAMBDA DISTRIBUTIONS

by

Joong K. Sohn
Purdue University

Technical Report #85-20

Department of Statistics
Purdue University

August 1985

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

This document has been approved for public release and sale; its distribution is unlimited.

DTIC
ELECTE
SEP 13 1985
S
D
A

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

DTIC COPY INSPECTED

INTRODUCTION

In many practical situations, the experimenter (or the decision-maker) is faced with the problem of comparing k ($\geq$ 2) populations, where each population is characterized by a real-valued parameter θ. In such situations, the classical approach is to test the hypothesis of homogeneity (equality) among the k parameters. On the other hand, the real interest (or goal) of the experimenter may be to identify the best population (defined by the experimenter in terms of, say, large value of θ) or to find a subset which contains the best population or a subset which contains all populations better than a control or standard. Thus, the test of homogeneity is inadequate in several aspects. Mosteller (1948), Paulson (1949), Bahadur (1950) and Bahadur and Robbins (1950) were among the earliest research workers to recognize this inadequacy. Since these early studies, the area of selection and ranking problems has been very active. It has seen tremendous growth over the last three and a half decades.

There have been mainly two formulations in selection and ranking problems, namely, the "indifference zone" approach and the "subset selection" approach. In the first formulation, due to Bechhofer (1954), the goal is to select one population (or a

fixed number t, $1 < t < k$) as the best population with a preassigned minimum 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 (1967), Mahamunulu (1967), Paulson (1967), Bechhofer, Kiefer and Sobel (1968), Desu and Sobel (1968, 1971), Dudewicz and Dalal (1975), Tamhane and Bechhofer (1977, 1979), among others.

In the second formulation, pioneered by Gupta (1956, 1965), the goal is to select a nonempty nontrivial subset of k populations so that the best population is included in the selected subset with a minimum guaranteed probability $P^*(\frac{1}{k} < P^* < 1)$ over the whole parameter space. The size of the selected subset is not determined in advance but is made to depend on the outcome of the experiment. Some recent contributions in this formulation have been made by Gupta and Studden (1970), Gupta and Nagel (1971), Gupta and Panchapakesan (1972), Santner (1975), Gupta and Huang (1975a, 1975b), Gupta and Huang (1976), Bickel and Yahav (1977), Gupta and Hsiao (1983), Gupta and Huang (1980), Lorenzen and McDonald (1981). Contributions to the nonparametric subset selection procedures have been made by Rizvi and Sobel (1967), Barlow and Gupta (1969), Nagel (1970), Gupta and McDonald (1970), Randles (1970), Ghosh (1973), Hsu (1978, 1981), Huang and Panchapakesan (1982).

Recently some contributions to the selection and ranking procedures based on isotonic estimators have been made by Gupta and

Yang (1984), Gupta and Huang (1983), Gupta and Leu (1983b), Huang (1984).

There have also been some contributions to the selection and ranking procedures in two stages. These are relevant when, for example, the experimenter wants to select a subset of populations (under investigation) which contains the populations of interest so that the populations in the selected subset can be examined further. Some important contributions in this direction have made by Santner (1976), Mukhopadhyay (1980), Gupta and Kim (1984) under the classical setting, and Miescke (1980, 1983), Gupta and Miescke (1982), Gupta and Miescke (1984) under the Bayesian setting.

For further developments in both formulations, reference can be made to Gupta and Panchapakesan (1979) (see also Gibbons, Olkin and Sobel (1977), Gupta and Huang (1981), and Dudewicz and Koo (1982)).

The main contribution of this thesis is to propose and study new subset selection procedures for some important and practical problems for the generalized family of lambda distributions. It should be pointed out that the family of Tukey's generalized lambda distributions is very broad and contains most well-known distributions as special cases.

Chapter I deals with selection and ranking procedures based on sample medians for the symmetric lambda distributions and applications of the lambda family of distributions. We investigate some properties of the lambda family of distributions. We also propose some selection procedures and study the properties of these procedures such as asymptotic relative efficiencies. An application of the lambda

distribution for approximating some constants used in the selection and ranking procedures for other symmetric theoretical distributions is made. Tables of associated constants for the proposed procedures are given in this chapter.

Chapter II deals with the problem of isotonic selection procedures for the family of lambda distributions and for logistic distributions. We propose and study some isotonic procedures for symmetric lambda distributions and for logistic distributions. In particular, we investigate the approximations of constants used in the proposed procedures. It is shown that the isotonic procedure is better than some classical procedures in terms of reducing the expected number of bad populations in the selected subset. Tables of associated constants for the proposed procedures are given in this chapter.

Chapter III deals with the problem of choosing the optimal score function for different nonparametric procedures proposed by Nagel (1970) and Gupta and McDonald (1970). The Tukey's lambda family of distributions is considered as the distribution for the score function. A Monte Carlo study for the optimal choice of the score function is carried out. This study indicates that the score function based on a uniform distribution is optimal and robust against possible deviations from the underlying distributions. Tables containing the values of score functions and the results of the simulations are given in this chapter.

Chapter IV deals with the problem of an elimination-type two-stage selection procedure under the Bayesian setting. We propose a

two-stage procedure $R(\alpha,d)$ which retains good populations at the first stage, and selects the best among selected populations. At Stage 2 we use a stopping rule to construct a $100(1-2\alpha)\%$ Highest Posterior Density (HPD) credible region with a common width 2d for the unknown means of selected populations. We study the properties of the rule $R(\alpha,d)$. Several figures are drawn to examine the performance of the procedure $R(\alpha,d)$. These figures are based on the results of a Monte Carlo study.

CHAPTER I

SELECTION AND RANKING PROCEDURES FOR TUKEY'S GENERALIZED LAMBDA DISTRIBUTIONS

1.1 Introduction

Tukey's generalized lambda distribution (hereafter called lambda distribution) was suggested by Tukey (1960) as a wide class of symmetric distributions and is defined in terms of its inverse cumulative distribution function. It has been generalized by Ramberg and Schmeiser (1972, 1974) so as to include both symmetric and asymmetric distributions. Originally, Ramberg and Schmeiser (1972, 1974) generalized and used the lambda distribution for the purpose of generation of continuous unimodal symmetric and asymmetric random variates since it is well known that the lambda distribution can be used to approximate many continuous theoretical distributions and empirical distributions. Therefore, since the work of Ramberg and Schmeiser (1972, 1974) the lambda distribution has been also used for Monte Carlo studies. Moberg, Ramberg and Randles (1978) have used the lambda distribution for Monte Carlo studies to check the robustness of the adaptive M-estimator for the selection problem under the indifference zone approach formulation. Also Ramberg, Tadikamalla, Dudewicz and Mykytka (1979) have used the lambda distribution to fit a distribution to a given set of data.

They also provided a useful table for various values of parameters of the lambda distribution for given combinations of skewness and kurtosis. Hogg, Fisher and Randles (1972) have studied the (empirical) power of the adaptive distribution-free test by using the lambda distribution for various combinations of skewness and kurtosis. Filliben (1969) has used the lambda distribution for estimating the location parameters of symmetric distributions. Joiner and Rosenblatt (1971) have studied the problem of the distribution of ranges of samples from the lambda distribution. Mykytka and Ramberg (1979) and Öztürk and Dale (1985) have studied the problem of estimating the parameters of the lambda distribution with a given data set.

If we confine ourselves to the class of unimodal continuous univariate distributions, skewness and kurtosis can be used as good measures to characterize a distribution. The lambda distribution is defined by values of its parameters which are determined by its first four central moments. The lambda distribution covers both symmetric and asymmetric distributions. The family of Burr distributions (1942, 1973) is also a general system of distributions, which is defined by two constants which determine the corresponding skewness, kurtosis, mean and variance. The Burr family, however, is much more difficult to handle than the lambda distribution family because the values of two constants of the Burr distribution do not provide a clear interpretation of its skewness and kurtosis. On the other hand, the lambda distribution is clearly defined by the location, scale and shape parameters which are directly related to the skewness and

kurtosis. The Pearson and Johnson systems (see Hahn and Shapiro (1967)), again, require several different functions to cover the classes of symmetric and asymmetric distributions. On the other hand, the lambda distribution family is defined by only one function and still it covers both symmetric and asymmetric distributions. Thus the family of lambda distributions is simple, flexible, and easy to use as well as it is quite broad and general. Hence the use of the lambda distribution as a model for selection and ranking problems provides results applicable to several parametric distributions, at least, to get approximate results. Also by changing the values of the parameters, we can examine the performance of the selection procedures by taking into consideration the given data. For example, if based on a given sample, one believes that the underlying distribution is a heavy-tail distribution, somewhere between the logistic and double exponential, then for this case one can assume the lambda distribution with several sets of values of parameters which are determined by the kurtosis, which, in this case, varies between 4.2 and 6.0. Again one can examine the robustness of any selection procedure due to several assumptions on the underlying distribution.

Recently several computer package programs in the field of selection and ranking have been developed by several authors. For example, the package RS-MCB is developed by Gupta and Hsu (1984a, 1984b) and Edwards (1984a, 1984b) has developed the package RANKSEL. But these package programs mainly deal with the normal models. But it is possible to modify these package programs to cover more models because the precision of the approximation in using the lambda distribution is

very good. We will discuss this further in Sections 2 and 4 of Chapter 1.

It is well known that for a symmetric distribution the sample median is an unbiased estimate of the location parameter and is robust in the presence of contamination from heavy-tailed distributions. Hence selection procedures based on the sample medians, under the formulation of the subset selection approach, have been developed for several distributions. Gupta and Leong (1979) have considered a procedure for selecting the largest of location parameters for the case of double exponential or Laplace distributions. Gupta and Singh (1980) have studied the case of normal distributions and Lorenzen and McDonald (1981) have considered the case of logistic distributions.

Here we consider some selection procedures based on sample medians for selecting the population associated with the largest location parameter among k populations whose observable characteristics follow lambda distributions.

In Section 1.2, we define the lambda distribution and also discuss some properties including tail-ordering.

In Section 1.3, the problem of selecting the population associated with the largest location parameter is studied for both the subset selection approach and the indifference zone approach for the symmetric lambda distribution. Some new selection procedures are proposed. The properties of these procedures such as asymptotic relative efficiencies (ARE) are studied. Also tables of constants necessary to carry out the procedures along with ARE's of the proposed selection

procedures are computed and tabulated. Comparisons of the rules based on medians with the selection rules based on sample means are provided for the case of symmetric lambda distributions with different values of parameters.

In Section 1.4, an application of the lambda distribution for approximating some constants used in the selection and ranking problems for other symmetric theoretical distributions is studied. Comparisons between exact values and approximated values are made for the case of logistic distributions.

As a closing remark, since the lambda distribution can be used to approximate theoretical continuous distributions, one can get many (approximate) results including evaluations of constants used in the various parametric situations for selection and ranking problems by using a lambda distribution by choosing values of its parameters properly.

At the end of this chapter, Table I.1 is provided for values of the scale and shape parameters for symmetric distributions for various values of the kurtosis ranging from 1.8 to 9.0 with steps of 0.1. This table gives 8 significant digits and this is an improvement over the table of Ramberg, Tadikamalla, Dudewicz and Mykytka (1979) in terms of both its scope and precision for the symmetric case.

1.2 Definition and Properties of the Lambda Distribution

The definition of the family of lambda distributions is as follows.

Definition 1.2.1. Let θ, β, γ_1, $\gamma_2 \in \mathbb{R}^1$, where $\beta \cdot \gamma_1 > 0$, $\beta \cdot \gamma_2 > 0$ and $\gamma_1 \cdot \gamma_2 > 0$. Let $F(\cdot)$ denote the cumulative distribution function (cdf) of a distribution and let $F^{-1}(\cdot)$ be its inverse. Then for $0 < p < 1$ and $x \in \mathbb{R}^1$, the lambda distribution $F(x)$ is defined by its inverse cdf as

$$x = F^{-1}(p) = \theta + \frac{1}{\beta}\{p^{\gamma_1} - (1-p)^{\gamma_2}\}, \tag{1.2.1}$$

where θ and β are location and scale parameters, respectively, and γ_1 and γ_2 are shape parameters.

If $\gamma_1 = \gamma_2$, the lambda distribution is symmetric. The moments and the support of the distribution depend upon β, γ_1 and γ_2. For example, for $\beta > 0$, $\gamma_1 > 0$ and $\gamma_2 > 0$, it has all positive moments of all order and its support is the interval $(\theta-1/\beta, \theta+1/\beta)$. On the other hand, for $\gamma_1 < -1$, $\gamma_2 > 1$ and $\gamma_1 > 1$, $\gamma_2 < -1$, there exist no positive moments. Ramberg, Tadikamalla, Dudewicz and Mykytka (1979) have studied these properties in detail and have provided some figures which characterize well-known continuous distributions by their standard third and fourth moments. Here we assume that the signs of both scale and shape parameters are the same for the symmetric case.

The mean, the variance, and the third and fourth central moments of the lambda distribution are given by

$$\mu_1 = \theta+(1/(\gamma_1+1) - 1/(\gamma_2+1))/\beta, \tag{1.2.2}$$

$$\mu_2 = \{[1/(2\gamma_1+1)-2Be(\gamma_1+1, \gamma_2+1) + 1/(2\gamma_2+1)] - [1/(\gamma_1+1) - 1/(\gamma_2+1)]^2\}/\beta^2, \tag{1.2.3}$$

$$(1.2.4)\quad \mu_3 = \{[1/(3\gamma_1+1) - 3Be(2\gamma_1+1, \gamma_2+1) + 3Be(\gamma_1+1, 2\gamma_2+1) - 1/(3\gamma_2+1)] - 3[1/(2\gamma_1+1) - 2Be(\gamma_1+1, \gamma_2+1) + 1/(2\gamma_2+1)][1/(\gamma_1+1) - 1/(\gamma_2+1)] + 2[1/(\gamma_1+1) - 1/(\gamma_2+1)]^3\}/\beta^3,$$

and

$$(1.2.5)\quad \mu_4 = \{[1/(4\gamma_1+1) - 4Be(3\gamma_1+1, \gamma_2+1) + 6Be(2\gamma_1+1, 2\gamma_2+1) - 4Be(\gamma_1+1, 3\gamma_2+1) + 1/(4\gamma_2+1)] - 4[1/(3\gamma_1+1) - 3Be(2\gamma_1+1, \gamma_2+1) + 3Be(\gamma_1+1, 2\gamma_2+1) - 1/(3\gamma_2+1)][1/(\gamma_1+1) - 1/(\gamma_2+1)] + 6[1/(2\gamma_1+1) - 2Be(\gamma_1+1,\gamma_2+1) + 1/2(\gamma_2+1)][1/(\gamma_1+1) - 1/(\gamma_2+1)]^2 - 3[1/(\gamma_1+1) - 1/(\gamma_2+1)]^4\}/\beta^4,$$

respectively, where Be(a,b) is the beta function with parameters a and b. For the symmetric case, i.e., $\gamma_1 = \gamma_2 = \gamma$, these can be simplified as

$$(1.2.6)\quad \mu_1 = \theta,$$

$$(1.2.7)\quad \mu_2 = 2[1/(2\gamma+1) - Be(\gamma+1, \gamma+1)]/\beta^2,$$

(1.2.8) $\mu_3 = 0,$

and

(1.2.9) $\mu_4 = 2[1/(4\gamma+1)-4Be(3\gamma+1, \gamma+1) + 3Be(2\gamma+1, 2\gamma+1)]/\beta^4.$

Hence the standardized fourth moment called kurtosis or a measure of peakedness, denoted by μ_4/μ_2^2 is

(1.2.10) $$\frac{\mu_4}{\mu_2^2} = \frac{1/(4\gamma+1) - 4Be(3\gamma+1, \gamma+1) + 3Be(2\gamma+1, 2\gamma+1)}{2[1/(2\gamma+1) - Be(\gamma+1, \gamma+1)]^2}.$$

Now we discuss some other properties of the family of lambda distributions. For this, we first discuss tail-ordering of distributions. The definition of a tail-ordering due to Doksum (1969) is as follows:

Definition 1.2.2. Let G and H be continuous distributions of random variables X and Y, respectively. Then G is said to be tail-ordered with respect to H, denoted by $G \underset{t}{\prec} H$, if and only if $G(0) = H(0) = \frac{1}{2}$ and $H^{-1}[G(x)] - x$ is non-decreasing on the support of G.

For symmetric continuous lambda distributions the following theorem holds.

Theorem 1.2.1. Let F and G be symmetric lambda distributions with location parameters $\theta_1 = \theta_2 = 0$, scale parameters β_1 and β_2, and shape parameters γ_1 and γ_2, respectively, where

$\gamma_1 \geq \gamma_2$. If $\beta_1/\gamma_1 \geq \beta_2/\gamma_2$, then

$$F \underset{t}{\prec} G.$$

Proof. Let $\Delta(x) = G^{-1}[F(x)] - x$. Then

$$\Delta(x) = \frac{1}{\beta_2}\left[F(x)^{\gamma_2} - (1-F(x))^{\gamma_2}\right] - x.$$

Thus

$$\Delta'(x) = \frac{d\Delta(x)}{x} = \frac{\gamma_2}{\beta_2}\left[F(x)^{\gamma_2-1} + (1-F(x))^{\gamma_2-1}\right]\frac{dF(x)}{dx} - 1.$$

Transforming $z = F(x)$, we have

$$\frac{dF(x)}{dx} = \frac{\beta_1}{\gamma_1\left(z^{\gamma_1-1}+(1-z)^{\gamma_1-1}\right)}$$

and thus, since $\gamma_1 \geq \gamma_2$, if $\beta_1/\gamma_1 \geq \beta_2/\gamma_2$,

$$\Delta'(z) = \frac{\beta_1\gamma_2}{\gamma_1\beta_2}\,\frac{z^{\gamma_2-1}+(1-z)^{\gamma_2-1}}{z^{\gamma_1-1}+(1-z)^{\gamma_1-1}} - 1$$

$$\geq \frac{z^{\gamma_2-1}(1-z^{\gamma_1-\gamma_2})+(1-z)^{\gamma_2-1}(1-(1-z)^{\gamma_1-\gamma_2})}{\left(z^{\gamma_1-1}+(1-z)^{\gamma_1-1}\right)}$$

$$\geq 0.$$

This completes the proof.

Ramberg and Schmeiser (1974) have derived the kth moment, denoted by μ_k', of the lambda distribution with $\theta = 0$, β, γ_1 and γ_2 as follows:

When μ_k' exists,

$$(1.2.11) \qquad \mu_k' = \beta^{-k} \sum_{i=0}^{k} \binom{k}{i} (-)^i Be(\gamma_1(k-i)+1,\ \gamma_2 i+1).$$

Here by using the method of moment generating functions, the first 4 moments of the sample mean based on n independent random samples from a lambda distribution with $\theta = 0$, β, γ_1 and γ_2, where β, γ_1 and γ_2 are chosen so that the moments exist, are given by the following theorem.

Theorem 1.2.2. Let $\bar{X}_n$ denote the sample mean based on n independent random samples from a lambda distribution with location parameter $\theta = 0$, scale parameter β and shape parameters γ_1 and γ_2. If values of β, γ_1 and γ_2 are such that μ_1', μ_2', μ_3' and μ_4' exist, then they are given by

$$(1.2.12) \qquad \mu_1' = \frac{SUM(1)}{\beta},$$

$$(1.2.13) \qquad \mu_2' = \frac{SUM(2)}{n\beta^2} + \frac{(n-1)}{n\beta^2} SUM^2(1),$$

$$(1.2.14) \qquad \mu_3' = \frac{SUM(3)}{n^2\beta^3} + \frac{(n-1)(n-2)SUM^3(1)}{n^2\beta^3},$$

and

$$(1.2.15)\quad \mu_4' = \frac{SUM(4)}{n^3\beta^4} + \frac{(3n-1)SUM^2(2)}{n^3\beta^4} + \frac{4(n-1)SUM(1)SUM(2)}{n^3\beta^4}$$

$$+ \frac{6(n-1)(n-2)SUM^2(1)SUM(2)}{n^3\beta^4} + \frac{(n-1)(n-2)(n-3)SUM^4(1)}{n^3\beta^4},$$

where

$$SUM(i) = \sum_{j=0}^{i} \binom{i}{j}(-)^j Be(\gamma_1(i-j)+1,\ \gamma_2 j+1).$$

Proof. From the fact that

$$\varphi_{\bar{X}_n}(t) = [\varphi_X(\tfrac{t}{n})]^n$$

and

$$\varphi_X(\tfrac{t}{n}) = \sum_{i=0}^{\infty} \frac{1}{i!}(\tfrac{t}{n\beta})^i SUM(i),$$

one can get the results by using standard methods, where $\varphi_X(t)$ is the moment generating function of a random variable X which has a lambda distribution with parameters $\theta = 0$, β, γ_1 and γ_2.

For a symmetric lambda distribution, i.e., $\gamma_1 = \gamma_2 = \gamma$, the following corollary holds.

Corollary 1.2.3. Under the same assumption as in Theorem 1.2.2 and letting $\gamma_1 = \gamma_2 = \gamma$, the following equations hold.

$$(1.2.16)\qquad \mu_1 = 0,$$

(1.2.17) $$\mu_2 = \frac{SUM(2)}{n\beta^2},$$

(1.2.18) $$\mu_3 = 0,$$

(1.2.19) $$\mu_4 = \frac{1}{n^3\beta^4}\{SUM(4) + 3(n-1)SUM^2(2)\},$$

and

(1.2.20) $$\frac{\mu_4}{\mu_2^2} = \frac{SUM(4) + 3(n-1)SUM^2(2)}{n\ SUM^2(2)}.$$

Proof. Since SUM(i) = 0 for all i odd for $\gamma_1 = \gamma_2 = \gamma$, one can get the results from Theorem 1.2.2 and hence the proof is omitted.

For a symmetric lambda distribution, the following remarks can be made.

Remarks:

(1) From Corollary 1.2.3, one can see that the limiting distribution of $\bar{X}_n$ has kurtosis 3 which is the same value as that of a normal distribution.

(2) The Corollary 1.2.3 can be utilized to approximate the distribution of the sample mean of some symmetric continuous distributions which are not infinitely divisible. Goel (1974) has derived the distribution of the sample mean from a logistic population as a series by using the method of characteristic functions and has provided tables the cdf for n = 2(1)12 at points 0.00(0.01)3.99 and n = 13(1)15 at points

1.2(0.01)3.89. Using the result of Corollary 1.2.3, the cdf of the logistic sample mean was approximated. It was seen that the maximum difference was less than 0.00155 for all values of n. This maximum error occurs at the point x = 0.6 for all the values of n. For $x \geq 1.0$, the error decreases as x increases and for $x \in [1.2, 3.9]$ the maximum error is less than 0.0007 for all n. The above discussion shows that the distribution of the sample mean of a logistic population can be approximated very well by using the lambda distribution.

1.3 Selecting the Population with the Largest Location Parameter Based on Sample Medians

1.3.1. The Proposed Rule R_T for Subset Selection - Symmetric Case

Let $\pi_1, \pi_2, \ldots, \pi_k$ be $k(\geq 2)$ independent populations which are characterized by observable random variables $X_1, X_2, \ldots, X_k$, respectively. Let X_i follow a symmetric lambda distribution with an unknown location parameter θ_i, and common known second and fourth central moments μ_2 and μ_4, $i = 1,2,\ldots,k$, respectively. This implies that the random variables X_i's have common known scale and shape parameters β and γ, respectively, given by equations (1.2.7) and (1.2.9). Also without loss of generality, we may assume $\mu_2 = 1$. Let $f(\cdot|\theta_i)$ and $F(\cdot|\theta_i)$ denote the probability density function (pdf) and cdf of a random variable X_i and let X_{ij}, $j = 1,2,\ldots,n$ be n independent observations from π_i, $i = 1,2,\ldots,k$, respectively. Let $\Omega = \{\underline{\theta} = (\theta_1,\ldots,\theta_k) \in \mathbb{R}^k\}$ be the parameter space and let $\Omega_0 = \{\underline{\theta} \in \Omega | \theta_1 = \ldots = \theta_k = \theta_0\}$. Let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ denote the ordered θ_i's. The population

associated with $\theta_{[k]}$ is called the best population. Also let $\pi_{(i)}$ denote the population corresponding to $\theta_{[i]}$. It is assumed that no prior knowledge is available for the correct pairing between $\theta_{[i]}$ and $\pi_{(i)}$, $i = 1,2,\ldots,k$. Our goal is to select a nontrivial (nonempty) subset including the best population so as to satisfy the P*-condition, i.e., $\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R) \geq P^*$, where CS stands for a correct selection i.e. a selection of any subset which includes the best. For convenience, let $n = 2m+1$, $m \geq 1$, and let $X_{i:m}$ be the sample median of π_i. Let $X_{[1]:m} \leq X_{[2]:m} \leq \cdots \leq X_{[k]:m}$ be ordered $X_{i:m}$'s. It is well known that a sample median $X_{i:m}$ has a pdf and a cdf

$$(1.3.1) \qquad g(x|\theta_i) = \frac{(2m+1)!}{(m!)^2} [F(x|\theta_i)]^m [1-F(x|\theta_i)]^m f(x|\theta_i)$$

and

$$(1.3.2) \qquad G(x|\theta_i) = I_{F(x|\theta_i)}(m+1, m+1),$$

respectively, where $I_x(a,b)$ is an incomplete beta function with parameters a and b. Let $X_{(i):m}$ be the sample median corresponding to $\theta_{[i]}$.

Now we propose the following selection rule R_T:

R_T: Select π_i if and only if $X_{i:m} \geq X_{[k]:m} - d_0$,

where d_0 (≥ 0) is chosen so as to satisfy the P*-condition. Without loss of generality, we can assume that $\mu_0 = 0$ in Ω_0. Under this assumption, let $G(\cdot)$ and $g(\cdot)$ denote the cdf and pdf of the sample median, respectively. Also under this assumption, let $f(\cdot)$ and $F(\cdot)$

denote the pdf and cdf of X_i, respectively. Then the following theorem holds.

Theorem 1.3.1. For the rule R_T,

$$(1.3.3)\quad \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_T) = \inf_{\underline{\theta}\in\Omega_0} P_{\underline{\theta}}(CS|R_T)$$

$$= \frac{(2m+1)!}{(m!)^2}\int_{-\infty}^{\infty} I^{k-1}_{F(x+d_0)}(m+1,m+1)[F(x)]^m \cdot [1-F(x)]^m f(x)dx.$$

Proof.
$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_T) = \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(\pi_{(k)} \text{ is selected}|R_T)$$

$$= \inf_{\underline{\theta}\in\Omega} \Pr\{X_{(k):m} \geq X_{(j):m}-d_0,\ j = 1,\ldots,k-1\}$$

$$= \inf_{\underline{\theta}\in\Omega} \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} G(x+\theta_{[k]}-\theta_{[j]}+d_0)g(x)dx$$

$$= \int_{-\infty}^{\infty} G^{k-1}(x+d_0)g(x)dx$$

$$= \frac{(2m+1)!}{(m!)^2}\int_{-\infty}^{\infty} I^{k-1}_{F(x+d_0)}(m+1,\ m+1)[F(x)]^m \cdot [1-F(x)]^m f(x)dx.$$

Hence the proof is complete.

Values of $d_0 \equiv d_0(k,m,P^*)$ can be obtained for various values of k,m and P* by solving for the smallest value of d_0 satisfying the following equation

$$(1.3.4)\qquad \frac{(2m+1)!}{(m!)^2}\int_{-\infty}^{\infty} I^{k-1}_{F(x+d_0)}(m+1,m+1)[F(x)]^m[1-F(x)]^m f(x)dx = P^*$$

or

$$(1.3.5)\qquad \frac{(2m+1)!}{(m!)^2}\int_{0}^{1} I^{k-1}_{F[\frac{1}{\beta}(t^\gamma-(1-t)^\gamma)+d_0]}(m+1,m+1)[t(1-t)]^m dt = P^*.$$

Using (1.3.5) values of d_0 were computed. These are given in Table I.2 for $m = 1(1)5$, $k = 2,3(2)9,10,11$, $P^* = 0.90, 0.95$ and for specified values of kurtosis $(\mu_4/\mu_2^2) = 4.6, 5.0, 5.6$ and 7.0 with $\mu_2 = 1$.

1.3.2. Properties and Performance of the Proposed Procedure R_T

Now we give some well-known definitions: Let p_i denote the probability that $\pi_{(i)}$ is selected by a selection rule R.

Definition 1.3.1.

(a) The rule R is strongly monotone in $\pi_{(i)}$ if p_i is nondecreasing in $\theta_{[i]}$ when all other components but $\theta_{[i]}$ are kept fixed and p_i is nonincreasing in $\theta_{[j]}$ for each $j \neq i$ when all other components are kept fixed.

(b) For $\underline{\theta} \in \Omega$, R is said to be monotone if $p_i \leq p_j$ for $1 \leq i < j \leq k$.

(c) For $\underline{\theta} \in \Omega$ and $1 \leq i < k$, R is said to be unbiased if $p_i \leq p_k$.

Note that strong monotonicity for all i $\Rightarrow$ monotonicity $\Rightarrow$ unbiasedness.

(d) Let $\phi_i(y_1,y_2,\ldots,y_k)$ be the probability that $\pi_{(i)}$ is selected by using any selection rule R based on statistics $y_1,y_2,\ldots,y_k$. Then R is said to be invariant (symmetric) if

$$\phi_i(y_1,\ldots,y_i,\ldots,y_j,\ldots,y_k) = \phi_j(y_1,\ldots,y_j,\ldots,y_i,\ldots,y_k).$$

Now we have the following theorem.

Theorem 1.3.2.

(a) The proposed selection procedure R_T is strongly monotone in $\pi_{(i)}$, for all $i = 1,2,\ldots,k$.

(b) The rule R_T is monotone and unbiased.

(c) The procedure R_T is invariant.

Proof. (a) The result follows from the fact that

$$(1.3.6) \qquad p_i = \Pr\{X_{(i):m} \geq X_{(j):m} - d_0,\ j = 1,\ldots,k,\ j\neq i\}$$

$$= \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j\neq i}}^{k} G(\alpha+\theta_{[i]}-\theta_{[j]}+d_0)dG(x).$$

Also the proofs of (b) and (c) follow from (1.3.6). Thus the proof is complete.

The expected size of the selected subset for the rule R_T, $E_{\underline{\theta}}(S|R_T)$, is given by

$$(1.3.7) \qquad E_{\underline{\theta}}(S|R_T) = \sum_{i=1}^{k} \Pr\{\pi_{(i)} \text{ is selected}\}$$

$$= \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j\neq i}}^{k} G(x+d_0+\theta_{[i]}-\theta_{[j]})dG(x).$$

Hence, by using the same argument as in Gupta (1965), one can prove the following theorem.

Theorem 1.3.3. For given k and P*(1/k < P* < 1),

$$(1.3.8)\qquad \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R_T) = \sup_{\underline{\theta}\in\Omega_0} E_{\underline{\theta}}(S|R_T) = k\int_{-\infty}^{\infty} G^{k-1}(x+d_0)dG(x) = kP^*.$$

Note that both $\inf_{\Omega} P(CS|R_T)$ and $\sup_{\Omega} E_{\underline{\theta}}(S|R_T)$ do not depend on the common $\theta_0 \in \Omega_0$. From (c) of Theorem 1.3.2 and Theorem 1.3.3, the following theorem holds.

Theorem 1.3.4. The procedure R_T is minimax among all invariant rules satisfying the P*-condition.

Proof. For $\underline{\theta}_0 \in \Omega_0$,

$$(1.3.9)\qquad \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_T) = \inf_{\underline{\theta}\in\Omega_0} P_{\underline{\theta}}(CS|R_T) = P_{\underline{\theta}_0}(CS|R_T) = P^*$$

and

$$(1.3.10)\qquad \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R_T) = \sup_{\underline{\theta}\in\Omega_0} E_{\underline{\theta}}(S|R_T) = E_{\underline{\theta}_0}(S|R_T) = kP^*.$$

Also for any invariant (symmetric) rule R and $\underline{\theta}_0 \in \Omega$,

$$\begin{aligned}(1.3.11)\qquad E_{\underline{\theta}_0}(S|R) &= \sum_{i=1}^{k} \Pr\{\pi_{(i)} \text{ being selected}|R\} \\ &= \sum_{i=1}^{k} \int_{-\infty}^{\infty} \phi_i(y_1,\ldots,y_k)\Big[\prod_{j=1}^{k} g(y_j)\Big]dy_1dy_2\ldots dy_k \\ &= \sum_{i=1}^{k} P_{\underline{\theta}_0}(CS|R).\end{aligned}$$

Hence for $\underline{\theta}_0 \in \Omega_0$,

$$(1.3.12) \quad E_{\underline{\theta}_0}(S|R)-E_{\underline{\theta}_0}(S|R_T) = k\{P_{\underline{\theta}_0}(CS|R)-P_{\underline{\theta}_0}(CS|R_T)\}.$$

Since the procedure R satisfies the P*-condition, from equation (1.3.12), one can see that

$$E_{\underline{\theta}_0}(S|R) \geq E_{\underline{\theta}_0}(S|R_T) = \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R_T)$$

so that

$$(1.3.13) \quad \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R) \geq \sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R_T).$$

Hence the proof is complete.

Now under a slippage configuration, that is, $\theta_{[1]} = \theta_{[k-1]} = \theta_{[k]}-\delta$, where $\delta > 0$, the asymptotic relative efficiency (ARE) of the proposed rule R_T relative to the Gupta-type procedure R_G, which will be defined later, will be discussed. First, the definition of the ARE is given as follows.

Definition 1.3.2. Under a slippage configuration with $\delta > 0$, let S' be the number of non-best populations selected. Also given $0 < \varepsilon < 1$, let $n_1(\varepsilon)$ and $n_2(\varepsilon)$ be minimum numbers of observations so that

$$(1.3.14) \quad E_{\underline{\theta}}(S'|R_i) = \varepsilon, \qquad i = 1,2,$$

for procedures R_1 and R_2. Then the ARE of the rule R_2 relative to R_1 is defined by

$$(1.3.15) \qquad ARE(R_2,R_1|\delta) = \lim_{\varepsilon \downarrow 0} \frac{n_1(\varepsilon)}{n_2(\varepsilon)},$$

provided that both procedures R_1 and R_2 satisfy the P*-condition. In the sequel, without loss of generality it will be assumed that $\theta_{[1]} = \theta_{[k-1]} = \theta_{[k]} - \delta = 0$. Also the Gupta-type procedure R_G is defined by

$$R_G: \text{ Select } \pi_i \text{ if and only if } \bar{X}_i \geq \max_j \bar{X}_j - d_G,$$

where $\bar{X}_i$'s are sample means and d_G is a nonnegative constant chosen so as to meet the P*-condition. Let n_T and n_G be the sample size for procedures R_T and R_G, respectively. Then as $n_T \to \infty$ and $n_G \to \infty$, one can see that, by use of the central limit theorem,

$$(1.3.16) \qquad \inf_{\underline{\theta} \in \Omega} P_{\underline{\theta}}(CS|R_G) \approx \int_{-\infty}^{\infty} \Phi^{k-1}(x+d_G\sqrt{n_G})d\Phi(x),$$

$$(1.3.17) \qquad \inf_{\underline{\theta} \in \Omega} P_{\underline{\theta}}(CS|R_T) \approx \int_{-\infty}^{\infty} \Phi^{k-1}(x + \frac{d_0}{\sigma_T})d\Phi(x),$$

$$(1.3.18) \qquad E_{\underline{\theta}}(S'|R_G) \approx (k-1)\int_{-\infty}^{\infty} \Phi^{k-2}(x+d_G\sqrt{n_G})\Phi(x-(\delta-d_G)\sqrt{n_G})d\Phi(x),$$

and

$$(1.3.19) \qquad E_{\underline{\theta}}(S'|R_T) \approx (k-1)\int_{-\infty}^{\infty} \Phi^{k-2}(x+d_0/\sigma_T)\Phi(x-(\delta-d_0)/\sigma_T)d\Phi(x),$$

where $\sigma_T^2 = 1/4n_T f^2(0)$.

As $\varepsilon \downarrow 0$, $n_T(\varepsilon)$ and $n_G(\varepsilon)$ become sufficiently large and thus from the equations (1.3.16) and (1.3.17), $d_G\sqrt{n_G} \approx d_0/\sigma_T$. Also the integrals of the right hand sides of equations (1.3.18) and (1.3.19) exist and integrands of both integrals are bounded and finite on $\mathbb{R}^1$. Thus

$$(1.3.20)\quad E_{\underline{\theta}}(S'|R_G) - E_{\underline{\theta}}(S'|R_T)$$

$$\approx \int_{-\infty}^{\infty} \Phi^{k-2}(x+d_G\sqrt{n_G})\{\Phi(x-(\delta-d_G)\sqrt{n_G})-\Phi(x-(\delta-d_0)/\sigma_T)\}d\Phi(x)$$

$$\approx 0.$$

Since $\Phi(x)$ is strictly increasing in x, it can be seen that

$$\frac{n_G(\varepsilon)}{n_T(\varepsilon)} \approx 4f^2(0) \qquad \text{for any } \delta > 0.$$

Hence the following theorem holds.

Theorem 1.3.5. Under the slippage configuration as defined above,

$$(1.3.21)\quad ARE(R_T, R_G|\delta) = f^2(0)$$

$$= 2^{2(\gamma-1)}\left(\frac{\beta}{\gamma}\right)^2.$$

The following table provides $ARE(R_T, R_G|\delta)$ for various values of β and γ for the following values of kurtosis μ_4/μ_2^2 = 1.8, 3.0, 4.2, 5.0(1.0) 9.0, with $\mu_2 = 1$.

Values of $ARE(R_T, R_G|\delta)$

μ_4/μ_2^2	β	γ	$ARE(R_T, R_G|\delta)$
1.8	.5744	1.0000	.3299
3.0	.1974	.1349	.6454
4.2	$-.0659\times10^{-2}$	$-.0363\times10^{-2}$	.8235
5.0	-.0870	-.0443	.9068
6.0	-.1686	-.0802	.9886
7.0	-.2306	-.1045	1.0532
8.0	-.2800	-.1233	1.0867
9.0	-.3203	-.1359	1.1503

It is already known that for the slippage configuration, ARE's of the median selection rules for the normal, logistic and double exponential distributions are 0.6366, 0.8225 and 1.0000, respectively. On the other hand, for values of kurtosis 3.0, 4.2, and 6.0 for the lambda distribution, the corresponding values of $ARE(R_T, R_G|\delta)$ are 0.6454, 0.8235 and 0.9886, respectively. These differences are mainly due to the approximation by lambda distributions with parameters β and γ for the corresponding distributions. Also one can see that when the tail of the distribution becomes heavier, $ARE(R_T, R_G|\delta)$ increases and thus the rule R_T becomes as efficient as the procedure R_G and the rule R_T is more efficient than the rule R_G for very heavy-tailed distributions.

Remark: From Theorem 1.2.1 and Theorem 1.3.5 one can see the following: With the same condition as in Theorem 1.2.1 and under a slippage configuration, the ARE(R_T, $R_G|\delta$) for a distribution F_1 is better (larger) than that of for a distribution F_2 when $F_1 \overset{\leq}{t} F_2$.

Now the performance of the rule R_T will be discussed in terms of $P_{\underline{\theta}}(CS|R_T)$, $E_{\underline{\theta}}(S'|R_T)$ and $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$. Recall that for $\underline{\theta} \in \Omega$,

$$(1.3.22)\quad P_{\underline{\theta}}(CS|R_T) = \frac{(2m+1)!}{(m!)^2}\int_0^1 \prod_{j=1}^{k-1} I_{F[\frac{1}{\beta}\{t^\gamma-(1-t)^\gamma\}+d_0+\theta_{[k]}-\theta_{[j]}]}(m+1,m+1) \cdot [t(1-t)]^m dt,$$

$$(1.3.23)\quad E_{\underline{\theta}}(S|R_T) = \sum_{i=1}^{k} P_{\underline{\theta}}\{\pi_{(i)} \text{ is selected}|R_T\}$$

$$= P_{\underline{\theta}}(CS|R_T) + E_{\underline{\theta}}(S'|R_T),$$

and

$$(1.3.24)\quad E(S'|R_T) = \sum_{i=1}^{k-1} \frac{(2m+1)!}{(m!)^2}\int_0^1 \prod_{\substack{j=1\\ j\neq i}}^{k} I_{F[\frac{1}{\beta}\{t^\gamma-(1-t)^\gamma\}+d_0+\theta_{[i]}-\theta_{[j]}]}(m+1,m+1)[t(1-t)]^m dt.$$

Here two configurations are considered, i.e., a slippage configuration $\theta_{[1]} = \theta_{[k-1]} = \theta_{[k]}-\delta$ and an equi-spaced configuration $\theta_{[1]} = \theta_{[2]}-\delta = \theta_{[i]}-(i-1)\delta = \theta_{[k]}-(k-1)\delta$, where $\delta > 0$. Under a slippage configuration equations (1.3.22) and (1.3.24) can be simplified as

$$P_{\underline{\theta}}(CS|R_T) = \frac{(2m+1)!}{(m!)^2}\int_0^1 I^{k-1}_{F[\frac{1}{\beta}\{t^\gamma-(1-t)^\gamma\}+\delta+d_0]}(m+1,m+1)[t(1-t)]^m dt$$

and

$$E_{\underline{\theta}}(S'|R_T) = (k-1)\frac{(2m+1)!}{(m!)^2}\int_0^1 I^{k-2}_{F[\frac{1}{\beta}\{t^\gamma-(1-t)^\gamma\}+d_0]}(m+1,m+1)$$

$$\cdot\, I_{F[\frac{1}{\beta}\{t^\gamma-(1-t)^\gamma\}+d_0-\delta]}(m+1,m+1)$$

$$\cdot\, [t(1-t)]^m dt.$$

Values of $P_{\underline{\theta}}(CS|R_T)$, $E_{\underline{\theta}}(S'|R_T)$, $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$ and $E_{\underline{\theta}}(S|R_T)$ under a slippage configuration are computed for δ = 0.1(0.2)0.5,1.0, m = 1(2)5, k = 2,5(2)9, P* = 0.90, 0.95 and kurtosis (μ_4/μ_2^2) = 4.6, 5.0, 5.6, 7.0 with μ_2 = 1. These are given in Table I.3. Similarly, under an equi-spaced configuration, values of $P_{\underline{\theta}}(CS|R_T)$, $E_{\underline{\theta}}(S'|R_T)$, $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$ and $E_{\underline{\theta}}(S|R_T)$ are computed. They are given in Table I.4 for δ = 0.1(0.2)0.5, m = 3,5, k = 5,7, P* = 0.90, 0.95 and kurtosis (μ_4/μ_2^2) = 4.6, 5.6, 7.0. Note that, for k = 2, values of $P_{\underline{\theta}}(CS|R_T)$, $E_{\underline{\theta}}(S'|R_T)$, $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$ and $E_{\underline{\theta}}(S|R_T)$ under an equi-spaced configuration are the same as those of under a slippage configuration. From Table I.3 and Table I.4, the following remarks can be made:

(1) As the value of kurtosis increases, values of $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$ increase and hence the proposed rule R_T can be more effective for heavy-tailed populations.

(2) Values of $P_{\underline{\theta}}(CS|R_T)/E_{\underline{\theta}}(S'|R_T)$ for P* = 0.90 are uniformly larger than those for P* = 0.95 for all combinations of values of k, m and δ for slippage configurations and also for equi-spaced configurations. This may be mainly the reason why an increase in the value of P*

causes R_T to select more non-best populations compared with the improvement on $P_{\underline{\theta}}(CS|R_T)$.

These tabulated values can help in an optimal choice of the value of P* in the sense of (approximate) maximizing the value of $P_{\underline{\theta}}(CS|R_T)$ and (approximate) minimizing the values of $E_{\underline{\theta}}(S'|R_T)$, simultaneously. (3) An increase in the values of δ decreases the values of $E_{\underline{\theta}}(S'|R_T)$ more significantly than an increase in the values of m for both configurations. Also values of $E_{\underline{\theta}}(S|R_T)$ decrease substantially as δ becomes larger for both configurations.

1.3.3. Selecting the t-Best Populations with Indifference Zone Approach-Symmetric Case

In Section 1.3.1 the subset selection approach for the selection of the population with the largest location parameter is considered. In this section, the indifference zone approach to select the t-best populations for the family of symmetric lambda distributions will be studied. Let the assumptions and notations be the same as those of Section 1.3.1 except for Ω and Ω_0, where for $\delta^* > 0$ and $1 \leq t < k$, let

$$\Omega(\delta^*: t) = \{\underline{\theta} \in \mathbb{R}^k | \theta_{[k-t+1]} - \theta_{[k-t]} \geq \delta^*\}$$

and

$$\Omega_0(\delta^*: t) = \{\underline{\theta} \in \mathbb{R}^k | \theta_{[1]} = \theta_{[k-t]} = \theta_{[k-t+1]} - \delta^* = \theta_{[k]} - \delta^*\}.$$

Then our goal is to select the t-best populations associated with $\theta_{[k-t+1]}, \ldots, \theta_{[k]}$ without regard to order, and to satisfy the condition that the probability of selecting t-best populations without regard to

order is at least P* for given δ^*, which is also called the P*-condition, where $P^* \in (1/\binom{k}{t},1)$ and δ^* are specified by the experimenter. Then the selection rule $R_I(t)$ is defined as follows.

$R_I(t)$: Select the t populations associated with $X_{[k-t+1]:m},\ldots,X_{[k]:m}$.

Then the following theorem holds.

Theorem 1.3.6. For $\delta^* > 0$,

$$\inf_{\underline{\theta}\in\Omega(\delta^*:t)} P_{\underline{\theta}}(CS|R_I(t)) = \inf_{\underline{\theta}\in\Omega_0(\delta^*:t)} P(CS|R_I(t)). \tag{1.3.25}$$

Proof. Proof is easy and hence omitted.

From Theorem 1.3.6, the least favorable configuration is $\Omega_0(\delta^*:t)$. Also the minimum size of samples n_t which guarantees the P*-condition is the smallest integer n such that

$$\inf_{\underline{\theta}\in\Omega_0(\delta^*:t)} P_{\underline{\theta}}(CS|R_I(t)) \geq P^*, \tag{1.3.26}$$

where

$$\inf_{\underline{\theta}\in\Omega_0(\delta^*,t)} P_{\underline{\theta}}(CS|R_I(t)) = t\int_{-\infty}^{\infty} G^{k-t}(x+\delta^*)(1-G(x))^{t-1}dG(x) \tag{1.3.27}$$

$$= \frac{t(2m+1)!}{(m!)^2}\int_0^1 I^{k-t}_{F[\frac{1}{\beta}(p^\gamma-(1-p)^\gamma)+\delta^*]}(m+1,m+1)[1-I_p(m+1,m+1)]^{t-1}$$

$$[p(1-p)]^m dp.$$

Remark. If μ_2 is not assumed equal to 1, δ^* in the equation (1.3.27) should be replaced with $\delta^*/\sqrt{\mu_2}$.

Table I.5 provides the minimum sample sizes for selected values of kurtosis μ_4/μ_2^2 = 3.0, 4.2, 5.6, 6.0, 7.0, P* = 0.90, 0.95, k = 2,3(2)7, 10, t = 1(1)3 (t < k), and δ^* = 0.5 and 1.0 with μ_2 = 1.

1.4. Applications of the Lambda Distribution

In this section, some applications of the lambda distribution for the evaluation of the d-values of subset selection approach in the selection and ranking problem are carried out. Here we restrict our attention to the symmetric case.

As mentioned in the introduction the lambda distribution can approximate theoretical continuous symmetric distributions if values of location, scale and shape parameters are chosen properly. The following table shows values of scale and shape parameters β and γ, respectively, with which the lambda distribution can be used to approximate some well-known symmetric distributions with μ_2 = 1.

distribution	μ_4/μ_2^2	β	γ
uniform	1.80	.5774	1.0000
normal	3.00	.1975	.1349
logistic	4.20	$-.0659\times10^{-2}$	$-.0363\times10^{-2}$
Laplace	6.00	-.1686	-.0802
t with 5 df	9.00	-.3202	-.1359
t with 10 df	4.00	.0261	.0148
t with 34 df	3.20	.1563	.1016
Cauchy	-	-3.0674	-1.0000

Remark: For the case of Cauchy distribution, entries come from the table of Ramberg and Schmeiser (1972).

Now we consider an approximation of values of d_G of the procedure R_G defined in Section 1.3.2 for the normal model. If one wants to use the selection rule R_G, one needs values of d_G and these values are provided by many authors (for example, Gupta (1956), Gupta (1963), Gupta, Nagel and Panchapakesan (1972), among others). But by using the lambda distribution one can approximate values of d_G, denoted by d_G', by solving the equation

$$\int_{-\infty}^{\infty} F^{k-1}(x+d_G')dF(x) = P^*, \tag{1.4.1}$$

where $F(\cdot)$ is a cdf of the lambda distribution with a scale parameter $p = 0.1975$ and a shape parameter $\gamma = 0.1349$. In the following table values of d_G come from Gupta, Nagel and Panchapakesan (1972) and values of d_G' are evaluated from the equation (1.4.1).

P^*	k	d_G	d_G'
0.90	2	1.8125	1.8126
	5	2.5997	2.6024
	9	2.9301	2.9339
0.95	2	2.3262	2.3279
	5	3.0551	3.0596
	9	3.3678	3.3728
0.99	2	3.2899	3.2931
	5	3.9196	3.9227
	9	4.1999	4.2015

From the above table, we see that the values of d_G' are fairly close to those of d_G. These agree to at least two decimal places. Furthermore, values of d_G' are conservative (larger than values of d_G); hence the P*-condition will not be violated if one uses d_G'-values in place of d_G-values.

Now we consider another approximation of the d-values of the subset selection procedures based on sample medians for the logistic distribution and compare those values with values from tables of Lorenzen and McDonald (1981). We know that a logistic distribution can be approximated by a lambda distribution with a scale parameter $\beta = -0.0659 \times 10^{-2}$ and a shape parameter $\gamma = -0.0363 \times 10^{-2}$. In the following table values of d_t come from the table of Lorenzen and McDonald (1981) and values of d_a are based on the approximation by using the lambda distribution.

m	P* / k	0.90		0.95	
		d_t	d_a	d_t	d_a
2	2	0.879	0.879	1.137	1.137
	5	1.274	1.273	1.510	1.510
	7	1.377	1.376	1.609	1.609
5	2	0.599	0.598	0.771	0.771
	5	0.863	0.863	1.019	1.018
	7	0.931	0.930	1.083	1.083
7	2	0.514	0.513	0.661	0.661
	5	0.740	0.739	0.872	0.872
	7	0.797	0.797	0.927	0.926
9	2	0.457	0.457	0.588	0.587
	5	0.657	0.657	0.775	0.774
	7	0.708	0.708	0.823	0.882

From the above table, we can see that the approximation by using the lambda distribution works fairly well. The values agree with each other at least to two decimal places and for many cases they agree up to three decimal places.

Based on the comparisons made so far it can be concluded that approximations based on the lambda distribution with proper values of scale and shape parameters work very well and we may not need tables for selection procedures for different distributions. More generally, for any (parametric) statistical inference problem, one may use the lambda distribution model to get approximate good results. This advantage may be useful for some package programs on selection and ranking problems mentioned in the introduction.

Table I.1

Values of β and γ of the Tukey's symmetric lambda distribution for given kurtosis and unit variance

kurtosis	β	γ	kurtosis	β	γ
1.8	.5773503	1.0000000	1.9	.5360259	.7315156
2.0	.4951808	.5843119	2.1	.4563041	.4839393
2.2	.4197244	.4092117	2.3	.3854375	.3506705
2.4	.3533229	.3032138	2.5	.3232217	.2637705
2.6	.2949687	.2303522	2.7	.2684053	.2016015
2.8	.2433846	.1765539	2.9	.2197734	.1545019
3.0	.1974514	.1349125	3.1	.1763108	.1173758
3.2	.1562549	.1015705	3.3	.1371972	.0872407
3.4	.1190600	.0741800	3.5	.1017736	.0622194
3.6	.0852749	.0512197	3.7	.0695075	.0410645
3.8	.0544199	.0316561	3.9	.0399657	.0229114
4.0	.0261027	.0147597	4.1	.0127925	.0071401
4.2	-.0006589	-.0003630	4.3	-.0123069	-.0067065
4.4	-.0241574	-.0130192	4.5	-.0355787	-.0189735
4.6	-.0465955	-.0246001	4.7	-.0572307	-.0299266
4.8	-.0675053	-.0349774	4.9	-.0774389	-.0397743
5.0	-.0870496	-.0443366	5.1	-.0963542	-.0486820
5.2	-.1053681	-.0528262	5.3	-.1141060	-.0567834
5.4	-.1225813	-.0605666	5.5	-.1308066	-.0641874
5.6	-.1387938	-.0676566	5.7	-.1465539	-.0709839
5.8	-.1540971	-.0741781	5.9	-.1614332	-.0772475
6.0	-.1685712	-.0801994	6.1	-.1755197	-.0830410
6.2	-.1822868	-.0857783	6.3	-.1888799	-.0884174
6.4	-.1953064	-.0909637	6.5	-.2015728	-.0934222
6.6	-.2076855	-.0957974	6.7	-.2136507	-.0980939
6.8	-.2194739	-.1003156	6.9	-.2251605	-.1024662
7.0	-.2307158	-.1045492	7.1	-.2361444	-.1065680
7.2	-.2414511	-.1085255	7.3	-.2466402	-.1104247
7.4	-.2517159	-.1122682	7.5	-.2566820	-.1140586
7.6	-.2615425	-.1157981	7.7	-.2663008	-.1174891
7.8	-.2709605	-.1191336	7.9	-.2755247	-.1207336
8.0	-.2799966	-.1222909	8.1	-.2843791	-.1238074
8.2	-.2886751	-.1252816	8.3	-.2928874	-.1267242
8.4	-.2970185	-.1281275	8.5	-.3010709	-.1294961
8.6	-.3050470	-.1308313	8.7	-.3089491	-.1321343
8.8	-.3127794	-.1334063	8.9	-.3165400	-.1346484
9.0	-.3202329	-.1358618			

Table I.2

Values of d_0 for the Procedure R_T with $\mu_2 = 1$.

$$\frac{\mu_4}{\mu_2^2} = 4.6$$

$$(\beta, \gamma) = (-0.0466, -0.0246)$$

m	P*	k 2	3	5	7	9	10	11
1	0.90	1.0970	1.3599	1.6026	1.7380	1.8317	1.8696	1.9033
	0.95	1.4282	1.6788	1.9139	2.0462	2.1382	2.1755	2.2088
2	0.90	0.8606	1.0640	1.2492	1.3511	1.4210	1.4491	1.4740
	0.95	1.1148	1.3064	1.4836	1.5821	1.6500	1.6774	1.7017
3	0.90	0.7305	0.9021	1.0571	1.1417	1.1996	1.2227	1.2433
	0.95	0.9440	1.1046	1.2520	1.3334	1.3893	1.4117	1.4316
4	0.90	0.6455	0.7966	0.9325	1.0064	1.0567	1.0768	1.0946
	0.95	0.8330	0.9739	1.1027	1.1734	1.2219	1.2413	1.2585
5	0.90	0.5846	0.7210	0.8434	0.9098	0.9549	0.9729	0.9883
	0.95	0.7537	0.8806	0.9963	1.0597	1.1030	1.1204	1.1357

$$\frac{\mu_4}{\mu_2^2} = 5.0$$

$$(\beta, \gamma) = (-0.0870, -0.0443)$$

m	P*	k 2	3	5	7	9	10	11
1	0.90	1.0798	1.3399	1.5813	1.7166	1.8107	1.8488	1.8827
	0.95	1.4085	1.6575	1.8924	2.0252	2.1180	2.1557	2.1893
2	0.90	0.8451	1.0455	1.2285	1.3295	1.3990	1.4270	1.4518
	0.95	1.0960	1.2853	1.4609	1.5589	1.6266	1.6539	1.6782
3	0.90	0.7165	0.8852	1.0380	1.1216	1.1788	1.2018	1.2221
	0.95	0.9267	1.0849	1.2305	1.3111	1.3665	1.3887	1.4085
4	0.90	0.6328	0.7811	0.9148	0.9876	1.2373	1.0572	1.0748
		0.8171	0.9557	1.0825	1.1524	1.2003	1.2195	1.2365
5	0.95	0.5728	0.7067	0.8270	0.8923	0.9367	0.9545	0.9702
		0.7389	0.8636	0.9774	1.0400	1.0826	1.0998	1.1150

Table I.2 (continued)

$$\frac{\mu_4}{\mu_2^2} = 5.6$$

$$(\beta, \gamma) = (-0.1389, -0.0667)$$

m	P*	k 2	3	5	7	9	10	11
1	0.90	1.0589	1.3156	1.5553	1.6905	1.7849	1.8233	1.8575
	0.95	1.3845	1.6315	1.8661	2.0000	2.0934	2.1315	2.1656
2	0.90	0.8264	1.0231	1.2035	1.2828	1.3506	1.4001	1.4023
	0.95	1.0732	1.2597	1.4334	1.5064	1.5727	1.5996	1.6234
3	0.90	0.6997	0.8649	1.0149	1.0973	1.1537	1.1764	1.1965
	0.95	0.9059	1.0611	1.2045	1.2840	1.3388	1.3609	1.3805
4	0.90	0.6175	0.7625	0.8135	0.9500	0.9980	1.0335	1.0344
	0.95	0.7979	0.9336	1.0582	1.1093	1.1558	1.1745	1.1910
5	0.90	0.5586	0.6894	0.8071	0.8712	0.9148	0.9323	0.9477
	0.95	0.7210	0.8430	0.9546	1.0160	1.0580	1.0749	1.0900

$$\frac{\mu_4}{\mu_2^2} = 7.0$$

$$(\beta, \gamma) = (-0.2306, -0.1045)$$

m	P*	k 2	3	5	7	9	10	11
1	0.90	1.0231	1.2736	1.5101	1.6448	1.7395	1.7782	1.8127
	0.95	1.3427	1.5861	1.8196	1.9540	2.0489	2.0877	2.1225
2	0.90	0.7947	0.9851	1.1608	1.2587	1.3266	1.3541	1.3785
	0.95	1.0345	1.2159	1.3862	1.4820	1.5488	1.5759	1.6000
3	0.90	0.6714	0.8306	0.9759	1.0560	1.1111	1.1334	1.1531
	0.95	0.8706	1.0209	1.1604	1.2380	1.2917	1.3134	1.3327
4	0.90	0.5917	0.7312	0.8576	0.9270	0.9744	0.9935	1.0104
	0.95	0.7656	0.8965	1.0172	1.0840	1.1300	1.1486	1.1650
5	0.90	0.5349	0.6605	0.7739	0.8357	0.8780	0.8949	0.9099
	0.95	0.6911	0.8086	0.9164	0.9758	1.0166	1.0330	1.0475

Table I.3

Performance of the Rule R_T under the slippage configuration $\underline{\theta} = (\theta,\theta,\ldots,\theta+\delta)$, where $\delta > 0$.

Kurtosis = 4.6

	P*				0.90				0.95	
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.10	2	1	.9181	.8788	1.0448	1.7970	.9601	.9378	1.0237	1.8979
		3	.9268	.8663	1.0699	1.7931	.9650	.9300	1.0377	1.8950
		5	.9328	.8564	1.0891	1.7892	.9684	.9237	1.0483	1.8921
	5	1	.9193	3.5760	.2571	4.4953	.9605	3.7865	.2537	4.7471
		3	.9291	3.5597	.2610	4.4888	.9661	3.7766	.2558	4.7427
		5	.9357	3.5462	.2639	4.4820	.9698	3.7684	.2573	4.7381
	7	1	.9195	5.3755	.1711	6.2940	.9606	5.6863	.1689	6.6469
		3	.9295	5.3580	.1735	6.2876	.9663	5.6758	.1703	6.6421
		5	.9363	5.3438	.1752	6.2801	.9701	5.6670	.1712	6.6371
	9	1	.9196	7.1751	.1282	8.0946	.9607	7.5861	.1266	8.5467
		3	.9298	7.1572	.1299	8.0870	.9665	7.5753	.1276	8.5418
		5	.9367	7.1421	.1312	8.0788	.9702	7.5661	.1282	8.5363
.30	2	1	.9464	.8266	1.1450	1.7730	.9750	.9060	1.0762	1.8810
		3	.9633	.7765	1.2405	1.7397	.9839	.8713	1.1293	1.8551
		5	.9727	.7347	1.3240	1.7074	.9886	.8407	1.1759	1.8292
	5	1	.9486	3.5080	.2704	4.4566	.9759	3.7472	.2604	4.7231
		3	.9668	3.4251	.2823	4.3919	.9854	3.6937	.2668	4.6790
		5	.9765	3.3466	.2918	4.3231	.9901	3.6408	.2720	4.6309
	7	1	.9490	5.3038	.1789	6.2528	.9760	5.6452	.1729	6.6213
		3	.9674	5.2103	.1857	6.1777	.9856	5.5859	.1765	6.5715
		5	.9772	5.1191	.1909	6.0962	.9904	5.5253	.1793	6.5157
	9	1	.9492	7.1010	.1337	8.0502	.9761	7.5439	.1294	8.5200
		3	.9678	7.0003	.1383	7.9682	.9858	7.4806	.1318	8.4664
		5	.9776	6.8991	.1417	7.8767	.9906	7.4140	.1336	8.4046

Table I.3 (continued)

Kurtosis = 4.6

	P*				0.90				0.95	
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.50	2	1	.9659	.7606	1.2699	1.7265	.9847	.8624	1.1419	1.8471
		3	.9830	.6585	1.4929	1.6414	.9931	.7835	1.2675	1.7766
		5	.9903	.5741	1.7251	1.5644	.9964	.7120	1.3994	1.7084
	5	1	.9682	3.4044	.2844	4.3726	.9856	3.6845	.2675	4.6701
		3	.9857	3.1912	.3089	4.1769	.9941	3.5364	.2811	4.5306
		5	.9926	2.9772	.3334	3.9698	.9972	3.3759	.2954	4.3731
	7	1	.9686	5.1898	.1866	6.1584	.9857	5.5774	.1767	6.5631
		3	.9862	4.9349	.1998	5.9211	.9943	5.4048	.1840	6.3991
		5	.9929	4.6655	.2128	5.6584	.9973	5.2078	.1915	6.2051
	9	1	.9688	6.9801	.1388	7.9488	.9858	7.4727	.1319	8.4585
		3	.9846	6.6938	.1474	7.6803	.9944	7.2817	.1366	8.2762
		5	.9932	6.3789	.1557	7.3720	.9974	7.0554	.1414	8.0528
1.00	2	1	.9901	.5461	1.8130	1.5362	.9959	.6949	1.4332	1.6908
		3	.9982	.3166	3.1524	1.3148	.9994	.4605	2.1704	1.4598
		5	.9996	.1805	5.5391	1.1800	.9999	.2936	3.4057	1.2935
	5	1	.9913	2.9307	.3382	3.9220	.9963	3.3633	.2962	4.3596
		3	.9987	2.1117	.4730	3.1104	.9995	2.6562	.3763	3.6558
		5	.9998	1.4392	.6947	2.4389	.9999	1.9724	.5070	2.9723
	7	1	.9915	4.6215	.2145	5.6130	.9964	5.2040	.1916	6.2003
		3	.9988	3.4941	.2859	4.4929	.9996	4.2693	.2341	5.2689
		5	.9998	2.4835	.4026	3.4833	.9999	3.2846	.3044	4.2846
	9	1	.9915	6.3411	.1564	7.3327	.9964	7.0613	.1411	8.0576
		3	.9988	4.9446	.2020	5.9435	.9996	5.9334	.1685	6.9330
		5	.9998	3.6118	.2768	4.6116	.9999	4.6713	.2141	5.6712

Table I.3 (continued)

Kurtosis = 5.0

	P*		0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.10	2	1	.9183	.8785	1.0453	1.7967	.9601	.9377	1.0239	1.8978
		3	.9272	.8656	1.0712	1.7927	.9652	.9296	1.0383	1.8947
		5	.9333	.8554	1.0910	1.7887	.9686	.9231	1.0493	1.8917
	5	1	.9193	3.5756	.2571	4.4949	.9605	3.7864	.2537	4.7468
		3	.9294	3.5587	.2612	4.4881	.9662	3.7760	.2559	4.7422
		5	.9362	3.5449	.2641	4.4811	.9700	3.7675	.2575	4.7375
	7	1	.9194	5.3749	.1711	6.2943	.9605	5.6859	.1689	6.6465
		3	.9298	5.3569	.1736	6.2867	.9664	5.6752	.1703	6.6416
		5	.9368	5.3421	.1753	6.2789	.9703	5.6662	.1712	6.6364
	9	1	.9195	7.1744	.1282	8.0939	.9606	7.5857	.1266	8.5462
		3	.9301	7.1557	.1300	8.0857	.9665	7.5746	.1276	8.5411
		5	.9371	7.1400	.1312	8.0771	.9704	7.5649	.1283	8.5353
.30	2	1	.9467	.8254	1.1469	1.7721	.9750	.9055	1.0768	1.8805
		3	.9638	.7735	1.2460	1.7374	.9841	.8694	1.1319	1.8535
		5	.9734	.7303	1.3328	1.7037	.9888	.8376	1.1805	1.8265
	5	1	.9486	3.5069	.2705	4.4555	.9758	3.7469	.2604	4.7227
		3	.9672	3.4206	.2828	4.3878	.9855	3.6911	.2670	4.6766
		5	.9770	3.3387	.2926	4.3157	.9903	3.6357	.2724	4.6260
	7	1	.9489	5.3026	.1789	6.2516	.9759	5.6449	.1729	6.6208
		3	.9678	5.2053	.1859	6.1732	.9858	5.5832	.1766	6.5690
		5	.9777	5.1098	.1913	6.0875	.9906	5.5198	.1795	6.5104
	9	1	.9491	7.1000	.1337	8.0491	.9759	7.5437	.1294	8.5196
		3	.9682	6.9947	.1384	7.9629	.9859	7.4777	.1319	8.4636
		5	.9781	6.8887	.1420	7.8668	.9908	7.4078	.1338	8.3986

Table I.3 (continued)

Kurtosis = 5.0

	P*		0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.50	2	1	.9661	.7581	1.2744	1.7242	.9847	.8612	1.1435	1.8459
		3	.9834	.6522	1.5078	1.6356	.9933	.7790	1.2751	1.7722
		5	.9908	.5652	1.7531	1.5559	.9965	.7046	1.4143	1.7011
	5	1	.9681	3.4016	.2846	4.3698	.9854	3.6837	.2675	4.6691
		3	.9860	3.1786	.3102	4.1645	.9942	3.5287	.2817	4.5229
		5	.9928	2.9546	.3360	3.9475	.9973	3.3597	.2968	4.3569
	7	1	.9684	5.1872	.1867	6.1556	.9855	5.5769	.1767	6.5624
		3	.9864	4.9203	.2005	5.9067	.9944	5.3961	.1843	6.3905
		5	.9932	4.6371	.2142	5.6303	.9974	5.1885	.1922	6.1859
	9	1	.9685	6.9778	.1388	7.9464	.9855	7.4725	.1319	8.4580
		3	.9867	6.6774	.1478	7.6641	.9945	7.2724	.1367	8.2669
		5	.9934	6.3457	.1566	7.3392	.9975	7.0331	.1418	8.0306
1.00	2	1	.9901	.5387	1.8381	1.5288	.9958	.6898	1.4437	1.6856
		3	.9983	.3043	3.2801	1.3026	.9994	.4473	2.2345	1.4466
		5	.9996	.1689	5.9200	1.1685	.9999	.2784	3.5912	1.2783
	5	1	.9911	2.9159	.3399	3.9070	.9962	3.3567	.2968	4.3528
		3	.9987	2.0623	.4843	3.0610	.9995	2.6143	.3823	3.6138
		5	.9998	1.3749	.7272	2.3747	.9999	1.9052	.5248	2.9051
	7	1	.9912	4.6055	.2152	5.5968	.9962	5.1981	.1916	6.1943
		3	.9988	3.4255	.2916	4.4243	.9996	4.2142	.2372	5.2138
		5	.9998	2.3849	.4192	3.3847	.9999	3.1876	.3137	4.1875
	9	1	.9913	6.3255	.1567	7.3168	.9962	7.0567	.1412	8.0529
		3	.9989	4.8590	.2056	5.8579	.9996	5.8674	.1704	6.8670
		5	.9998	3.4803	.2873	4.4801	.9999	4.5456	.2200	5.5456

Table I.3 (continued)

Kurtosis = 5.6

p*			0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.10	2	1	.9185	.8782	1.0459	1.7967	.9602	.9376	1.0240	1.8978
		3	.9277	.8648	1.0727	1.7925	.9654	.9292	1.0390	1.8946
		5	.9339	.8543	1.0933	1.7882	.9689	.9224	1.0504	1.8913
	5	1	.9194	3.5757	.2571	4.4951	.9605	3.7866	.2537	4.7470
		3	.9299	3.5581	.2613	4.4880	.9664	3.7758	.2559	4.7422
		5	.9368	3.5436	.2644	4.4804	.9703	3.7669	.2576	4.7371
	7	1	.9195	5.3753	.1711	6.2948	.9605	5.6866	.1689	6.6472
		3	.9303	5.3566	.1737	6.2869	.9666	5.6750	.1703	6.6416
		5	.9374	5.3410	.1755	6.2785	.9705	5.6655	.1713	6.6361
	9	1	.9196	7.1750	.1282	8.0945	.9605	7.5863	.1266	8.5469
		3	.9305	7.1555	.1300	8.0859	.9667	7.5745	.1276	8.5412
		5	.9377	7.1392	.1313	8.0769	.9707	7.5645	.1283	8.5352
.30	2	1	.9470	.8240	1.1493	1.7711	.9751	.9050	1.0775	1.8801
		3	.9646	.7699	1.2529	1.7345	.9844	.8672	1.1352	1.8516
		5	.9743	.7248	1.3441	1.6991	.9892	.8338	1.1864	1.8230
	5	1	.9487	3.5062	.2706	4.4549	.9757	3.7470	.2604	4.7227
		3	.9678	3.4154	.2834	4.3832	.9857	3.6884	.2673	4.6741
		5	.9777	3.3289	.2937	4.3066	.9906	3.6297	.2729	4.6203
	7	1	.9489	5.3024	.1790	6.2513	.9758	5.6457	.1728	6.6215
		3	.9684	5.2002	.1862	6.1685	.9860	5.5804	.1767	6.5664
		5	.9784	5.0989	.1919	6.0773	.9909	5.5131	.1797	6.5040
	9	1	.9490	7.1001	.1337	8.0491	.9758	7.5446	.1293	8.5203
		3	.9687	6.9894	.1386	7.9581	.9861	7.4750	.1319	8.4611
		5	.9788	6.8771	.1423	7.8559	.9910	7.4010	.1339	8.3921

Table I.3 (continued)

Kurtosis = 5.6

	P*		0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.50	2	1	.9665	.7550	1.2801	1.7214	.9848	.8598	1.1454	1.8445
		3	.9840	.6444	1.5271	1.6284	.9935	.7734	1.2846	1.7668
		5	.9913	.5539	1.7896	1.5452	.9967	.6952	1.4337	1.6919
	5	1	.9681	3.3988	.2848	4.3668	.9853	3.6832	.2675	4.6685
		3	.9864	3.1629	.3119	4.1493	.9943	3.5192	.2826	4.5135
		5	.9932	2.9257	.3395	3.9189	.9974	3.3392	.2987	4.3366
	7	1	.9683	5.1850	.1867	6.1533	.9853	5.5774	.1767	6.5628
		3	.9868	4.9025	.2013	5.8893	.9945	5.3856	.1847	6.3801
		5	.9936	4.6013	.2159	5.5948	.9975	5.1636	.1932	6.1611
	9	1	.9683	6.9763	.1388	7.9447	.9853	7.4734	.1318	8.4587
		3	.9870	6.6580	.1482	7.6450	.9946	7.2613	.1370	8.2558
		5	.9938	6.3045	.1576	7.2983	.9976	7.0055	.1424	8.0031
1.00	2	1	.9902	.5292	1.8710	1.5194	.9958	.6833	1.4573	1.6791
		3	.9984	.2891	3.4531	1.2875	.9994	.4307	2.3205	1.4301
		5	.9997	.1549	6.4544	1.1545	.9999	.2598	3.8494	1.2596
	5	1	.9909	2.8973	.3420	3.8883	.9960	3.3487	.2974	4.3447
		3	.9988	1.9996	.4995	2.9984	.9995	2.5607	.3904	3.5602
		5	.9998	1.2947	.7722	2.2945	.9999	1.8204	.5493	2.8203
	7	1	.9910	4.5861	.2161	5.5771	.9960	5.1923	.1918	6.1883
		3	.9988	3.3382	.2992	4.3371	.9996	4.1432	.2413	5.1427
		5	.9998	2.2614	.4421	3.2612	.9999	3.0631	.3264	4.0630
	9	1	.9910	6.3064	.1571	7.2974	.9960	7.0523	.1412	8.0483
		3	.9989	4.7502	.2103	5.7491	.9996	5.7824	.1729	6.7819
		5	.9998	3.3152	.3016	4.3151	.9999	4.3856	.2280	5.3855

Table I.3 (continued)

Kurtosis = 7.0

			0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.10	2	1	.9189	.8776	1.0470	1.7965	.9603	.9374	1.0243	1.8977
		3	.9286	.8633	1.0756	1.7919	.9658	.9284	1.0404	1.8942
		5	.9351	.8521	1.0974	1.7872	.9695	.9212	1.0525	1.8907
	5	1	.9196	3.5756	.2572	4.4951	.9604	3.7866	.2536	4.7471
		3	.9306	3.5566	.2617	4.4872	.9668	3.7752	.2561	4.7419
		5	.9380	3.5411	.2649	4.4791	.9708	3.7656	.2578	4.7364
	7	1	.9196	5.3752	.1711	6.2948	.9604	5.6866	.1689	6.6470
		3	.9310	5.3551	.1738	6.2861	.9669	5.6744	.1704	6.6413
		5	.9385	5.3382	.1758	6.2767	.9711	5.6642	.1714	6.6352
	9	1	.9196	7.1751	.1282	8.0947	.9604	7.5865	.1266	8.5469
		3	.9312	7.1540	.1302	8.0852	.9670	7.5739	.1277	8.5409
		5	.9388	7.1367	.1315	8.0755	.9712	7.5633	.1284	8.5345
.30	2	1	.9477	.8215	1.1537	1.7692	.9753	.9039	1.0789	1.8792
		3	.9659	.7632	1.2656	1.7291	.9850	.8630	1.1414	1.8479
		5	.9757	.7148	1.3651	1.6905	.9898	.8267	1.1973	1.8166
	5	1	.9489	3.5043	.2708	4.4531	.9755	3.7468	.2604	4.7223
		3	.9688	3.4053	.2845	4.3742	.9861	3.6830	.2677	4.6691
		5	.9790	3.3107	.2957	4.2896	.9911	3.6184	.2739	4.6095
	7	1	.9489	5.3010	.1790	6.2499	.9755	5.6455	.1728	6.6211
		3	.9693	5.1892	.1868	6.1586	.9863	5.5748	.1769	6.5611
		5	.9795	5.0777	.1929	6.0573	.9913	5.5004	.1802	6.4918
	9	1	.9489	7.0992	.1337	8.0481	.9755	7.5449	.1293	8.5204
		3	.9696	6.9783	.1390	7.9479	.9864	7.4693	.1321	8.4557
		5	.9799	6.8545	.1430	7.8344	.9915	7.3876	.1342	8.3791

Table I.3 (continued)

Kurtosis = 7.0

			0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
.50	2	1	.9671	.7491	1.2909	1.7162	.9848	.8570	1.1492	1.8418
		3	.9850	.6301	1.5634	1.6151	.9938	.7627	1.3030	1.7566
		5	.9922	.5336	1.8592	1.5258	.9970	.6779	1.4707	1.6749
	5	1	.9681	3.3927	.2854	4.3608	.9850	3.6816	.2675	4.6667
		3	.9871	3.1332	.3150	4.1202	.9946	3.5009	.2841	4.4955
		5	.9938	2.8719	.3460	3.8657	.9976	3.3005	.3023	4.2981
	7	1	.9681	5.1798	.1869	6.1480	.9850	5.5764	.1766	6.5614
		3	.9874	4.8677	.2028	5.8551	.9947	5.3651	.1854	6.3598
		5	.9941	4.5331	.2193	5.5272	.9977	5.1165	.1950	6.1143
	9	1	.9681	6.9721	.1389	7.9402	.9850	7.4734	.1318	8.4583
		3	.9876	6.6201	.1492	7.6076	.9947	7.2395	.1374	8.2342
		5	.9943	6.2259	.1597	7.2202	.9978	6.9523	.1435	7.9501
1.00	2	1	.9903	.5120	1.9344	1.5023	.9957	.6709	1.4841	1.6667
		3	.9985	.2628	3.7996	1.2613	.9995	.4010	2.4924	1.4005
		5	.9997	.1320	7.5732	1.1317	.9999	.2281	4.3831	1.2280
	5	1	.9907	2.8614	.3462	3.8521	.9957	3.3322	.2988	4.3280
		3	.9989	1.8857	.5297	2.8845	.9996	2.4605	.4063	3.4600
		5	.9998	1.1562	.8647	2.1560	.9999	1.6691	.5991	2.6690
	7	1	.9907	4.5475	.2178	5.5381	.9957	5.1771	.1923	6.1728
		3	.9989	3.1765	.3145	4.1754	.9996	4.0092	.2493	5.0088
		5	.9998	2.0437	.4892	3.0436	.9999	2.8382	.3523	3.8382
	9	1	.9906	6.2678	.1581	7.2584	.9957	7.0402	.1414	8.0359
		3	.9990	4.5473	.2197	5.5462	.9996	5.6207	.1778	6.6203
		5	.9998	3.0217	.3309	4.0215	.9999	4.0927	.2443	5.0927

Table I.4

Performance of the Rule R_T under the equally-spaced configuration, $\underline{\theta} = (\theta, \theta+\delta, \ldots, \theta+(k-1)\delta)$, where $\delta>0$

Kurtosis = 4.6

		P*	0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
0.1	5	3	.9550	3.4127	.2798	4.3677	.9794	3.6812	.2661	4.6606
		5	.9633	3.3275	.2895	4.2908	.9837	3.6205	.2717	4.6041
	7	3	.9653	4.9594	.1946	5.9247	.9844	5.4077	.1821	6.3921
		5	.9723	4.7363	.2054	5.7092	.9882	5.2354	.1887	6.2236
0.3	5	3	.9875	2.4713	.3996	3.4588	.9948	2.9066	.3423	3.9013
		5	.9921	2.0643	.4806	3.0564	.9969	2.4834	.4014	3.4803
	7	3	.9914	2.8956	.3424	3.8869	.9964	3.4703	.2871	4.4668
		5	.9947	2.3105	.4305	3.3051	.9979	2.7799	.3590	3.7778
0.5	5	3	.9956	1.5375	.6476	2.5330	.9983	1.8930	.5273	2.8913
		5	.9979	1.1687	.8538	2.1666	.9992	1.4495	.6894	2.4487
	7	3	.9970	1.6943	.5884	2.6914	.9988	2.0555	.4859	3.0543
		5	.9986	1.2889	.7748	2.2875	.9995	1.5700	.6366	2.5695

Table I.4(continued)

Kurtosis = 5.6

		P*	0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
0.1	5	3	.9559	3.4019	.2810	4.3577	.9798	3.6751	.2666	4.6548
		5	.9645	3.3085	.2915	4.2730	.9842	3.6080	.2728	4.5921
	7	3	.9661	4.9318	.1959	5.8979	.9847	5.3906	.1827	6.3753
		5	.9738	4.6870	.2078	5.6608	.9885	5.1997	.1901	6.1882
0.3	5	3	.9879	2.4037	.4110	3.3916	.9947	2.8465	.3495	3.8414
		5	.9926	1.9819	.5008	2.9745	.9970	2.3973	.4159	3.3943
	7	3	.9916	2.7911	.3553	3.7827	.9965	3.3584	.2967	4.3549
		5	.9950	2.2084	.4506	3.2033	.9980	2.6628	.3748	3.6608
0.5	5	3	.9958	1.4670	.6788	2.4628	.9983	1.8144	.5502	2.8127
		5	.9981	1.1040	.9040	2.1021	.9993	1.3747	.7269	2.3739
	7	3	.9971	1.6189	.6160	2.6160	.9989	1.9705	.5069	2.9694
		5	.9987	1.2199	.8187	2.2186	.9995	1.4912	.6703	2.4907

Table I.4 (continued)

Kurtosis = 7.0

		P*	0.90				0.95			
δ	k	m	P(CS)	E(S')	P(CS)/E(S')	E(S)	P(CS)	E(S')	P(CS)/E(S')	E(S)
0.1	5	3	.9568	3.3907	.2822	4.3475	.9801	3.6687	.2672	4.6488
		5	.9656	3.2891	.2936	4.2547	.9846	3.5952	.2739	4.5798
	7	3	.9668	4.9027	.1972	5.8695	.9849	5.3726	.1833	6.3576
		5	.9747	4.6358	.2103	5.6105	.9889	5.1623	.1916	6.1512
0.3	5	3	.9884	2.2362	.4231	3.3246	.9950	2.7849	.3573	3.7799
		5	.9931	1.9029	.5219	2.8960	.9972	2.3131	.4311	3.3103
	7	3	.9919	2.6909	.3686	3.6828	.9966	3.2497	.3067	4.2462
		5	.9953	2.1131	.4710	3.1084	.9981	2.5533	.3909	3.5514
0.5	5	3	.9960	1.4010	.7109	2.3970	.9984	1.7399	.5738	2.7382
		5	.9982	1.0443	.9559	2.0425	.9993	1.3055	.7655	2.3049
	7	3	.9973	1.5479	.6443	2.5452	.9989	1.8907	.5283	2.8896
		5	.9988	1.1558	.8642	2.1547	.9996	1.4181	.7048	2.4177

Table I.5

Values of sample sizes for the Rule $R_I(t)$ with unit variance

Kurtosis	P*	δ \ k (t)	2	3	5		7		10		
			1	1	1	2	1	2	1	2	3
3.0	0.90	0.5	21	31	43	51	49	61	55	69	75
		1.0	5	9	11	13	13	15	15	17	19
	0.95	0.5	35	47	59	67	65	77	73	87	93
		1.0	9	11	15	17	17	19	19	23	23
4.2	0.90	0.5	17	25	33	41	39	47	45	55	59
		1.0	5	7	9	11	11	13	11	15	15
	0.95	0.5	27	37	47	53	53	61	57	69	73
		1.0	7	9	13	15	13	17	15	19	19
5.6	0.90	0.5	15	21	29	35	33	41	39	47	51
		1.0	5	7	9	9	9	11	11	13	13
	0.95	0.5	23	31	41	47	45	53	51	59	63
		1.0	7	9	11	13	13	15	13	15	17
6.0	0.90	0.5	15	21	29	35	33	41	37	45	51
		1.0	5	7	9	9	9	11	11	13	13
	0.95	0.5	23	31	39	45	43	51	49	57	61
		1.0	7	9	11	13	13	13	13	15	17
7.0	0.90	0.5	13	21	27	33	31	37	35	43	47
		1.0	5	5	7	9	9	11	9	11	13
	0.95	0.5	21	29	37	43	41	49	47	55	59
		1.0	7	9	11	11	11	13	13	15	15

CHAPTER II

ISOTONIC PROCEDURES FOR SELECTING POPULATIONS BETTER THAN A CONTROL FOR TUKEY'S GENERALIZED LAMBDA DISTRIBUTIONS AND LOGISTIC DISTRIBUTIONS

2.1 Introduction

The problem of selecting a subset containing all populations better than a control or standard has been considered by many authors under different formulations. Dunnett (1955), Gupta and Sobel (1958), Gupta (1965), Rizvi, Sobel and Woodworth (1968), Bechhofer (1968), Huang (1974), Naik (1975), Turnbull (1976), Broström (1977), and Gupta and Singh (1979) have studied this problem. Using a decision-theoretic Bayesian approach, Gupta and Kim (1980), Gupta and Hsiao (1981), Gupta and Miescke (1984) have also considered this problem. For further references, see Gupta and Panchapakesan (1979) and Dudewicz and Koo (1982). However, most of these papers assume that there is no knowledge about the correct ordering among unknown parameters. But in practice, there are cases where the experimenter may know the correct ordering even though the values of parameters are unknown. For example, in the pharmacological studies, a higher amount of acetaminophen in the pain reliever will result in a quicker effect on relieving fever. In this situation, when the experimenter

considers the time taken to reduce the temperature to a certain degree as a measurement of the effect, the experimenter knows the correct ordering among several pain relievers with different amounts of acetaminophen even though the true values of the times are unknown. For this case then, it is reasonable to assume an ordering prior. Selection procedures under the assumption of ordering priors are, in general, concerned with isotonic inference. Recently Gupta and Yang (1984) have considered isotonic selection procedures for the case of normal populations. They have also considered some isotonic procedures under the assumption of partial ordering. Gupta and Huang (1983) have studied isotonic procedures for the case of binomial populations and Gupta and Leu (1983b) have proposed and studied isotonic selection procedures for unknown guarantee lifetimes in the case of two-parameter exponential populations. Huang (1984) has also proposed and studied a nonparametric isotonic selection procedure.

In this chapter we investigate isotonic selection procedures for the family of lambda distributions and for the logistic populations. As pointed out earlier, the lambda family of distribution was defined by Tukey (1960) and generalized by Ramberg and Schmeiser (1972, 1974). It is well known that the lambda family of distributions can be used to approximate many univariate continuous distributions very well as shown in Chapter 1. For further discussion relating to the lambda family of distributions, reference should be made to Section 1.2 of Chapter 1. Here we restrict

ourselves to the family of symmetric lambda distributions. We also study the logistic distribution which is frequently used as a model in biological assay problems, (see for example, Berkson (1944, 1951, 1953) and Finney (1947)).

In Section 2.2, we introduce notations and definitions used in this chapter.

In Section 2.3, some isotonic selection procedures are proposed and studied for symmetric lambda populations and for the logistic populations. Especially, we investigate the approximations of constants used in the proposed procedures mainly because of difficulties involved in obtaining the exact distribution of sums of sample medians. For both the lambda distribution and the logistic distribution, moments of sums of sample medians are derived.

2.2 Preliminaries

Let $\pi_0, \pi_1, \ldots, \pi_k$ be (k+1) independent populations, where π_0 can be regarded as a control or standard population. Let a random variable X_i be the observable characteristic of π_i and let X_{ij}, $j = 1,2,\ldots,n$ be n independent random samples from π_i, $i = 1,\ldots,k$, respectively. Let $F(\cdot|\theta_i, \underline{\xi})$ be a cumulative distribution function (cdf) of the random variable X_i, where θ_i is an unknown location parameter that we are interested in and $\underline{\xi}$ is a vector of nuisance parameters which are assumed to be common and known. For the lambda populations, $\underline{\xi}$ is a vector of the common known scale and shape parameters and for the logistic populations, $\underline{\xi}$ is a common known variance. The value of θ_0 associated with π_0 may or may not be known. A population π_i is said to be "good" ("bad") if $\theta_i \geq (<)\theta_0$.

Assume that we have a simple ordering prior of $\theta_1,\ldots,\theta_k$. Without loss of generality, let $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$. Of course, the true values of θ_i's are unknown. Our goal is to select a nontrivial subset which includes all good populations with the requirement that the minimum probability of a correct selection (CS) be at least equal to a preassigned number P*.

Let $\Omega = \{\underline{\theta} = (\theta_0,\theta_1,\ldots,\theta_k)' \; -\infty < \theta_1 \leq \theta_2 \leq \ldots \leq \theta_k < \infty, \; -\infty < \theta_0 < \infty\}$ be the parameter space, where $\Omega \subseteq \mathbb{R}^{k+1}$. Also let us define

$$\Omega_0 = \{\underline{\theta} \in \Omega | \theta_k < \theta_0\},$$

$$\Omega_i = \{\underline{\theta} \in \Omega | \theta_{k-i} < \theta_0 \leq \theta_{k-i+1}\}, \quad i = 1,2,\ldots,k-1,$$

and

$$\Omega_k = \{\underline{\theta} \in \Omega | \theta_0 < \theta_1\}.$$

Then Ω_i's are mutually disjoint sets and $\Omega = \bigcup_{i=0}^{k} \Omega_i$. We now give some definitions.

Definition 2.2.1. A selection procedure R is called isotonic if and only if whenever it selects π_i with π_i, it also selects π_j when

$\pi_i \quad \pi_j$.

Definition 2.2.2. A real-valued function f defined on a poset $(S, \lesssim)$, where $\lesssim$ denotes a binary partial order on a set S, is called isotonic if f preserves the partial order on S.

Definition 2.2.3. Let g be a given function on $(S, \lesssim)$ and let W be a given positive function on $(S, \lesssim)$. An isotonic function g* on

$(S, \underset{\sim}{\leq})$ is called an isotonic regression of g with weights W if it minimizes the sum $\sum_{x\in S}[g(x)-g^*(x)]^2W(x)$ over a class of all isotonic functions on S.

From Barlow, Bartholomew, Bremner and Brunk (1972), it is known that there exists one and only one isotonic regression of a given g with weights W on S when S is simply ordered. Also the isotonic estimator of θ_i can be found by using the max-min formulas given by Ayer, Brunk, Ewing, Reid and Silverman (1955) as follows.

Let $\tilde{X}_i$ be the sample median of π_i based on n independent random samples $X_{i1},\ldots,X_{in}$, i = 1,2,...,k, respectively. For convenience, let n = 2m+1, $m \geq 0$, and let the common known variance be 1 for both lambda and logistic populations. Also let C^2 denote the common known variance of $\tilde{X}_i$. Let us define a finite set $S = \{\theta_1,\ldots,\theta_k | \theta_1 \leq \ldots \leq \theta_k\}$ and let $W(\theta_i) \equiv w_i = n$, i = 1,2,...,k, respectively. Then by the max-min formulas, the isotonic regression of g with weight W is g*, where

$$g^*(\theta_i) = \max_{1\leq s\leq i} \min_{s\leq t\leq k} \left\{ \frac{\tilde{X}_s + \ldots + \tilde{X}_t}{t-s+1} \right\}.$$

Hence the isotonic estimator $\hat{X}_{i:k}$ of θ_i is

$$\hat{X}_{i:k} = \max_{1\leq s\leq i} \hat{\tilde{X}}_{s:k},$$

and

$$\hat{\tilde{X}}_{s:k} = \min\left\{ \tilde{X}_s, \frac{\tilde{X}_s+\tilde{X}_{s+1}}{2}, \ldots, \frac{\tilde{X}_s + \ldots + \tilde{X}_k}{k-s+1} \right\},$$

for i = 1,2,...,k, respectively.

We give the following definition for the sake of completeness.

Definition 2.2.4. Let $F(\cdot|\theta_i, \underline{\xi})$ be a symmetric lambda family of distributions. Then, for $\underline{\xi} = (\beta, \gamma)$ and $0 \leq u \leq 1$,

$$F^{-1}(u) = \theta_i + \frac{1}{\beta}[u^{\gamma}-(1-u)^{\gamma}], \tag{2.2.1}$$

where θ_i is a location parameter, β is a scale parameter and γ is a shape parameter.

For further discussion on the properties of the family of lambda distributions, reference should be made to Section 1.2 of Chapter 1.

2.3. Proposed Procedures R_1 and R_2.

We confine ourselves to the class of isotonic procedures which satisfy the P*-condition, i.e., for an isotonic rule R,

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R) \geq P^*. \tag{2.3.1}$$

2.3.1. Definitions of the Proposed Rules R_1 and R_2

The cases of both θ_0 known and θ_0 unknown are considered.

(A) θ_0 known

Since θ_0 is known, no samples need to be taken from the control population π_0. Now the rule R_1 is proposed as follows:

Procedure R_1: Steps $i = 1,2,\ldots,k-1$, are defined as follows:

Step i. Select the subset $\{\pi_i,\ldots,\pi_k\}$ and stop if

$$\hat{X}_{i:k} \geq \theta_0 - Cd^{(1)}_{i:k},$$

otherwise reject π_i and go to Step i+1, and

Step k. Select π_k if

$$\hat{X}_{k:k} \geq \theta_0 - Cd^{(1)}_{k:k},$$

otherwise reject π_k and decide that none of k populations are good.

Here $d^{(1)}_{i:k}$, $i = 1,2,\ldots,k$ are chosen to be the smallest non-negative constants so that the procedure R_1 is isotonic and meets the P*-condition. Since

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_1) = \inf_{1\leq i\leq k} \inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_1), \tag{2.3.2}$$

the P*-condition is equivalent to

$$\inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_1) \geq P^*, \quad \text{for } i = 1,\ldots,k. \tag{2.3.3}$$

Also, for any $\underline{\theta} \in \Omega_i$, $1 \leq i \leq k$, let

$$\hat{\tilde{Z}}_{i:k} = \min\left\{\tilde{Z}_i, \frac{\tilde{Z}_i+\tilde{Z}_{i+1}}{2}, \ldots, \frac{\tilde{Z}_i+\ldots+\tilde{Z}_k}{k-i+1}\right\},$$

where

$$\tilde{Z}_i = \frac{\tilde{X}_i - \theta_i}{C}, \quad i = 1,\ldots,k.$$

Then

$$P_{\underline{\theta}}(CS|R_1) = P_{\underline{\theta}}\left\{\bigcup_{j=1}^{k-i+1} (\hat{X}_{j:k} \geq \theta_0 - Cd^{(1)}_{j:k})\right\} \tag{2.3.4}$$

$$= P_{\underline{\theta}}\left\{\bigcup_{j=1}^{k-i+1} \bigcup_{\ell=1}^{j} (\hat{\tilde{X}}_{\ell:k} \geq \theta_0 - Cd^{(1)}_{j:k})\right\}$$

$$\geq \Pr\left\{\bigcup_{j=1}^{k-i+1} \bigcap_{\ell=1}^{j} (\hat{\hat{Z}}_{\ell:k} + \frac{\theta_\ell - \theta_0}{C} \geq -d^{(1)}_{j:k})\right\}$$

which is non-decreasing in θ_ℓ, $\ell = 1,\ldots,k-i+1$. Thus

$$\inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_1) \geq \Pr\{\hat{\hat{Z}}_{k-i+1:k} \geq -d^{(1)}_{k-i+1:k}\}. \tag{2.3.5}$$

Also one can see that

$$\inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_1) \leq P_{\underline{\theta}^*}\left\{\bigcup_{j=1}^{k-i+1} (\hat{X}_{j:k} \geq \theta_0 - Cd^{(1)}_{j:k})\right\} \tag{2.3.6}$$

$$= \Pr\{\hat{\hat{Z}}_{k-i+1:k} \geq -d^{(1)}_{k-i+1:k}\},$$

where $\underline{\theta}^* = (\theta_0, -\infty, \ldots, -\infty, \underbrace{\theta_0, \ldots, \theta_0}_{i \text{ terms}})$.

Since $\hat{\hat{Z}}_{k-i+1:k}$ has the same distribution as $\hat{\hat{Z}}_{1:i}$, the following theorem holds.

Theorem 2.3.1. For given $P^*(0 < P^* < 1)$ and $\underline{\theta} \in \Omega_i$,

$$\inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_1) = \Pr\{\hat{\hat{Z}}_{1:i} \geq -d^{(1)}_{k-i+1:k}\}, \quad i = 1,\ldots,k. \tag{2.3.7}$$

From Theorem 2.3.1, one can get the following corollary.

Corollary 2.3.1. For a given $P^*(0 < P^* < 1)$, $d^{(1)}_{k-i+1:k}$ which is the solution of the equation

$$\Pr(\hat{\hat{Z}}_{1:i} \geq -z) = P^*$$

satisfies the P^*-condition for the procedure R_1.

Proof. The proof is straightforward and hence omitted.

The evaluation of the constants $d^{(1)}_{k-i+1:k}$ will be discussed in the next section.

Remarks:

(1) Since $\hat{\tilde{Z}}_{k-i+1:k}$ has the same distribution as $\hat{\tilde{Z}}_{1:i}$, $d^{(1)}_{k-i+1:k} = d^{(1)}_{1:i}$, $i = 1,2,\ldots,k$.

(2) It can be seen that $d^{(1)}_{1:i}$ is increasing in i.

(B) θ_0 unknown

Since θ_0 is unknown, n independent observations $X_{01},\ldots,X_{0n}$ from the control population π_0 are taken. Let $\tilde{X}_0$ denote the median of the above samples. Then the selection procedure R_2 is defined as follows:

Procedure R_2: Steps $i = 1,\ldots,k-1$, are defined as follows:

Step i. Select the subset $\{\pi_i,\ldots,\pi_k\}$ and stop if

$$\hat{X}_{i:k} \geq \tilde{X}_0 - Cd^{(2)}_{i:k},$$

otherwise reject π_i and go to Step i+1, and

Step k. Select π_k only and stop if

$$\hat{X}_{k:k} \geq \tilde{X}_0 - Cd^{(2)}_{k:k},$$

otherwise reject π_k and decide that none of them are good populations.

Now similar to Theorem 2.3.1, the following theorem holds.

Theorem 2.3.2. For given $P^*(0 < P^* < 1)$ and $\underline{\theta} \in \Omega_i$,

$$\inf_{\underline{\theta}\in\Omega_i} P_{\underline{\theta}}(CS|R_2) = \Pr\{\hat{\tilde{Z}}_{1:i} \geq \tilde{Z}_0 - d^{(2)}_{k-i+1:k}\}, \quad i = 1,\ldots,k, \tag{2.3.8}$$

where $\tilde{Z}_0 = (\tilde{X}_0 - \theta_0)/C$.

Proof. The proof is analogous to that of Theorem 2.3.1 and hence omitted.

Corollary 2.3.2. For given $P^*(0 < P^* < 1)$, $d^{(2)}_{k-i+1:k}$, which is the solution of the equation

$$\Pr\{\hat{\tilde{Z}}_{1:i} \geq \tilde{Z}_0 - t\} = P^*, \tag{2.3.9}$$

satisfies the P^*-condition for the rule R_2.

Proof. The proof is straightforward and hence omitted.

The evaluation of the constants $d^{(2)}_{k-i+1:k}$ will be discussed in the following section.

Remark: It can be seen that for $i = 1,\ldots,k$, $d^{(2)}_{k-i+1:k} = d^{(2)}_{1:i}$ and also $d^{(2)}_{1:i}$ is increasing in i.

2.3.2. The Evaluation of Constants $d^{(1)}_{k-i+1:k}$ and $d^{(2)}_{k-i+1:k}$.

Since the evaluation of constants $d^{(2)}_{k-i+1:k}$ is similar to that of constants $d^{(1)}_{k-i+1:k}$, we will discuss here only the evaluation of constants $d^{(1)}_{k-i+1:k}$.

Now to solve the equation

$$\Pr\{\hat{\hat{Z}}_{1:i} \geq -z\} = P^*, \tag{2.3.10}$$

the following lemmas are needed. First the lemma due to Gupta and Yang (1984) based on the theory of random walk will be cited without *proof.*

Lemma 2.3.1. Suppose $U_1, U_2, \ldots$ are iid random variables whose distribution is not concentrated on a half-axis. Let $S_0 = 0$, $S_j = U_1 + \ldots + U_j$, $j = 1,2,\ldots$, respectively and let $U_i = T_i - x$, where $E(T_i) = 0$, for $i = 1,2,\ldots$, respectively. Let $V_j = \min_{1 \leq r \leq j} \frac{1}{r} S_r$. Then

$$\Pr(V_{\ell+1} \geq x) = \frac{1}{\ell+1} \sum_{j=0}^{\ell} \Pr(V_j \geq x)\Pr(S_{\ell-j+1} \geq 0), \tag{2.3.11}$$

where $\Pr(V_0 \geq x) \equiv 1$ for all x.

To use Lemma 2.3.1, first it is necessary to evaluate the quantity $\Pr(S_{\ell-j+1} \geq 0)$, where for ease of notation S_j denotes the sum of j iid sample medians for both symmetric lambda and logistic populations. To find the exact and closed form of distribution of S_j is very difficult. Hence one can consider several ways to approximate the quantity $\Pr(S_{\ell-j+1} \geq 0)$, for example, (i) Cornish-Fisher expansion (ii) Monte Carlo Method (iii) Approximation by using a lambda distribution. Since the lambda family of distributions can be used to approximate many theoretical distributions very well, provided that the values of scale and shape parameters are properly chosen (based on the standardized second and fourth

moments), the method of approximation by a lambda distribution will be applied. Hence it is necessary to compute the second and fourth central moments of the sum of k sample medians from k iid symmetric lambda distributions with mean 0 and variance 1. The same problem for the case of logistic distributions will be discussed later.

Lemma 2.3.2. Let μ_r be the rth central moments of the sum of k sample medians from k iid distributions based on a common sample size $n = 2m+1$, $m \geq 0$. Then for k symmetric lambda distributions with common scale and shape parameters β and γ, respectively,

$$(2.3.12) \qquad \mu_2 = \frac{2k\,\Gamma(2m+2)}{\beta^2[\Gamma(m+1)]^2}\,\frac{[\Gamma(m+1)\Gamma(m+1+2\gamma)-[\Gamma(m+1+\gamma)]^2]}{\Gamma(2m+2+2\gamma)},$$

and

$$(2.3.13) \qquad \mu_4 = \frac{12k(k-1)}{\beta^4}\left\{\frac{\Gamma(2m+2)}{[\Gamma(m+1)]^2}\right\}^2\left\{\frac{\Gamma(m+1)\Gamma(m+1+2\gamma)-[\Gamma(m+1+2\gamma)]^2}{\Gamma(2m+2+2\gamma)}\right\}^2 +$$

$$+\frac{2k\Gamma(2m+2)}{\beta^4[\Gamma(m+1)]^2\Gamma(2m+2+4\gamma)}\{\Gamma(m+1)\Gamma(m+1+4\gamma)$$

$$-4\Gamma(m+1+\gamma)\Gamma(m+1+3\gamma)+3[\Gamma(m+1+2\gamma)]^2\},$$

where $\Gamma(\cdot)$ is a gamma function.

Proof. Let $\varphi_k(t)$ be the moment generating function of the sum of k iid sample medians. Then it is well known that $\varphi_k(t) = [\varphi_1(t)]^k$. Also one can get that

$$(2.3.14)\quad \varphi_1(t) = \frac{\Gamma(2m+2)}{[\Gamma(m+1)]^2} \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{(-)^{\ell} t^{\ell+j}}{\ell! j! \beta^{j+\ell}} \mathrm{Be}(m+1+j\gamma,\ m+1+\ell\gamma),$$

where Be(a,b) is a complete beta function with parameters a and b. Thus by the standard method, one gets the result. Hence the proof is complete.

Remark:

In addition to Lemma 2.3.2, μ_6 is computed and is given as follow:

$$(2.3.15)\quad \mu_6 = 15k(k-1)(k-2)\left\{\frac{\Gamma(2m+2)}{[\Gamma(m+1)]^2}\right\}^3 A_1^3 +$$

$$+ 15k(k-1)\left\{\frac{\Gamma(2m+2)}{[\Gamma(m+1)]^2}\right\}^2 A_1 A_2 + kA_3 \frac{\Gamma(2m+2)}{[\Gamma(m+1)]^2},$$

where

$$(2.3.16)\quad A_1 = \frac{2}{\beta^2}\{\mathrm{Be}(m+1,\ m+1+2\gamma) - \mathrm{Be}(m+1+\gamma,\ m+1+\gamma)\},$$

$$(2.3.17)\quad A_2 = \frac{2}{\beta^4}\{\mathrm{Be}(m+1,\ m+1+4\gamma) - 4\mathrm{Be}(m+1+\gamma,\ m+1+3\gamma) + 3\mathrm{Be}(m+1+2\gamma,\ m+1+2\gamma)\},$$

and

$$(2.3.18)\quad A_3 = \frac{2}{\beta^6}\{\mathrm{Be}(m+1,\ m+1+6\gamma) - 6\mathrm{Be}(m+1+\gamma,\ m+1+5\gamma) + 15\mathrm{Be}(m+1+2\gamma,\ m+1+4\gamma) - 10\mathrm{Be}(m+1+3\gamma,\ m+1+3\gamma)\}.$$

This result for μ_6 (and higher moments) can be used if one wants to use the Cornish-Fisher expansion.

To find the proper values of the scale and shape parameters of a lambda distribution from Lemma 2.3.2, values of kurtosis for the sum of k sample medians based on n = 2m+1 samples from lambda distribution with mean 0 and variance 1 are given in Table II.1 for k = 1(1)5(2)11, 15, 20 and m = 0(1)5(2)9,10(5)20,30,50 when the underlying lambda distributions have common kurtosis 4.6, 6.0 and 7.0. Furthermore, based on Lemma 2.3.1 and Lemma 2.3.2, values of $d^{(1)}_{k-i+1;k}$ for the lambda populations are computed. They are given in Table II.2 for m = 0(1)3(2)9,10, P* = 0.75, 0.90, 0.95, .099 when the underlying lambda populations have common variance 1 and common kurtosis 4.6, 6.0 and 7.0.

For the case of logistic population, the following lemma, which is similar to Lemma 2.3.2, holds.

Lemma 2.3.3. Let $n = 2m+1$, $m \geq 0$ be the common sample size of k iid logistic populations. Then the second and fourth central moments of the sum of k sample medians from k logistic population are:

$$\mu_2 = \frac{2k}{a^2}\left(\frac{1}{6}\pi^2 - \sum_{i=1}^{m}\frac{1}{i^2}\right) \tag{2.3.19}$$

and

$$\mu_4 = \frac{12k(k+1)}{a^4}\left[\frac{\pi^4}{90} - \sum_{i=1}^{m}\frac{1}{i^4}\right] + \frac{12k^2}{a^2}\left[\left(\frac{\pi^2}{6} - \sum_{i=1}^{m}\frac{1}{i^2}\right)^2 - \left(\frac{\pi^4}{90} - \sum_{i=1}^{m}\frac{1}{i^4}\right)\right], \tag{2.3.20}$$

where $a = \pi/\sqrt{3}$.

Proof. Noting the fact that

$$(2.3.21)\quad \varphi_k(t) = \prod_{i=1}^{\infty} [1 - (\frac{t/a}{m+i})^2]^{-k},$$

the proof is analogous to that of Lemma 2.3.2 and hence omitted.

Similar to the case of lambda populations, values of kurtosis for the sum of k sample medians based on n = 2m+1 samples from logistic distributions with common variance 1 are computed. These are given in Table II.3 for k = 1(1)5(2)11, 15, 20 and m = 0(1)5(2)9, 10(5)20, 30, 50. Also based on Lemma 2.3.1 and Lemma 2.3.3 values of $d_{1:k}^{(1)}$ for the logistic populations are computed. These are tabulated in Table II.4 for m = 0(1)3(2)9, 10, P* = 0.75, 0.90, 0.95, 0.99 and k = 1(1)7.

2.3.3. Expected Number of Bad Populations in the Selected Subset.

Suppose θ_0 is known and thus, without loss of generality, let $\theta_0 = 0$. Let B be the random size of bad populations in the subset selected by the procedure R_1. Then the expected number of bad populations due to the selection procedure R_1, denoted by $E_{\underline{\theta}}(B|R_1)$, can be used as a measure of the efficiency of the rule R_1. Now for any j, $0 \le j \le k$,

$$(2.3.22)\quad \sup_{\underline{\theta}\in\Omega_{k-j}} E_{\underline{\theta}}(B|R_1) = \sup_{\underline{\theta}\in\Omega_{k-j}} \sum_{k=1}^{j} P_{\underline{\theta}}\{\bigcup_{i=1}^{r} (\hat{X}_{i:k} \ge -Cd_{i:k}^{(1)})\}$$

$$= \sum_{r=1}^{j} \Pr\{\bigcup_{i=1}^{r} (\hat{Z}_{i:j} \ge -d_{i:k}^{(1)})\}.$$

Also under the same assumption as that of the rule R_1 let us consider an alternative rule R_3 which uses a fixed constant d_3 and selects a subset simultaneously. This rule R_3 is

R_3: Select π_i if and only if $X_{i:k} \geq \theta_0 - Cd_3$ for $i = 1,2,\ldots,k$, where $d_3(\geq 0)$ is chosen so as to satisfy the P*-condition. Then one can see that $d_3 = d^{(1)}_{1:k}$ and also

$$(2.3.23) \quad \sup_{\underline{\theta} \in \Omega_{k-j}} E_{\underline{\theta}}(B|R_3) = \sum_{r=1}^{j} \Pr\left(\bigcap_{i=1}^{r} (Z_{i:j} \geq -d_3)\right).$$

Now the following theorem holds.

Theorem 2.3.3. For any j, $0 \leq j \leq k$,

$$(2.3.24) \quad \sup_{\underline{\theta} \in \Omega_{k-j}} E_{\underline{\theta}}(B|R_1) \leq \sup_{\underline{\theta} \in \Omega_{k-j}} E_{\underline{\theta}}(B|R_3).$$

Proof. The proof is straightforward and is based on the fact that

$$d^{(1)}_{j:k} = d^{(1)}_{1:k-j+1} \leq d^{(1)}_{1:k} = d_3.$$

From the above theorem R_1 is uniformly better than R_3 in terms of the number of bad populations in the selected subset.

2.3.4. Another Procedure R_M.

Since the lambda family of distributions is not infinitely divisible, it is very hard to find the exact closed form of the distribution of the mean of samples from the lambda distribution.

This is also true for the logistic distribution. But as we have discussed in Chapter 1, the lambda distribution can be used to approximate a univariate continuous theoretical distribution precise enough, and thus we can use a lambda distribution to approximate the distribution of the sample mean by computing its second and fourth moments. Thus, when this kind of approximation is acceptable, we can consider another isotonic procedure R_M based on sample means instead of sample medians. Here we consider the case of lambda populations with θ_0 known. Now we define the isotonic procedure R_M as follows:

Procedure R_M: Steps i = 1,...,k-1, are defined as follows:

Step i. Select a subset $\{\pi_i,\ldots,\pi_k\}$ and stop if

$$\hat{X}^M_{i:k} \geq \theta_0 - C_M d^M_{i:k},$$

otherwise rejct π_i and go to Step i+1,

and

Step k. Select π_k and stop if

$$\hat{X}^M_{k:k} \geq \theta_0 - C_M d^M_{k:k},$$

otherwise reject π_k and decide that none of populations are good, where

$$\hat{\lambda}^M_{i:k} = \max_{1\le s\le i} \hat{\bar{X}}^M_{s:k},$$

$$\hat{\bar{X}}^M_{s:k} = \min\{\bar{X}_s, \ldots, \frac{\bar{X}_s+\ldots+\bar{X}_k}{k-s+1}\},$$

$$\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij},$$

and

$$\mathrm{Var}(\bar{X}_i) = C_M^2.$$

Here $d^M_{i:k}$ are the smallest nonnegative constants such that the procedure R_M is isotonic and meets the P*-condition.

Now similar to that for the procedure R_1, the following theorem holds.

Theorem 2.3.4. For given $P^*(0 < P^* < 1)$, $d^M_{k-i+1:k}$ which is the solution of the equation

$$\Pr\{\hat{Z}^M_{1:i} \ge -z\} = P^* \tag{2.3.26}$$

satisfies the P*-condition for the procedure R_M, where

$$Z_i = \frac{\bar{X}_i - \theta_i}{C_M},$$

and

$$\hat{Z}^M_{1:i} = \min\left\{Z_1, \ldots, \frac{Z_1+\ldots+Z_i}{i}\right\}.$$

Proof. The proof is similar to that of Corollary 2.3.1 and hence omitted.

To solve the equation (2.3.26), we can use the same method as that in Section 2.3.2 and thus it is necessary to compute second and fourth moments of the sum of k sample means based on n independent observations from each of the k populations. Then the following theorem holds.

Theorem 2.3.5. Let μ_i be the ith central moment of the sum of k sample means based on n independent samples from each of the k lambda distributions with a common scale parameter β and a common shape parameter γ. Assume that the common variance of k lambda distributions is 1. Then

$$\mu_2 = \frac{k\ \mathrm{sum}(2)}{n\beta^2},$$

$$\mu_4 = \frac{k}{n^3\beta^4}\{\mathrm{sum}(4) + 3(kn-1)\mathrm{sum}^2(2)\},$$

where

$$\mathrm{sum}(i) = \sum_{j=0}^{i} \binom{i}{j}(-)^j \mathrm{Be}(\gamma(i-j)+1,\ \gamma j+1),$$

where Be(a,b) is a complete Beta function with parameters a and b.

Proof. The proof is straightforward.

Table II.1

Kurtosis of the sum of k sample medians based on n samples from the underlying lambda distributions which have common scale and shape parameters β and γ

Kurtosis = 4.6

m	k	1	2	3	4	5	7	9	11	15	20
0		4.600	3.800	3.533	3.400	3.320	3.229	3.178	3.145	3.107	3.080
1		3.728	3.364	3.243	3.182	3.146	3.104	3.081	3.066	3.049	3.036
2		3.456	3.278	3.152	3.114	3.091	3.065	3.051	3.041	3.030	3.023
3		3.329	3.165	3.110	3.082	3.066	3.047	3.037	3.030	3.022	3.016
4		3.257	3.128	3.086	3.064	3.051	3.037	3.029	3.023	3.017	3.013
5		3.210	3.105	3.070	3.053	3.042	3.030	3.023	3.019	3.014	3.011
7		3.154	3.077	3.051	3.039	3.031	3.022	3.017	3.014	3.010	3.008
9		3.122	3.061	3.041	3.030	3.024	3.017	3.014	3.011	3.008	3.006
10		3.110	3.055	3.037	3.028	3.022	3.016	3.012	3.010	3.007	3.006
15		3.074	3.037	3.025	3.019	3.015	3.011	3.008	3.007	3.005	3.004
20		3.055	3.027	3.018	3.014	3.011	3.008	3.006	3.005	3.004	3.003
30		3.030	3.015	3.010	3.007	3.006	3.004	3.003	3.003	3.002	3.001

Table II.1 (continued)

Kurtosis = 6.0

m	k	1	2	3	4	5	7	9	11	15	20
0		6.000	4.500	4.000	3.750	3.600	3.429	3.333	3.273	3.200	3.150
1		4.097	3.548	3.366	3.274	3.219	3.157	3.122	3.100	3.073	3.055
2		3.648	3.324	3.216	3.162	3.130	3.093	3.072	3.059	3.043	3.032
3		3.456	3.228	3.152	3.114	3.091	3.065	3.051	3.041	3.030	3.023
4		3.351	3.175	3.117	3.088	3.070	3.050	3.039	3.032	3.023	3.018
5		3.285	3.142	3.095	3.071	3.057	3.041	3.032	3.026	3.019	3.014
7		3.206	3.103	3.069	3.052	3.041	3.029	3.023	3.019	3.014	3.010
9		3.162	3.081	3.054	3.040	3.032	3.023	3.018	3.015	3.011	3.008
10		3.146	3.073	3.049	3.037	3.029	3.021	3.016	3.013	3.010	3.007
15		3.098	3.049	3.033	3.025	3.020	3.014	3.011	3.009	3.007	3.005
20		3.074	3.037	3.025	3.018	3.015	3.011	3.008	3.007	3.005	3.004
30		3.049	3.025	3.016	3.012	3.010	3.007	3.005	3.004	3.003	3.002

Table II.1 (continued)

Kurtosis = 7.0

m	k	1	2	3	4	5	7	9	11	15	20
0		7.000	4.999	4.333	3.999	3.800	3.571	3.444	3.363	3.267	3.200
1		4.295	3.647	3.432	3.324	3.259	3.185	3.144	3.118	3.086	3.065
2		3.744	3.372	3.248	3.186	3.149	3.106	3.083	3.068	3.050	3.037
3		3.517	3.259	3.172	3.129	3.103	3.074	3.057	3.047	3.034	3.026
4		3.395	3.198	3.132	3.099	3.079	3.056	3.044	3.036	3.026	3.020
5		3.320	3.160	3.107	3.080	3.064	3.046	3.036	3.029	3.021	3.016
7		3.231	3.115	3.077	3.058	3.046	3.033	3.026	3.021	3.015	3.012
9		3.180	3.090	3.060	3.045	3.036	3.026	3.020	3.016	3.012	3.009
10		3.163	3.081	3.054	3.041	3.033	3.023	3.018	3.015	3.011	3.008
15		3.109	3.054	3.036	3.027	3.022	3.016	3.012	3.010	3.007	3.005
20		3.082	3.041	3.027	3.020	3.016	3.012	3.009	3.007	3.005	3.004
30		3.055	3.027	3.018	3.014	3.011	3.008	3.006	3.005	3.004	3.003

Table II.2

Values of $d_{1:k}^{(1)}$ for the case of symmetric lambda populations with common kurtosis and common variance 1

Kurtosis = 4.6

m	k	P*	0.75	0.90	0.95	0.99
0	1		0.5920	1.1949	1.6141	2.5688
	2		0.7382	1.2879	1.6796	2.5929
	3		0.7836	1.3087	1.6899	2.5938
	4		0.8029	1.3148	1.6918	2.5939
	5		0.8123	1.3167	1.6922	2.5939
	6		0.8174	1.3174	1.6922	2.5939
1	1		0.3860	0.7614	1.0081	1.5278
	2		0.4745	0.8109	1.0393	1.5358
	3		0.5005	0.8209	1.0433	1.5360
	4		0.5111	0.8236	1.0439	1.5360
	5		0.5161	0.8242	1.0440	1.5360
	6		0.5187	0.8244	1.0440	1.5360
2	1		0.3077	0.6008	0.7885	1.1670
	2		0.3758	0.6368	0.8100	1.1747
	3		0.3954	0.6437	0.8126	1.1748
	4		0.4032	0.6454	0.8130	1.1748
	5		0.4069	0.6459	0.8130	1.1748
	6		0.4088	0.6460	0.8130	1.1748
3	1		0.2634	0.5115	0.6682	0.9804
	2		0.3259	0.5408	0.6853	0.9839
	3		0.3369	0.5463	0.6873	0.9839
	4		0.3433	0.5477	0.6875	0.9839
	5		0.3463	0.5480	0.6875	0.9839
	6		0.3478	0.5481	0.6875	0.9839
5	1		0.2127	0.4107	0.5339	0.7743
	2		0.2579	0.4331	0.5466	0.7767
	3		0.2706	0.4372	0.5480	0.7767
	4		0.2756	0.4382	0.5482	0.7767
	5		0.2780	0.4384	0.5482	0.7767
	6		0.2791	0.4385	0.5482	0.7767

Table II.2 (continued)

Kurtosis = 4.6

m	k	P*	0.75	0.90	0.95	0.99
7	1		0.1832	0.3527	0.4573	0.6795
	2		0.2217	0.3715	0.4679	0.6614
	3		0.2325	0.3749	0.4690	0.6614
	4		0.2367	0.3757	0.4691	0.6614
	5		0.2387	0.3759	0.4691	0.6614
	6		0.2397	0.3759	0.4691	0.6614
9	1		0.1633	0.3138	0.4064	0.5840
	2		0.1974	0.3303	0.4155	0.5856
	3		0.2069	0.3333	0.4165	0.5856
	4		0.2107	0.3340	0.4166	0.5856
	5		0.2124	0.3342	0.4166	0.5856
	6		0.2132	0.3342	0.4166	0.5856
10	1		0.1555	0.2987	0.3866	0.5548
	2		0.1879	0.3143	0.3952	0.5563
	3		0.1970	0.3171	0.3961	0.5563
	4		0.2005	0.3178	0.3962	0.5563
	5		0.2021	0.3179	0.3962	0.5563
	6		0.2029	0.3180	0.3962	0.5563

Kurtosis = 6.0

0	1		0.5591	1.1526	1.5863	2.6451
	2		0.7055	1.2573	1.6683	2.6867
	3		0.7537	1.2834	1.6837	2.6897
	4		0.7751	1.2920	1.6874	2.6897
	5		0.7860	1.2951	1.6883	2.6897
	6		0.7920	1.2963	1.6885	2.6897
1	1		0.3619	0.7218	0.9650	1.4973
	2		0.4480	0.7731	0.9991	1.5078
	3		0.4740	0.7839	1.0040	1.5080
	4		0.4847	0.7869	1.0048	1.5080
	5		0.4899	0.7878	1.0049	1.5080
	6		0.4927	0.7881	1.0049	1.5080

Table II.2 (continued)

Kurtosis = 6.0

m	k	P*	0.75	0.90	0.95	0.99
2	1		0.2883	0.5671	0.7490	1.1286
	2		0.3537	0.6032	0.7714	1.1340
	3		0.3729	0.6103	0.7742	1.1341
	4		0.3806	0.6121	0.7747	1.1341
	5		0.3843	0.6127	0.7747	1.1341
	6		0.3862	0.6128	0.7747	1.1341
3	1		0.2468	0.4819	0.6324	0.9384
	2		0.3014	0.5107	0.6497	0.9422
	3		0.3171	0.5162	0.6518	0.9422
	4		0.3234	0.5176	0.6520	0.9422
	5		0.3264	0.5180	0.6521	0.9422
	6		0.3279	0.5181	0.6521	0.9422
5	1		0.1992	0.3861	0.5035	0.7356
	2		0.2421	0.4078	0.5160	0.7380
	3		0.2543	0.4118	0.5174	0.7380
	4		0.2591	0.4128	0.5176	0.7380
	5		0.2614	0.4131	0.5176	0.7380
	6		0.2625	0.4132	0.5176	0.7380
7	1		0.1716	0.3312	0.4305	0.6242
	2		0.2080	0.3493	0.4407	0.6261
	3		0.2183	0.3526	0.4419	0.6261
	4		0.2223	0.3534	0.4420	0.6261
	5		0.2242	0.3536	0.4420	0.6261
	6		0.2251	0.3536	0.4420	0.6261
9	1		0.1530	0.2946	0.3822	0.5515
	2		0.1852	0.3103	0.3910	0.5531
	3		0.1942	0.3132	0.3919	0.5531
	4		0.1977	0.3139	0.3920	0.5531
	5		0.1994	0.3141	0.3921	0.5531
	6		0.2002	0.3141	0.3921	0.5531

Table II.2 (continued)

Kurtosis = 6.0

m	k	P*	0.75	0.90	0.95	0.99
10	1		0.1457	0.2803	0.3634	0.5235
	2		0.1762	0.2952	0.3717	0.5250
	3		0.1848	0.2979	0.3726	0.5250
	4		0.1882	0.2985	0.3727	0.5250
	5		0.1897	0.2987	0.3727	0.5250
	6		0.1905	0.2987	0.3727	0.5250

Kurtosis = 7.0

0	1		0.5437	1.1317	1.5708	2.6758
	2		0.6894	1.2413	1.6607	2.7282
	3		0.7387	1.2702	1.6792	2.7331
	4		0.7610	1.2802	1.6840	2.7331
	5		0.7727	1.2840	1.6854	2.7331
	6		0.7793	1.2856	1.6857	2.7331
1	1		0.3507	0.7031	0.9441	1.4808
	2		0.4355	0.7550	0.9795	1.4925
	3		0.4614	0.7663	0.9848	1.4929
	4		0.4723	0.7694	0.9857	1.4929
	5		0.4776	0.7704	0.9859	1.4929
	6		0.4803	0.7708	0.9859	1.4929
2	1		0.2794	0.5513	0.7303	1.1081
	2		0.3435	0.5874	0.7530	1.1139
	3		0.3624	0.5946	0.7560	1.1140
	4		0.3701	0.5965	0.7564	1.1140
	5		0.3738	0.5970	0.7565	1.1140
	6		0.3756	0.5972	0.7565	1.1140
3	1		0.2391	0.4681	0.6156	0.9181
	2		0.2925	0.4967	0.6329	0.9221
	3		0.3079	0.5022	0.6351	0.9221
	4		0.3141	0.5036	0.6354	0.9221
	5		0.3171	0.5040	0.6354	0.9221
	6		0.3186	0.5041	0.6354	0.9221

Table II.2 (continued)

Kurtosis = 7.0

m	k	P*	0.75	0.90	0.95	0.99
5	1		0.1930	0.3747	0.4893	0.7172
	2		0.2349	0.3961	0.5017	0.7197
	3		0.2468	0.4001	0.5031	0.7197
	4		0.2515	0.4010	0.5033	0.7197
	5		0.2537	0.4013	0.5033	0.7197
	6		0.2548	0.4014	0.5033	0.7197
7	1		0.1662	0.3213	0.4181	0.6076
	2		0.2017	0.3390	0.4281	0.6095
	3		0.2117	0.3423	0.4293	0.6095
	4		0.2157	0.3431	0.4294	0.6095
	5		0.2175	0.3433	0.4294	0.6095
	6		0.2184	0.3433	0.4294	0.6095
9	1		0.1482	0.2857	0.3709	0.5363
	2		0.1795	0.3011	0.3796	0.5379
	3		0.1883	0.3039	0.3806	0.5379
	4		0.1918	0.3046	0.3807	0.5379
	5		0.1934	0.3048	0.3807	0.5379
	6		0.1942	0.3048	0.3807	0.5379
10	1		0.1411	0.2718	0.3526	0.5090
	2		0.1786	0.2864	0.3508	0.5104
	3		0.1792	0.2890	0.3617	0.5104
	4		0.1825	0.2896	0.3618	0.5104
	5		0.1840	0.2898	0.3618	0.5104
	6		0.1847	0.2898	0.3618	0.5104

Table II.3

Kurtosis of the sum of k sample medians based on n samples from the logistic distributions with common variance 1

m \ k	1	2	3	4	5	7	9	11	15	20
0	4.2000	3.6000	3.4000	3.3000	3.2400	3.1714	3.1333	3.1091	3.0800	3.0600
1	3.5938	3.2969	3.1979	3.1484	3.1188	3.0848	3.0660	3.0540	3.0396	3.0297
2	3.3813	3.1906	3.1271	3.0953	3.0763	3.0545	3.0424	3.0347	3.0254	3.0191
3	3.2785	3.1392	3.0928	3.0696	3.0557	3.0398	3.0309	3.0253	3.0186	3.0139
4	3.2187	3.1094	3.0729	3.0547	3.0437	3.0312	3.0243	3.0199	3.0146	3.0109
5	3.1798	3.0899	3.0599	3.0450	3.0360	3.0257	3.0200	3.0164	3.0120	3.0090
7	3.1325	3.0662	3.0442	3.0331	3.0265	3.0189	3.0147	3.0120	3.0088	3.0066
9	3.1048	3.0524	3.0349	3.0262	3.0210	3.0150	3.0116	3.0095	3.0070	3.0052
10	3.0948	3.0474	3.0316	3.0237	3.0190	3.0135	3.0105	3.0086	3.0063	3.0047
15	3.0641	3.0320	3.0214	3.0160	3.0128	3.0092	3.0071	3.0058	3.0043	3.0032
20	3.0483	3.0241	3.0161	3.0121	3.0097	3.0069	3.0054	3.0044	3.0032	3.0024
30	3.0319	3.0160	3.0106	3.0080	3.0064	3.0046	3.0035	3.0029	3.0021	3.0016
50	3.0182	3.0091	3.0061	3.0046	3.0036	3.0026	3.0020	3.0017	3.0012	3.0009

Table II.4

Values of $d^{(1)}_{1:k}$ for the logistic populations with common variance 1

m	k	P*	0.75	0.90	0.95	0.99
0	1		0.6047	1.2120	1.6240	2.5349
	2		0.7516	1.2983	1.6836	2.5523
	3		0.7957	1.3186	1.6900	2.5523
	4		0.8135	1.3227	1.6930	2.5523
	5		0.8223	1.3227	1.6930	2.5523
	6		0.8276	1.3227	1.6930	2.5523
	7		0.8298	1.3227	1.6930	2.5523
1	1		0.3961	0.7776	1.0253	1.5385
	2		0.4854	0.8263	1.0552	1.5456
	3		0.5114	0.8358	1.0590	1.5457
	4		0.5219	0.8382	1.0595	1.5457
	5		0.5269	0.8389	1.0596	1.5457
	6		0.5294	0.8391	1.0596	1.5457
	7		0.5308	0.8392	1.0596	1.5457
2	1		0.3158	0.6147	0.8046	1.1862
	2		0.3849	0.6506	0.8258	1.1901
	3		0.4047	0.6573	0.8282	1.1901
	4		0.4126	0.6591	0.8286	1.1901
	5		0.4162	0.6595	0.8286	1.1901
	6		0.4181	0.6596	0.8286	1.1901
	7		0.4191	0.6597	0.8286	1.1901
3	1		0.2704	0.5239	0.6830	0.9973
	2		0.3286	0.5533	0.6999	1.0006
	3		0.3451	0.5587	0.7018	1.0006
	4		0.3516	0.5601	0.7021	1.0006
	5		0.3546	0.5604	0.7021	1.0006
	6		0.3562	0.5605	0.7021	1.0006
	7		0.3570	0.5605	0.7021	1.0006
5	1		0.2183	0.4209	0.5465	0.7902
	2		0.2645	0.4436	0.5593	0.7925
	3		0.2774	0.4478	0.5607	0.7925
	4		0.2825	0.4488	0.5608	0.7925
	5		0.2850	0.4490	0.5609	0.7925
	6		0.2861	0.4491	0.5609	0.7925
	7		0.2867	0.4491	0.5609	0.7925

Table II.4 (continued)

m	k	P*	0.75	0.90	0.95	0.99
7	1		0.1881	0.3616	0.4685	0.6740
	2		0.2274	0.3807	0.4791	0.6759
	3		0.2384	0.3842	0.4803	0.6759
	4		0.2428	0.3850	0.4804	0.6759
	5		0.2447	0.3852	0.4804	0.6759
	6		0.2457	0.3852	0.4804	0.6759
	7		0.2463	0.3852	0.4804	0.6759
9	1		0.1617	0.3219	0.4165	0.5974
	2		0.2026	0.3387	0.4258	0.5990
	3		0.2123	0.3417	0.4258	0.5990
	4		0.2160	0.3424	0.4268	0.5990
	5		0.2178	0.3426	0.4268	0.5990
	6		0.2187	0.3426	0.4268	0.5990
	7		0.2192	0.3426	0.4268	0.5990
10	1		0.1596	0.3064	0.3963	0.5678
	2		0.1928	0.3223	0.4050	0.5692
	3		0.2021	0.3251	0.4059	0.5693
	4		0.2057	0.3258	0.4061	0.5693
	5		0.2073	0.3260	0.4061	0.5693
	6		0.2082	0.3260	0.4061	0.5693
	7		0.2086	0.3260	0.4061	0.5693

CHAPTER III

NONPARAMETRIC SELECTION PROCEDURES AND THEIR EFFICIENCY COMPARISONS

3.1. Introduction

Since the selection and ranking problems were introduced and formulated, many papers have been concerned with nonparametric selection procedures. Since, in practice, there are many situations in which one cannot observe the complete samples because of lack of resources, such as time, budget, unexpected accidents, but one can at least observe ranks. This kind of difficulty occurs in life-testing very frequently. Also realistically the underlying distributions of populations are almost unknown to the experimenter and hence sometimes a parametric approach to the testing hypotheses problems or other inference problems is sensitive to the assumptions on the underlying distributions. Thus, to avoid these deficiencies of the parametric approaches, nonparametric approaches are frequently used. These can provide robustness against deviations from the assumptions about the underlying distributions.

Some nonparametric selection procedures in terms of quantiles were considered by Rizvi and Sobel (1967), Barlow and Gupta (1969), among others. Also nonparametric subset selection procedures based

on ranks were studied by Nagel (1970), McDonald (1969, 1972, 1973, 1975), Gupta and McDonald (1970), Hsu (1978, 1981), Gupta, Huang and Nagel (1979), Huang and Panchapakesan (1982), Gupta and Leu (1983a), Gupta and Liang (1984) and Matsui (1984), among others. Also, Bartlett and Govindarajulu (1968) have studied locally optimal procedures based on ranks even though the functional forms of the underlying distributions are assumed to be known.

Nagel (1970) and Gupta and McDonald (1970) proposed and studied some nonparametric subset selection procedures for the location and scale models which choose a subset including the best population among k populations. The latter authors considered locally optimal selection procedures based on some functions. But the optimal choice of the score function for these procedures has not been studied. Since the rank sum statistic is easy to deal with, many proposed nonparametric subset selection proceducres are based on this statistic.

In this chapter we consider the problem of choosing the optimal score function for different procedures proposed by Nagel (1970) and Gupta and McDonald (1970). The Tukey's lambda family of distributions is considered as the distribution for the score function because this family of distributions can be used to approximate many theoretical (unimodal) continuous distributions and hence it is easy to deal with.

In Section 3.2, the problem of selection and ranking with nonparametric subset selection procedures is formulated and notations and definitions including proposed procedures are given.

In Section 3.3, we evaluate those procedures and compute constants which are necessary to carry out the procedures. Also the score function which leads the procedures to be locally optimal in the neighborhood of some points is introduced and evaluated.

A Monte Carlo study for the optimal choice of the score function is carried out in Section 3.4. This study indicates that the score function based on uniform distribution is optimal and robust against possible deviations from the underlying distributions. Also the score function which is a weighted sum of ranks turn out to be optimal for some procedures. Furthermore, it shows that the Gupta-type procedure is almost uniformly better than another available procedure. This is not the same conclusion as that in Gupta and McDonald (1970). The reason why these results are different is due to the lack of number of simulations in Gupta and McDonald (1970) for various underlying populations. Also it is due to the fact that they only use the rank sum statistics. Some tables including the values of score functions are constructed. Also some tables containing the results of simulations are provided.

3.2 Formulation

Let $\pi_1,\ldots,\pi_k$ be $k(\geq 2)$ independent populations and let X_i be an observable characteristic of π_i, $i = 1,2,\ldots,k$, respectively. Assume that a random variable X_i follows a continuous distribution $F(\cdot|\theta_i)$, and that the family $\{F(\cdot|\theta)\}$ is stochastically increasing in θ. Here we assume that the θ_i are unknown location parameters. Let X_{ij}, $j = 1,\ldots,n$ be n independent random observations from

π_i, $i = 1,2,\ldots,k$. Let R_{ij} denote the rank of the observation X_{ij} in the pooled sample of kn observations. Define

$$(3.2.1) \qquad nH_i = \sum_{j=1}^{n} a(R_{ij}), \quad i = 1,2,\ldots,k,$$

where $a(r)$ is a score function defined by

$$-\infty < a(r) = E(T(r)|G) < \infty,$$

where $T(1) \leq T(2) \leq \ldots \leq T(N)$ is an ordered sample of size $N = nk$ from a continuous distribution G. Let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ be the ordered θ_i's. Since the family $\{F(x|\theta)\}$ is stochastically increasing in θ,

$$F(x|\theta_{[1]}) \geq F(x|\theta_{[2]}) \geq \ldots \geq F(x|\theta_{[k]})$$

for any $x \in \mathbb{R}^1$.

The population associated with $\theta_{[k]}$, i.e. $F(x|\theta_{[k]})$, is called the best. In case several populations have the same largest value $\theta_{[k]}$, randomly one of them is tagged as the best. Our goal is to select a subset which contains the best with the usual requirement on the probability of a correct selection (PCS), i.e., for any procedure R,

$$(3.2.2) \qquad \inf_{\underline{\theta} \in \Omega} P_{\underline{\theta}}(CS|R) \geq P^*,$$

where $\Omega = \{\underline{\theta} | \underline{\theta} = (\theta_1,\ldots,\theta_k), \underline{\theta} \in \mathbb{R}^k\}$ is the parameter space.

Gupta and McDonald (1970) proposed procedures $R_1(G)$ and $R_3(G)$, which choose a subset containing the best, and which depend on the choice of G, as follows:

$R_1(G)$: Select π_i if and only if $H_i \geq \max_j H_j - d$, $i = 1,2,\ldots,k$,

and

$R_3(G)$: Select π_i if and only if $H_i \geq D$, $i = 1,2,\ldots,k$,

where $d(\geq 0)$ and $D(-\infty < D < \infty)$ are chosen so as to meet the P*-condition.

Note that rules $R_1(G)$ and $R_3(G)$ are equivalent if $k = 2$. Also the rule $R_3(G)$ may select an empty set. A usual choice of G is a uniform distribution which is appealing because of simplicity.

Let $\pi_{(i)}$ be the population associated with $\theta_{[i]}$. It is easy to see that, for rules $R_1(G)$ and $R_3(G)$,

$$(3.2.3) \quad Pr(CS|R_1(G)) = Pr(H_{(k)} \geq \max_j H_{(j)} - d, \quad j = 1,\ldots,k-1)$$

and

$$(3.2.4) \quad Pr(CS|R_3(G)) = Pr(H_{(k)} \geq D),$$

where $H_{(i)}$ is the H_i associated with $\pi_{(i)}$, $i = 1,2,\ldots,k$, respectively.

3.3. Comparison of the Procedures $R_1(G)$ and $R_3(G)$.

In order to compare $R_1(G)$ and $R_3(G)$ for various choices of G, we need first the results relating to the infimum of the PCS and evaluation of necessary constants.

3.3.1. PCS for $R_1(G)$ and $R_3(G)$ and Evaluation of Associated Constants

We state below (without proof) the results regarding the infimum of PCS for rules $R_1(G)$ and $R_3(G)$ obtained by Gupta and McDonald (1970).

Theorem 3.3.1. For procedures $R_1(G)$ and $R_3(G)$,

$$(3.3.1) \qquad \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_j(G)) = \inf_{\underline{\theta}\in\Omega_k} P_{\underline{\theta}}(CS|R_j(G)), \quad j = 1,3,$$

and further, for the procedure $R_3(G)$,

$$(3.3.2) \qquad \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R_3(G)) = \inf_{\underline{\theta}\in\Omega_0} P_{\underline{\theta}}(CS|R_3(G)),$$

where $\Omega_k = \{\underline{\theta} \in \Omega | \theta_{[k-1]} = \theta_{[k]}\}$ and $\Omega_0 = \{\underline{\theta} \in \Omega | \theta_{[1]} = \ldots = \theta_{[k]}\}$

Remark: When $\underline{\theta} \in \Omega_0$, procedures $R_1(G)$ and $R_3(G)$ are distribution-free in the sense that the distributions of the statistics $\max_{1\le j\le k} H_j - H_i$ and H_i do not depend upon the underlying distribution $F(\cdot|\theta)$.

In general, the least favorable configuration (LFC) of the rule $R_1(G)$ is unknown except for $k = 2$; however, it is known (see Rizvi and Woodworth (1970)) that the LFC need not occur in Ω_0. In order to compare rules $R_1(G)$ and $R_3(G)$, for various choices of G, the constants d and D are chosen to yield approximately the same P* when $\underline{\theta} \in \Omega_0$. The ratio $EFF(R) \equiv P(CS|R)/E(S|R)$ is used to compare the rules, where $E(S|R)$ is the expected size of the subset selected.

Now, taking G to be a symmetric lambda distribution with location parameter α, scale parameter ϱ and shape parameter γ, for

$\underline{\theta} \in \Omega_0$, we have the following:

$$(3.3.3)\quad a(r) = E(T(r)|G)$$

$$= \alpha + \frac{\Gamma(N+1)}{B\Gamma(r)\Gamma(N-r+1)} \left\{ \frac{\Gamma(r+\gamma)\Gamma(N-r+1)-\Gamma(r)\Gamma(N+\gamma-r+1)}{\Gamma(N+\gamma+1)} \right\},$$

$$(3.3.4)\quad \sum_{r=1}^{N} a(r) = \alpha N,$$

and

$$(3.3.5)\quad \sum_{i=1}^{k} H_i = \alpha k.$$

Now, let $a(r) = \alpha+\xi_r$. When $N = 2m+1$, $m \geq 0$, we have from (3.3.3)

$$\xi_{2m+1} = -\xi_1, \ldots, \xi_{m+2} = -\xi_m, \ \xi_{m+1} = 0.$$

In this case, we obtain

$$(3.3.6)\quad E(H_i) = \alpha,$$

$$(3.3.7)\quad n^2 Var(H_i) = \frac{2N(k-1)}{k^2(N-1)} \sum_{j=m+2}^{N} \xi_j^2,$$

$$(3.3.8)\quad n^2 Cov(H_i,H_j) = -\frac{-2\sum_{j=m+2}^{N} \xi_j^2}{k(N-1)} - \frac{\alpha^2 N(n-1)}{k},$$

and

$$(3.3.9)\quad -\frac{1}{k-1} \leq Cov(H_i,H_j) < 0.$$

On the other hand, when $N = 2m$, $m > 0$, we get

$$\xi_{2m} = -\xi_1, \ldots, \xi_{m+1} = -\xi_m.$$

Consequently, in this case also we obtain results (3.3.6) through (3.3.9) except that the summations in (3.3.7) and (3.3.8) will be from m+1 to N instead of m+2 to N.

Gupta and McDonald (1970) derived the exact distribution of $\max_{1\le j\le k} H_j - H_i$ for the case of $a(R_{ij}) = R_{ij}$ for k = 3 and n = 2(1)5. Also, for $a(R_{ij}) = R_{ij}$, H_i is the well-known Mann-Whitney U-statistic. But in general the distribution of $\max_{1\le j\le k} H_j - H_i$ is not known since it depends on G. However, with a(r) defined as in (3.3.3), for k = 3 and $d \ge 0$,

$$\Pr\{\max_{1\le j\le 3} H_j - H_i \le d\} = \Pr\{H_2-H_1 \le d,\ H_3-H_1 \le d\}$$

can be evaluated on the computer. Without loss of generality, one can assume that $\alpha = 0$. Table III.1, Table III.2, and Table III.3 provide, respectively, the values of a(r), d-values for the procedure $R_1(G)$, and D-values for the rule $R_3(G)$, respectively, for k = 3, n = 3,5, and (β, γ) = (0.57735, 1.00000), (0.19745, 0.13491), (-0.0006589, -0.0003630), (-0.16857, -0.080199). In Tables III.2 and III.3, we choose P* = 0.75, 0.90, 0.95, 0.975 and 0.99. The four choices of (β, γ) specified above correspond to the cases where the lambda distribution can be used to approximate uniform, normal, logistic and double exponential distributions, respectively, each with mean 0 and variance 1. Accordingly, these choices are denoted in the tables by U, N, L, and D, respectively.

Finally, we briefly discuss how approximate values of d and D can be obtained using asymptotic theory.

Theorem 3.3.2. For $\underline{\theta} \in \Omega_0$ and for the rule $R_1(G)$,

$$P(CS|R_1(G)) \approx \int_{-\infty}^{\infty} \phi^{k-1}(x + \frac{nd}{v})\, d\phi(x),$$

where $v^2 = Var(H_i) - C_v$, C_v is common covariance between H_i and H_j for $i \neq j$, and $\phi(x)$ is the cdf of a standard normal distribution.

Proof. By checking Lindeberg's condition, one can show that $nH_i/\sqrt{Var(H_i)-C_v}$ is asymptotically normally distributed. Hence the result follows.

The value of d satisfying

$$\int_{-\infty}^{\infty} \phi^{k-1}(x + \frac{nd}{v})\, d\phi(x) = P^*$$

can be obtained from the tables of Gupta (1963), Gupta, Nagel and Panchapakesan (1969) or Gupta, Panchapakesan and Sohn (1985), who have tabulated $h = nd/\sqrt{2}v$.

Similarly the following theorem holds for the rule $R_3(G)$.

Theorem 3.3.3. For $\underline{\theta} \in \Omega_0$ and $N = 2m+1$,

$$P(CS|R_3(G)) \approx \phi^k(\frac{D}{nw}),$$

where $w^2 = \frac{2(k-1)}{nk(kn-1)} \sum_{j=n+2}^{kn} \xi_i^2$.

Proof. Proof is analogous to that of Theorem 3.3.2 and hence omitted.

From the above theorem, we have $D \approx \phi^{-1}(nwP^{*1/k})$.

3.3.2 Evaluation of Constants for $R_1(G)$ and $R_3(G)$ using scores $a_0^*(r)$.

In this section, we use a score function $a_0^*(r)$ (to be defined later) in the rules $R_1(G)$ and $R_3(G)$ and evaluate the associated constants d and D.

In order to define the scores $a_0^*(r)$, consider the density $d(x,\theta)$, $\theta \in \Theta$, on an interval containing the origin, satisfying the following regularity conditions.

(i) $d(x,\theta)$ is absolutely continuous in θ for almost every x:

(ii) the limit

$$\dot{d}(x,0) = \lim_{\theta \to 0} \frac{1}{\theta} [d(x,\theta) - d(x,0)]$$

exists for almost every x:

(iii) $$\lim_{\theta \to 0} \int_{-\infty}^{\infty} |\dot{d}(x,\theta)| dx = \int_{-\infty}^{\infty} |\dot{d}(x,0)| dx < \infty$$

holds, with $\dot{d}(x,\theta)$ denoting the partial derivative with respect to θ. Note that the existence of $\dot{d}(x,\theta)$ for almost every θ is insured at every point x such that $d(x,\theta)$ is absolutely continuous in θ. This, however, does not make the condition (ii) superfluous.

In deriving locally most powerful tests for equality of location Gupta, Huang and Nagel (1979) used the score function $a_0^*(r)$ defined by

$$a_0^*(r) = E\left\{ \frac{\dot{d}(X_N^{(r)},0)}{d(X_N^{(r)},0)} \right\}, \tag{3.3.10}$$

where $X_N^{(r)}$ denotes the r-th order statistic in a sample of size N from the distribution with density $d(x,0)$. For the location parameter case, $a_0^*(r)$ can be written as

$$(3.3.11) \qquad a_0^*(r) = E\left\{\frac{\dot{f}(F^{-1}(U_{(r)},0),0)}{f(F^{-1}(U_{(r)},0),0)}\right\},$$

where $U_{(r)}$ denotes the r-th order statistic in a sample of size N from the uniform distribution. Now, specifying $d(x,\theta)$ to be the symmetric lambda density with parameters θ(location), γ(scale) and (scale), we obtain

$$a_0^*(r) = \begin{cases} \int_0^1 N\binom{N-1}{r-1} \frac{\beta(\gamma-1)u^{r-1}(1-u)^{N-r}(u^{\gamma}-(1-u)^{\gamma-2})}{\gamma^2(u^{\gamma-1}+(1-u)^{\gamma-1})^2}\,du, & \beta \geq 0, \\ \int_0^1 N\binom{N-1}{r-1} \frac{(-\beta)(\gamma-1)u^{r-1}(1-u)^{N-r}(u^{\gamma-1}-(1-u)^{\gamma-2})}{\gamma^2(u^{\gamma-1}+(1-u)^{\gamma-1})^2}\,du, & \beta<0. \end{cases}$$

For k = 3, n = 3,5, and selected values of (β,γ) which were denoted by U, N, L and D earlier in Section 3.3.2, the values of $a_0^*(r)$ are tabulated in Table III.4. For the same values of k, n and (β,γ), the constants d and D are given in Tables III.5 and III.6, respectively, with P* = 0.75, 0.90, 0.95, 0.975, 0.99 in each case.

Remark: Nagel (1970) and Gupta, Huang and Nagel (1979) have derived locally optimal subset selection procedures. It follows from their results that the rule $R_3(G)$ is locally optimal in the sense that the rule maximizes the PCS in a neighborhood of any $\underline{\theta} \in \Omega_0$ among all rules which satisfy $\inf_{\underline{\theta} \in \Omega_0} P(CS|R) = P^*$.

3.3.3. Comparisons of the Procedures $R_1(G)$ and $R_3(G)$.

As we have stated in Section 3.3.1, the procedures $R_1(G)$ and $R_3(G)$ are compared in terms of EFF(R), which is used as a measure of efficiency. A large value indicates high efficiency.

For a proper comparison of the two procedures, we should have the constants d and D such that the two procedures will have the PCS approximately equal to P* for $\underline{\theta} \in \Omega_0$. In our Monte Carlo studies with k=3, this led to the choice of P* = .90, 0.95, 0.975 for n=3, and P* = 0.75, 0.90, 0.95, 0.975 for n=5. Further, we considered normal, logistic, and double exponential distributions all with variance 1, as three possible choices of the underlying distributions. Let θ_1, θ_2, θ_3 be the means of the three populations π_1, π_2, π_3. We considered four different configurations of $\underline{\theta} = (\theta_1,\theta_2,\theta_3)$, namely,

I: $\underline{\theta} = (0,0,0.1)$, II: $\underline{\theta} = (0,0,0.5)$,

III: $\underline{\theta} = (0,0,1)$, IV: $\underline{\theta} = (0,0.5,1.0)$.

For comparisons using the score function a(r), we chose the four choices of the parameter (β,γ) of the lambda distribution, referred to by U, N, L, and D in Section 3.3.1. For comparisons using $a_0^*(r)$, the choice of (β,γ), denoted by UD, is made so that the lambda distribution can be used to approximate the underlying distributions with variance 1.

For each choice of the underlying distribution, random samples were generated by using the random number generator RVP, developed by Professor Rubin at Purdue University. Our results are based on 1000 simulations in the case of n = 3 and 500 simulations in

the case of n=5. Table III.7 is reproduced for the cases where the underlying distributions are normal and logistic distributions with the mean configuration II for $(n,P^*) = (3,0.90)$; the patterns in the other case are similar.

Besides comparing the efficiencies of the rules $R_1(G)$ and $R_3(G)$ under each choice of G, we are also interested in comparing the different choices of G for each rule. Based on the Monte Carlo study, our conclusions are summarized below.

(1) When the means are close to each other, no rule performs uniformly better than the other when the underlying distributions are normal or double exponential; however, as $P^* \to 1$, the rule $R_3(G)$ performs slightly better than the rule $R_1(G)$. With means close to each other, the situation changes when the underlying distributions are uniform or logistic: Then, the rule $R_3(G)$ performs almost uniformly better than the rule $R_1(G)$.

(2) When the largest mean is sufficiently away from the next largest, the rule $R_1(G)$ generally performs better than the rule $R_3(G)$ no matter what the choice of G is. This behavior becomes more clear as n increases. Also, when P^* is close to 1, the difference in the performances of the two rules narrows down, even though $R_1(G)$ still is better.

(3) Generally, the rule $R_1(G)$ performs better than the rule $R_3(G)$ when the choices of G are the lambda distribution to be the uniform and the underlying distribution F (i.e., G is U or UD) both with variance 1.

(4) Considering the efficiency of the procedure $R_1(G)$, the best choice of G is the lambda distribution which approximates the uniform

distribution with unit variance (i.e., G is U).

(5) For the rule $R_3(G)$, the best choice of G is the lambda distribution approximating the underlying distribution with unit variance. This is all the more clear when the underlying distributions are normal or double exponential with their means close to each other.

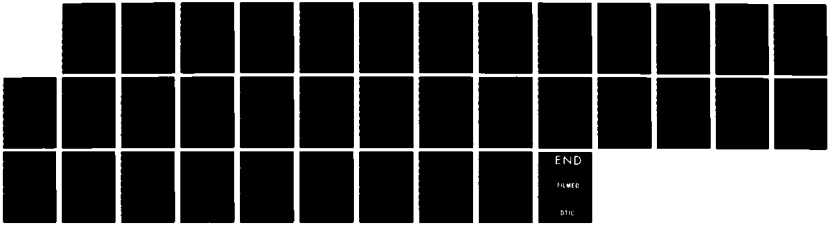
Considering all the findings of the study, the overall recommendations will be:

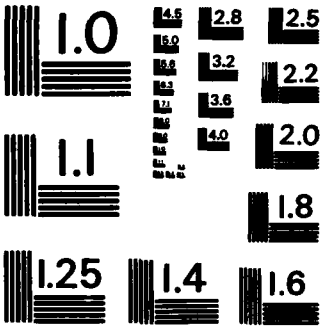
(1) When the means of the underlying distributions are expected to be close to each other, use either the rule $R_1(G)$ with U as the choice for G or the rule $R_3(G)$ with UD as the choice for G.

(2) When the largest mean is expected to be sufficiently away from the next largest, use the rule $R_1(G)$ with U as the choice for G.

AD-A159 153 MULTIPLE DECISION PROCEDURES FOR TUKEY'S GENERALIZED LAMBDA DISTRIBUTIONS(U) PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS J K SOHN AUG 85 TR-85-20 2/2

UNCLASSIFIED N00014-84-C-0167 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Table III.1

Values of a(r) under Ω_0 for k=3,
where $\Omega_0 = \{\underline{\theta} \in \Omega | \theta_1 = \theta_2 = \theta_3\}$

n	a(r)	U	N	L	D
3	a(9)	1.38552	1.48669	1.49804	1.49582
	a(8)	1.03914	0.93118	0.87778	0.83529
	a(7)	0.69276	0.57013	0.52348	0.48933
	a(6)	0.34638	0.27334	0.24800	0.22992
	a(5)	0.	0.	0.	0.
5	a(15)	1.51541	1.73896	1.79233	1.81764
	a(14)	1.29893	1.24834	1.20149	1.15927
	a(13)	1.08240	0.94605	0.88346	0.83506
	a(12)	0.86595	0.71257	0.65382	0.61080
	a(11)	0.64936	0.51350	0.46595	0.43213
	a(10)	0.43298	0.33363	0.30065	0.27756
	a(9)	0.21649	0.16441	0.14759	0.13591
	a(8)	0.	0.	0.	0.

Note For n=3, a(i) = -a(10-i), i=1,...,4 and for n=5, a(i) = -a(16-i), i=1,...,7.

Table III.2

d-values of the procedure $R_1(G)$ under

$\Omega_0 = \{\underline{\theta} \in \Omega | \theta_1 = \theta_2 = \theta_3\}$ for k=3

		U		N		L		D	
P*	n	3	5	3	5	3	5	3	5
0.75		2.423	3.156	2.431	3.173	2.402	3.135	2.388	3.094
0.90		3.809	4.887	3.644	4.825	3.597	4.750	3.538	4.684
0.95		4.501	5.843	4.264	5.744	4.227	5.648	4.114	5.556
0.975		4.848	6.619	4.747	6.490	4.644	6.370	4.545	6.249
0.99		5.194	7.485	5.131	7.288	5.026	7.124	4.920	6.984

Table III.3

D-values of the rule $R_3(G)$ for k=3
under $\Omega_0 = \{\underline{\theta} \in \Omega | \theta_1=\theta_2=\theta_3\}$

		U		N		L		D	
P*	n	3	5	3	5	3	5	3	5
0.75		-1.03914	-1.29883	-0.93118	-1.22607	-0.87778	-1.20182	-0.83529	-1.17884
0.90		-1.73190	-2.38126	-1.77465	-2.28712	-1.74604	-2.25795	-1.72574	-2.22446
0.95		-2.07828	-2.81444	-2.14453	-2.87711	-2.12782	-2.84109	-2.10119	-2.79435
0.975		-2.42466	-3.46370	-2.41787	-3.36773	-2.37582	-3.31218	-2.33111	-3.26350
0.99		-3.11742	-2.89678	-2.98800	-3.89562	-2.89930	-3.80070	-2.82044	-3.72936

Table III.4

Values of $a_0^*(r)$ for some values of (β,γ) and n=3,5

n	$a_0^*(r)$	N	L	D
3	$a_0^*(9)$	10.95367	3997.81042	18.30010
	$a_0^*(8)$	6.96000	2999.62692	14.97188
	$a_0^*(7)$	4.31341	2000.12774	10.39644
	$a_0^*(6)$	2.08126	1000.15792	5.30546
	$a_0^*(5)$	0.0	0.0	0.0
5	$a_0^*(15)$	12.76184	4371.83812	19.28142
	$a_0^*(14)$	9.24459	3748.92094	18.05537
	$a_0^*(13)$	7.09456	3124.76751	15.75637
	$a_0^*(12)$	5.39158	2500.15400	12.98432
	$a_0^*(11)$	3.90891	1875.28911	9.93465
	$a_0^*(10)$	2.54966	1250.26286	6.71000
	$a_0^*(9)$	1.25921	625.15168	3.38000
	$a_0^*(8)$	0.0	0.0	0.0

Note n=3, $a_0^*(1) = -a_0^*(9),\ldots,a_0^*(4) = -a_0^*(6)$. Also for n=5, $a_0^*(1) = -a_0^*(15),\ldots,a_0^*(7) = -a_0^*(9)$.

Table III.5

Values of d of the rule

$R_1(G)$ with $a_0^*(r)$

n	P*	N	L	D
3	0.75	18.05	6996.0	33.47
	0.90	27.09	10994.0	52.07
	0.95	31.50	12994.0	62.17
	0.975	35.35	13996.0	69.02
	0.99	38.04	14996.0	74.33
5	0.75	23.51	9368.0	44.30
	0.90	35.74	14365.0	67.90
	0.95	42.59	16868.0	81.24
	0.975	48.15	19365.0	92.21
	0.99	54.10	21865.0	104.62

Table III.5

Values of d of the rule

$R_1(G)$ with $a_0^*(r)$

n	P*	N	L	D
3	0.75	18.05	6996.0	33.47
	0.90	27.09	10994.0	52.07
	0.95	31.50	12994.0	62.17
	0.975	35.35	13996.0	69.02
	0.99	38.04	14996.0	74.33
5	0.75	23.51	9368.0	44.30
	0.90	35.74	14365.0	67.90
	0.95	42.59	16868.0	81.24
	0.975	48.15	19365.0	92.21
	0.99	54.10	21865.0	104.62

Table III.6

Values of D of the rule

$R_3(G)$ with $a_0^*(r)$

n	P*	N	L	D
3	0.75	-6.96000	-2997.84061	-13.20912
	0.90	-13.18583	-4997.96834	-23.60556
	0.95	-15.83242	-5999.91257	-30.67377
	0.975	-17.91368	-6998.09608	-34.00199
	0.99	-22.22709	-8997.56508	-43.66842
5	0.75	-9.20044	-3746.81187	-16.98243
	0.90	-16.89421	-6870.99386	-32.07069
	0.95	-21.33906	-8125.32180	-40.66679
	0.975	-25.02452	-9996.04816	-47.27144
	0.99	-29.05684	-11249.13155	-56.14283

Table III.7

Comparisons of the Procedures $R_1(G)$ and $R_3(G)$

under the configuration $\underline{\theta} = (0,0,0.5)$ and $P^* = 0.90$

(a) n=3

Underlying distribution	Normal					Logistic				
G	U	N	L	D	UD	U	N	L	D	UD
$P(CS\|R_1(G))$	0.969 (0.005)	0.971 (0.005)	0.971 (0.005)	0.971 (0.005)	0.971 (0.005)	0.927 (0.008)	0.939 (0.008)	0.940 (0.008)	0.940 (0.008)	0.927 (0.008)
$P(CS\|R_3(G))$	0.985 (0.004)	0.975 (0.005)	0.975 (0.005)	0.975 (0.005)	0.973 (0.005)	0.947 (0.007)	0.947 (0.007)	0.947 (0.007)	0.947 (0.007)	0.937 (0.008)
$E(S\|R_1(G))$	2.583 (0.019)	2.607 (0.018)	2.604 (0.018)	2.604 (0.018)	2.607 (0.018)	2.668 (0.018)	2.704 (0.017)	2.699 (0.017)	2.699 (0.017)	2.668 (0.018)
$E(S\|R_3(G))$	2.712 (0.014)	2.658 (0.015)	2.658 (0.015)	2.658 (0.015)	2.627 (0.015)	2.753 (0.014)	2.726 (0.014)	2.726 (0.014)	2.726 (0.014)	2.719 (0.014)
$EFF(R_1(G))$	0.400 (0.005)	0.394 (0.005)	0.393 (0.005)	0.393 (0.005)	0.394 (0.005)	0.357 (0.005)	0.355 (0.004)	0.355 (0.004)	0.355 (0.004)	0.357 (0.005)
$EFF(R_3(G))$	0.374 (0.003)	0.378 (0.003)	0.378 (0.003)	0.378 (0.003)	0.382 (0.003)	0.348 (0.003)	0.353 (0.003)	0.353 (0.003)	0.753 (0.003)	0.349 (0.004)

Table III.7 (continued)

(b) n=5

Underlying distribution	Normal					Logistic				
G	U	N	L	D	UD	U	N	L	D	UD
$P(CS\|R_1(G))$	0.988 (0.005)	0.984 (0.006)	0.986 (0.005)	0.986 (0.005)	0.984 (0.006)	0.952 (0.010)	0.952 (0.010)	0.946 (0.010)	0.948 (0.010)	0.952 (0.010)
$P(CS\|R_3(G))$	0.990 (0.004)	0.986 (0.005)	0.986 (0.005)	0.988 (0.005)	0.986 (0.005)	0.954 (0.009)	0.948 (0.010)	0.948 (0.010)	0.950 (0.010)	0.952 (0.010)
$E(S\|R_1(G))$	2.528 (0.029)	2.534 (0.022)	2.542 (0.022)	2.546 (0.022)	2.532 (0.022)	2.732 (0.023)	2.726 (0.023)	2.726 (0.023)	2.734 (0.022)	2.732 (0.023)
$E(S\|R_3(G))$	2.590 (0.022)	2.584 (0.022)	2.604 (0.022)	2.612 (0.022)	2.586 (0.022)	2.726 (0.020)	2.716 (0.020)	2.710 (0.020)	2.710 (0.020)	2.712 (0.020)
$EFF(R_1(G))$	0.431 (0.008)	0.427 (0.008)	0.426 (0.008)	0.425 (0.008)	0.428 (0.008)	0.359 (0.006)	0.361 (0.006)	0.357 (0.006)	0.357 (0.006)	0.359 (0.006)
$EFF(R_3(G))$	0.397 (0.004)	0.396 (0.004)	0.392 (0.004)	0.392 (0.004)	0.396 (0.004)	0.356 (0.005)	0.355 (0.005)	0.356 (0.005)	0.356 (0.005)	0.357 (0.005)

CHAPTER IV

A TWO-STAGE PROCEDURE FOR SELECTING THE BEST AMONG GOOD POPULATIONS

4.1. Introduction

Since the early work of Bechhofer, Dunnett and Sobel (1954) on the two-sample (two-stage) problem for selecting the population associated with the largest unknown mean from k ($\geq$ 2) normal populations, several types of two-stage procedures have been studied. Among them elimination type procedures, which select a subset of populations of interests at stage 1 and finally select the best population at stage 2, are important. Under the non-Bayesian formulation Alam (1970) has studied the known variances case and Tamhane and Bechhofer (1977, 1979), using a minimax criterion, also have studied the known variances case. Gupta and Kim (1984) and Tamhane (1975) have considered the common unknown variance case. Recently Gupta and Miescke (1982, 1983), among others, have studied the problem under the decision-theoretic Bayesian framework.

In this chapter, we propose an elimination type procedure under the Bayesian setting. At stage 1 we use a noninformative prior for unknown parameters. To select the best population at stage 2, we use a stopping rule to construct a $100(1-2\alpha)\%$ Highest Posterior Density (HPD) credible region with a common width 2d.

In Section 4.2 we give notations and definitions including the definition of the $100(1-2\alpha)\%$ HPD credible region.

In Section 4.3 we propose a procedure $R(\alpha,d)$ which selects the best after retaining a subset of populations at stage 1 and investigate its properties.

4.2. Framework

Let $\pi_1,\ldots,\pi_k$ be k independent normal populations with unknown means $\theta_1,\ldots,\theta_k$, respectively and unknown common variance σ^2 $(0 < \sigma^2 < \infty)$. Also let a random variable X_i be the observable characteristic associated with π_i. For $i = 1,2,\ldots,k$, let $\underline{X}_i' = (X_{i1},\ldots,X_{in})$ be a vector of n independent observations from π_i, $i = 1,2,\ldots,k$, respectively. Assuming that very little is known to the experimenter about the prior distribution of $(\theta_1,\theta_2,\ldots,\theta_k, \sigma^2)$, we may use a locally uniform joint prior density $\tau(\theta_1,\theta_2,\ldots,\theta_k,\sigma^2) = \sigma^{-2}I_{(0,\infty)}(\sigma^2)$, which is also a noninformative prior for the model, where $I_A(x)$ is the usual indicator function. Let $\tau_1(\theta_1,\ldots,\theta_k|\underline{X}_1,\ldots,\underline{X}_k)$ be the marginal joint posterior distribution of $\underline{\theta}' = (\theta_1,\ldots,\theta_k)$ given $\underline{X}' = (\underline{X}_1,\ldots,\underline{X}_k)$.

π_i is said to be 'good' ('bad') if $\theta_i \geq \theta_0$ $(\theta_i < \theta_0)$, where θ_0 is a control or standard which is specified a priori by the experimenter. Let $\underline{\delta}^{(1)}(\underline{X}) = (\delta_1^{(1)}(\underline{X}_1),\ldots,\delta_k^{(1)}(\underline{X}_k))$, where $\delta_i^{(1)}(\underline{X}_i)$ is a nonrandomized decision rule for π_i at stage 1, i.e., $\delta_i^{(1)}(\underline{X}_i) = 1$ if π_i is accepted as a good population and $\delta_i^{(1)}(\underline{X}_i) = 0$ if π_i is rejected as a bad one. Let the loss function

$L^{(1)}(\underline{\theta}, \underline{\delta}^{(1)}(\underline{X}))$ at stage 1 be as follow:

$$(4.2.1) \qquad L^{(1)}(\underline{\theta},\underline{\delta}^{(1)}(\underline{X})) = \sum_{i=1}^{k} L_i^{(1)}(\theta_i,\delta_i^{(1)}(\underline{X}_i)),$$

where $L_i^{(1)}(\theta_i, \delta_i^{(1)}(\underline{X}_i))$ is loss due to the decision $\delta_i^{(1)}(\underline{X}_i)$ about π_i such that

$$(4.2.2) \qquad L_i^{(1)}(\theta_i,\delta_i^{(1)}(\underline{X}_i)) = \begin{cases} k_0 & \text{if } \delta_i^{(1)}(\underline{X}_i) = 1 \quad \text{and } \theta_i \le \theta_0 \\ k_1 & \text{if } \delta_i^{(1)}(\underline{X}_i) = 0 \quad \text{and } \theta_i > \theta_0 \\ 0 & \text{otherwise,} \end{cases}$$

in other words, a loss due to selecting each bad population is k_0 and a loss due to rejecting each good population is k_1.

Remarks:

One might question the suitability of a loss of this kind in this problem. However, a loss function of this kind can be proper for the two-component decision problems, because the loss function of this kind can reflect the importance of two types of possible misclassification errors. For our situation, at stage 1, we 'only' want to classify populations into possible good and bad populations. Thus at stage 1 our problem can be regarded as the k two-component decision problems. Problems of this type have been investigated by Lehmann (1957).

Let our final nonrandomized decision $\delta^{(2)}(\underline{Y})$ at stage 2 be $\delta^{(2)}(\underline{Y}) = \{j: j \in S\}$, where $\underline{Y}' = (\underline{Y}_1,\ldots,\underline{Y}_s)$ are combined samples from stage 1 and stage 2 for populations in S where S is a selected

subset at stage 1 with size s. Let a loss due to the decision $\delta^{(2)}(\underline{Y})$ be

$$(4.2.3) \qquad L^{(2)}(\underline{\theta}, \delta^{(2)}(\underline{Y})) = I\{\theta_j \neq \theta_{[k]}\},$$

Now we give the definition of the $100(1-2\alpha)\%$ HPD credible region which we will use at stage 2.

Let $\tau_1(\theta|\underline{X})$ be the marginal posterior density of θ given $\underline{X}$.

Definition 4.2.1 (see Berger (1980)). The $100(1-2\alpha)\%$ HPD credible region for θ is the subset $C_{(1-2\alpha)}$ of the parameter space Θ of the form

$$(4.2.4) \qquad C_{(1-2\alpha)} = \{\theta \in \Theta;\ \tau_1(\theta|\underline{X} = \underline{x}) \geq \xi_{2\alpha}\},$$

where $\xi_{2\alpha}$ is the largest constant such that

$$(4.2.5) \qquad Pr(C_{(1-2\alpha)}|\underline{X} = \underline{x}) \geq 1-2\alpha.$$

Remark:

If $\tau_1(\theta|\underline{X})$ is not unimodal, then the credible region $C_{(1-2\alpha)}$ may consist of several disjoint intervals.

4.3. Goal and a Proposed Procedure R(α,d).

Assume that no knowledge is available concerning the correct pairing between populations and the ordered θ_i's. Our goal is to select the population associated with the largest unknown mean, if any, from the set of good populations. The procedure R(α,d) is designed to meet the goal.

4.3.1. Definition of the Procedure R(α,d).

Stage 1. Take $n_0 = \max\{2, [Z_{(1-\alpha)}/d] + 1\}$ observations from each population π_i, where $Z_{(1-\alpha)}$ is the $100(1-\alpha)$ percentile of the standard normal distribution and [a] is the largest integer $\leq a$. Note that 2d corresponds to the width of the $100(1-2\alpha)$% HPD credible region for θ, which is to be specified by the experimenter.

Now based on first stage samples, we select a subset S by the following rule.

At stage 1, for $i = 1,2,\ldots,k$, $\delta_i^{(1)}(\underline{X}_i) = 1$ if and only if

$$G_\nu\left(\frac{\theta_0-\bar{X}_i}{V}\right) \leq \frac{k_1}{k_0+k_1},$$

where $G_\nu(\cdot)$ is the cdf of a Student's t distribution with $\nu = k(n_0-1)$ degrees of freedom, $\bar{X}_i = \sum_{j=1}^{n_0} X_{ij}/n_0$ and

$$V^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_0} (X_{ij}-\bar{X}_i)^2/kn_0(n_0-1).$$

Now with a selected subset S with its size s,

(1) if s = 0, we decide that none of the populations are good and stop,

(2) if s = 1, we decide that the population selected is the only good one and hence it is the best and stop.

(3) if $s \geq 2$, we proceed to stage 2.

Stage 2. Take one observation at a time from each population in S till $N-n_0$ observations are taken such that

(4.3.1) $\quad N = \inf\{n: \ n \geq \max\{n_0, [t_\alpha^2 V_1^2/d^2 q]+1\}$,

where t_α is the 100α lower percentile of the Student's t distribution with $q = (k-s)(n_0-1) + s(n-1)$ degrees of freedom and

$$V_1^2 = \sum_{i \notin S} \sum_{j=1}^{n_0} (X_{ij} - \bar{X}_i)^2 + \sum_{i \in S} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_i)^2, \text{ and } \bar{Y}_i = \sum_{j=1}^{n} Y_{ij}/n.$$

Then our final decision at stage 2 is

$$\delta^2(\underline{Y}) = \{j: \ j \in S \text{ and } \bar{Y}_j = \max_{1 \leq \ell \leq s} \bar{Y}_\ell\},$$

that is, to select the population associated with the largest overall sample mean and claim it to be the best population among good populations.

4.3.2. Properties of the Procedure $R(\alpha,d)$.

It is easy to verify that the marginal joint posterior joint density $\tau_1(\theta_1,\ldots,\theta_k|\underline{X}_1,\ldots,\underline{X}_k)$ at stage 1 follows a multivariate t distribution with variance-covariance matrix $W = V^2 I$, where I is a $k \times k$ identity matrix. Hence the marginal posterior density of θ_i given $\underline{X}_1,\ldots,\underline{X}_k$ at stage 1 follows a Student's t distribution with $k(n_0-1)$ degrees of freedom, a location parameter $\bar{X}_i$ and a scale parameter V. Similarly, at stage 2 the marginal posterior density of θ_i of π_i in S given $\{\underline{X}_i, \ i \notin S\}$ and $\underline{Y}$ follows a Student's t distribution with $q = (k-s)(n_0-1) + s(N-1)$ degrees of freedom, a location parameter $\bar{Y}_i$ and a scale parameter Q, where

$$(4.3.2) \qquad Q^2 = \frac{\sum_{i \notin S} \sum_{j=1}^{n_0} (X_{ij} - \bar{X}_i)^2 + \sum_{i \in S} \sum_{j=1}^{N} (Y_{ij} - \bar{Y}_i)^2}{qN}.$$

Hence the following theorem holds.

Theorem 4.3.1. The stopping rule N provides the $100(1-2\alpha)\%$ HPD credible region with a common width 2d for each selected population at stage 1.

Proof. The proof is straightfoward and hence omitted.

Remark:

Since the loss $L^{(1)}(\underline{\theta},\underline{\delta}^{(1)}(\underline{X}))$ at stage 1 is linear and additative, the decision rule $\underline{\delta}^{(1)}(\underline{X})$ is Bayes. This follows from the fact that $E[L_i^{(1)}(\theta_i, \{1\})] = k_0 Pr\{\theta_i \leq \theta_0|\underline{X}\}$ and $E[L_i^{(1)}(\theta_i, \{0\})] = k_1 Pr\{\theta_i > \theta_0|\underline{X}\}$, for $i = 1,\ldots,k$, respectively.

Theorem 4.3.2. Let $n = \sigma^2 Z^2_{(1-\alpha)}/d^2$. Then for a fixed $\sigma^2 (0 < \sigma^2 < \infty)$ and the stopping rule N,

(a) $N/n \to 1$ a.s. as $d \to 0$

and

(b) $\lim_{d\to 0} E(N/n) = 1$ (asymptotic efficiency).

Proof. From the definitions of n_0 and N, one can get the following inequalities;

$$(4.3.3) \qquad \frac{t_\alpha^2 V_1^2}{d^2 q} \leq N \leq \frac{t_\alpha^2 V_1^2}{d^2 q} + \frac{Z_{(1-\alpha)}}{d} + 4.$$

Since $n_0 \to \infty$ and $N \to \infty$ as $d \to 0$ hence $S^2 \to \sigma^2$ a.s.. Thus (a) and (b) follow.

To examine the performance of the procedure $R(\alpha,d)$ a Monte Carlo study was carried out for $k = 5$, $\alpha = 0.025, 0.05$ with 300 simulations. To generate normal random variates with common variance 1, the random number generator RVP developed by Professor Rubin was used. As underlying configurations of means (supposed to be unknown to the experimenter), we chose four different configurations with $d = 0.4$, namely,

(I) $\underline{\theta} = (-0.2,0,0,0.2,0.4)$ (II) $\underline{\theta} = (-0.2,-0.2,0,0.2,0.4)$

(III) $\underline{\theta} = (-0.2,-0.2,0,0,0.2)$ (IV) $\underline{\theta} = (-0.2,-0.2,-0.2,0,0.2)$.

The value of θ_0 was supposed to be 0. As a special case under the configuration (IV), $d = 0.2$ was also chosen and is called configuration V. Basically four statistics were simulated: (a) the expected subset size S at stage 1 (E(S)), (b) the expected value of the overall sample size N (E(N)), (c) the expected loss at stage 1 (E(L1)) and (d) the probability of selecting the population associated with the largest mean (PSB). For the loss function, $(k_0,k_1) = (1,1), (1,2), (2,1), (1,5)$ and $(5,1)$ were considered. The results are shown in several figures, where each figure contains five different configurations for $\alpha = 0.025$. In each of four figures, the abscissa is the ratio k_1/k_0. Thus Figure 1 is E(S) versus k_1/k_0; Figure 2 is E(N); Figure 3 is PSB; and Figure 4 is E(L1). Figures for $\alpha = 0.05$ are similar to these figures drawn for $\alpha = 0.025$ and hence are omitted.

The results indicate:

(1) As k_1/k_0 increases, the values of PSB increases.

(2) In general, the value of E(N) increases as k_1/k_0 increases.

(3) Values of k_1/k_0 are irrelevant to the values of E(L1).

(4) When the number of good populations among five populations decreases, the value of E(S) decreases but the value of E(L1) increases slightly.

(5) When the value of d decreases, the value of PSB increases. But when the overall sample size required and the value of E(S) are taken into consideration, the rule $R(\alpha,d)$ does not provide vast improvement on PSB. This is mainly due to the fact that an elimination-type procedure cannot recover the best population at stage 2 if it has been eliminated at stage 1.

(6) For fixed values of the ratio k_1/k_0, as the distance between the largest mean and the smallest mean increases, the values of PSB increase and the values of E(L1) decrease (slightly).

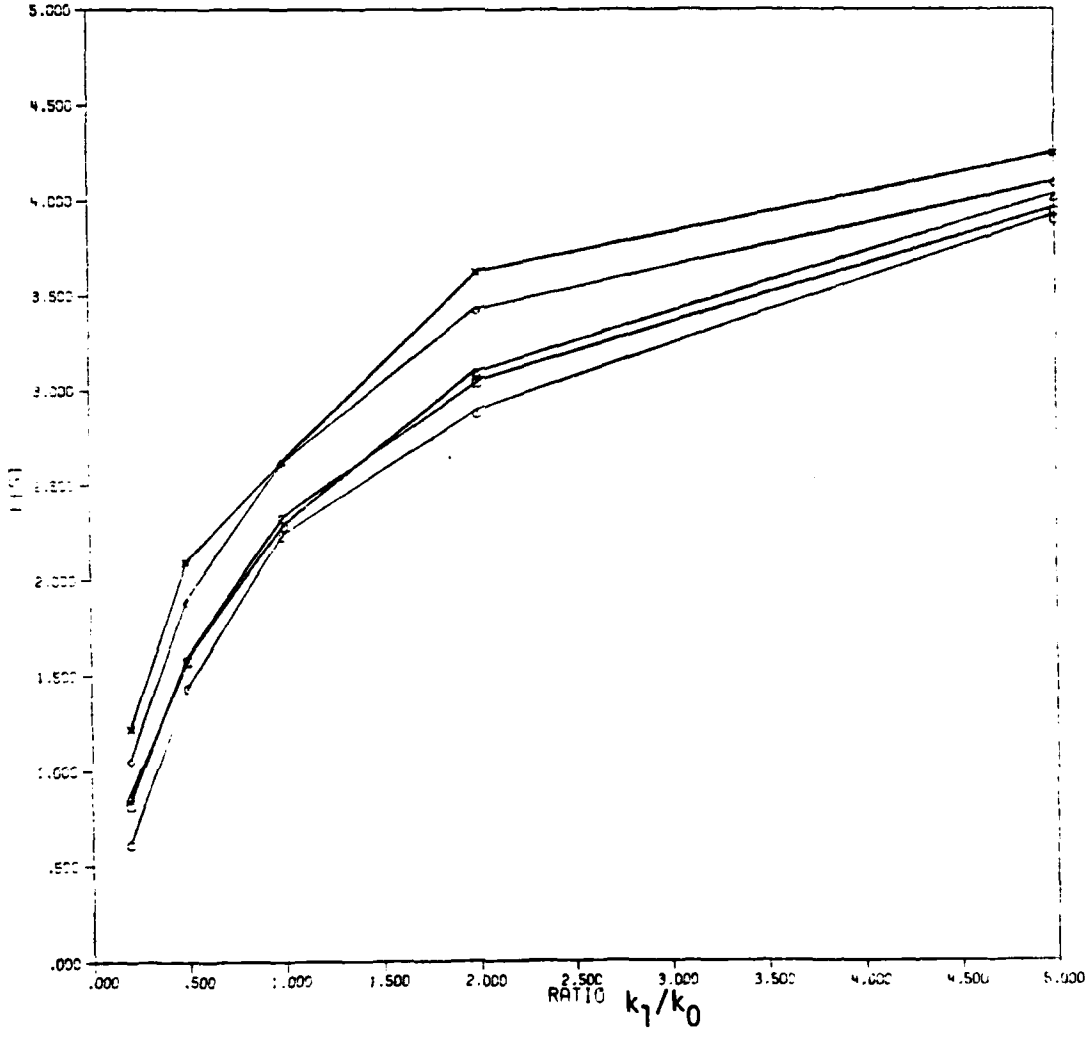


Figure 1. E[S] versus the ratio k_1/k_0 for five configurations.

Legend of Configurations

(I) $\underline{\theta}$ = (-0.2,0,0,0.2,0.4) with d = 0.4

(II) $\underline{\theta}$ = (-0.2,-0.2,0,0.2,0.4) with d = 0.4

(III) $\underline{\theta}$ = (-0.2,-0.2,0,0,0.2) with d = 0.4

(IV) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.4

(V) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.2

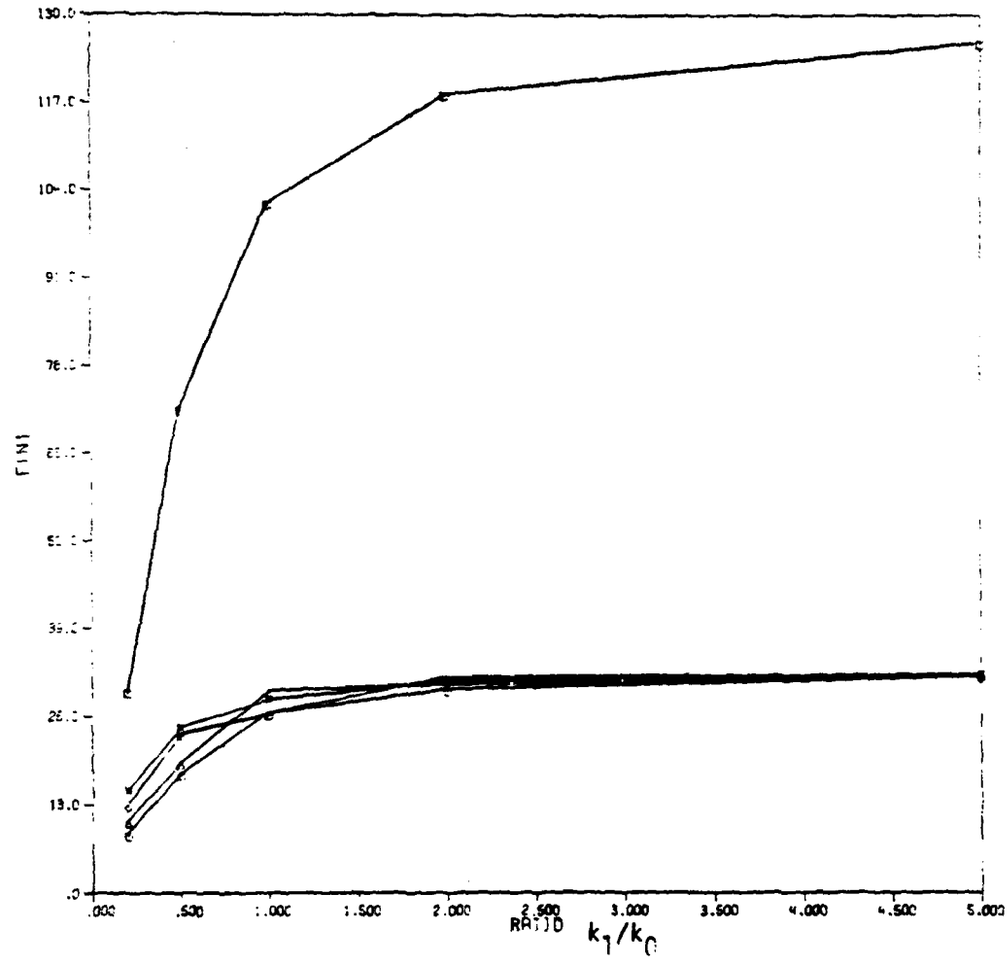


Figure 2. E[N] versus the ratio k_1/k_0 for five configurations.

Legend of Configurations

(I) $\underline{\theta}$ = (-0.2,0,0,0.2,0.4) with d = 0.4

(II) $\underline{\theta}$ = (-0.2,-0.2,0,0.2,0.4) with d = 0.4

(III) $\underline{\theta}$ = (-0.2,-0.2,0,0,0.2) with d = 0.4

(IV) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.4

(V) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.2

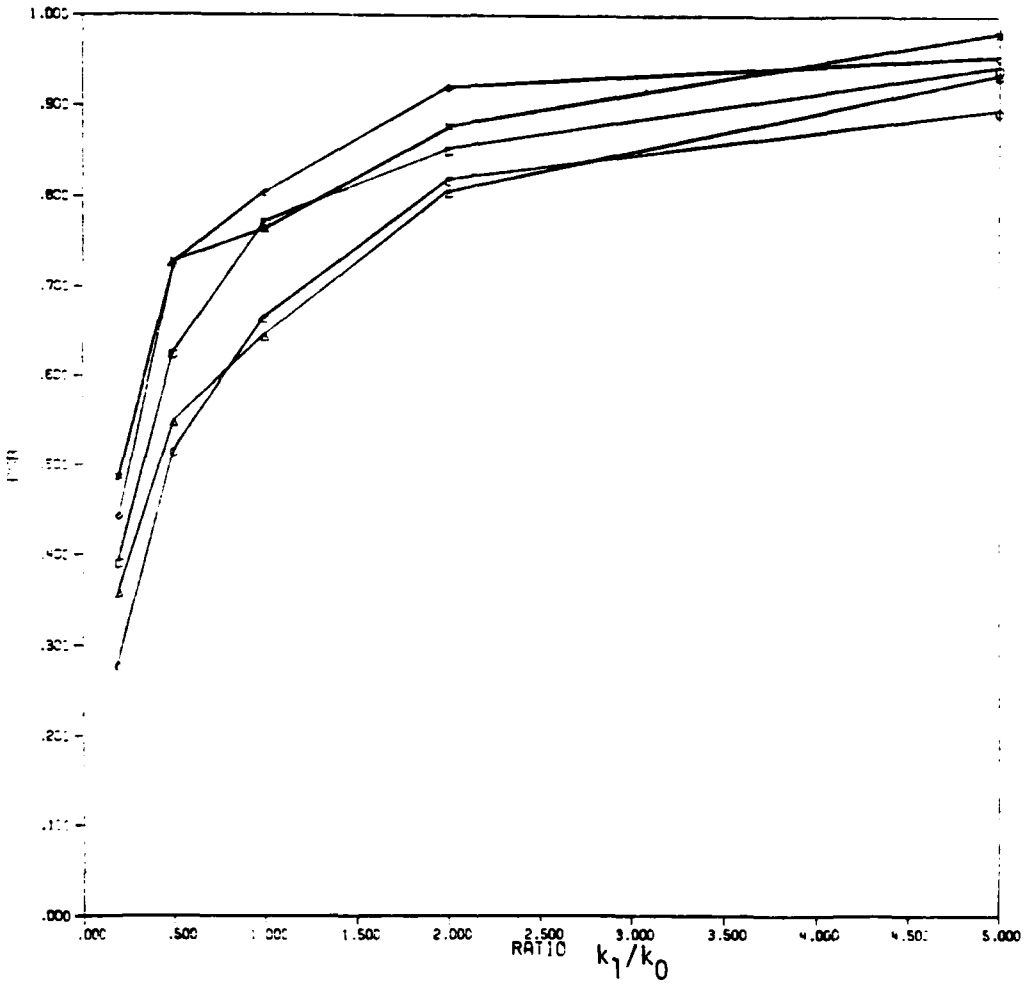


Figure 3. PSB versus the ratio k_1/k_0 for five configurations.

Legend of Configurations

(I) $\underline{\theta}$ = (-0.2,0,0,0.2,0.4) with d = 0.4

(II) $\underline{\theta}$ = (-0.2,-0.2,0,0.2,0.4) with d = 0.4

(III) $\underline{\theta}$ = (-0.2,-0.2,0,0,0.2) with d = 0.4

(IV) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.4

(V) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.2

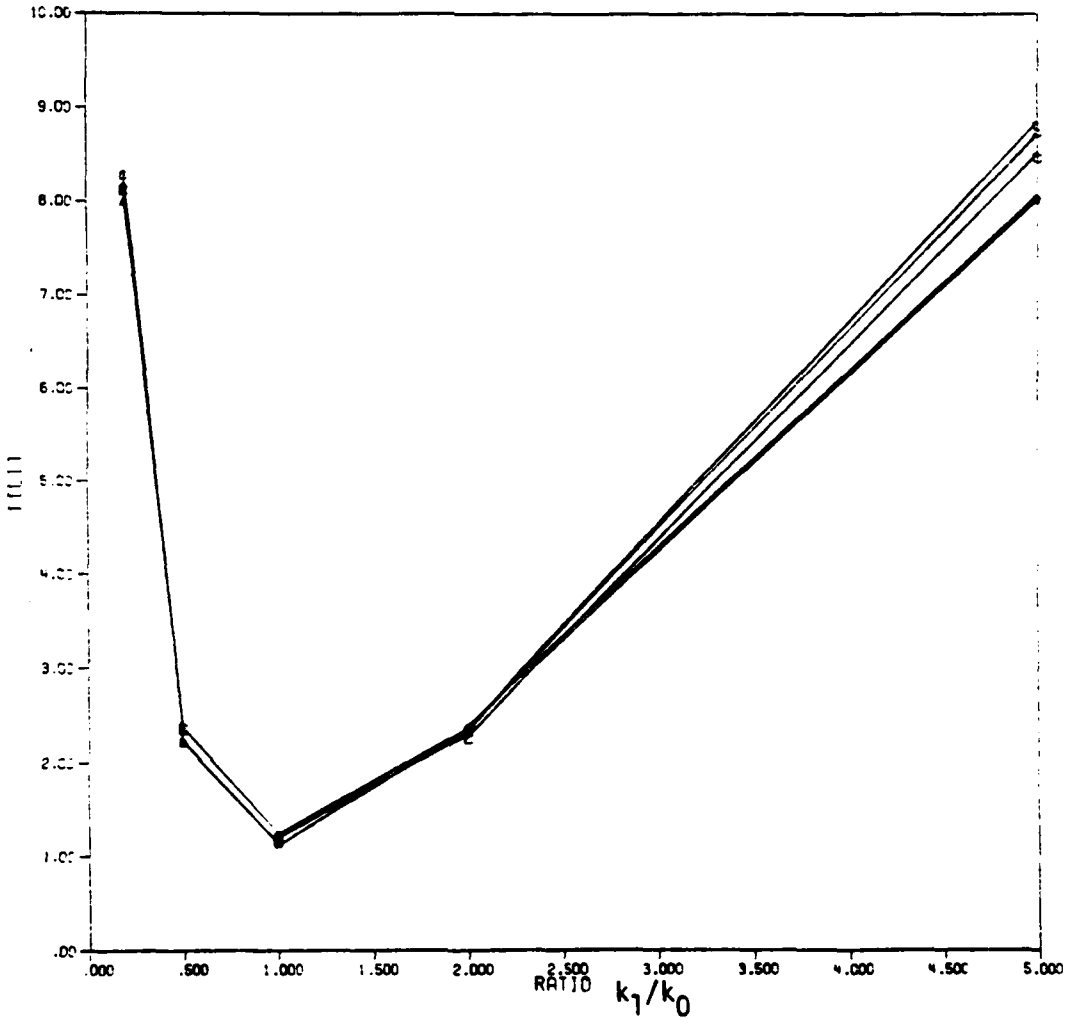


Figure 4. E[L1] versus the ratio k_1/k_0 for five configurations.

Legend of Configurations

(I) $\underline{\theta}$ = (-0.2,0,0,0.2,0.4) with d = 0.4

(II) $\underline{\theta}$ = (-0.2,-0.2,0,0.2,0.4) with d = 0.4

(III) $\underline{\theta}$ = (-0.2,-0.2,0,0,0.2) with d = 0.4

(IV) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.4

(V) $\underline{\theta}$ = (-0.2,-0.2,-0.2,0,0.2) with d = 0.2

BIBLIOGRAPHY

BIBLIOGRAPHY

Alam, K. (1970). A two-sample procedure for selecting the population with the largest mean from k normal populations. Ann. Inst. Statist. Math. 22, 127-136.

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. Ann. Math. Statist. 26, 641-647.

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

Bahadur, R. R., and Robbins, H. (1950). The problem of the greater mean. Ann. Math. Statist. 21, 469-487. Correction, 22 (1951), 310.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference under Order Restrictions. John Wiley & Sons, New York.

Barlow, R. E., and Gupta, S. S. (1969). Selection procedures for restricted families of distributions. Ann. Math. Statist. 40, 905-917.

Bartlett, N. S., and Govindarajulu, Z. (1968). Some distribution-free statistics and their application to the selection problem. Ann. Inst. Statist. Math. 20, 79-97.

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

Bechhofer, R. E. (1968). Multiple comparisons with a control for multiply-classified variances of normal populations. Technometrics 10, 715-718.

Bechhofer, R. E., Dunnett, C. W., and Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. Biometrika 41, 170-176.

Bechhofer, R. E., Kiefer, J., and Sobel, M. (1968). Sequential Identification and Ranking Procedures. The University of Chicago Press, Chicago.

Bechhofer, R. E., and Sobel, M. (1954). A single-sample multiple-decision procedure for ranking variances of normal populations. Ann. Math. Statist. 25, 273-289.

Berger, J. O. (1980). Statistical Decision Theory: Foundations, Concepts, and Methods. Springer-Verlag, New York-Heidelberg-Berlin.

Berkson, J. (1944). Application of the logistic function to bio-assay. J. Amer. Statist. Assoc., 39, 357-365.

Berkson, J. (1951). Why I prefer logits to probits. Biometrics, 7, 327-339.

Berkson, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay and quantal response, based on the logistic function. J. Amer. Statist. Assoc., 48, 565-599.

Bickel, P. J., and Yahav, J. A. (1977). On selecting a subset of good populations. Statistical Decision Theory and Related Topics-II (S. S. Gupta and D. S. Moore, eds.), Academic, New York, 37-55.

Broström, G. (1977). An improved procedure for selecting all populations better than a standard. Tech. Report 1977-3, Inst. of Math. and Statist., Univ. of Umea, Sweden.

Burr, I. W. (1942). Cumulative frequency functions. Ann. Math. Statist., 13, 215-232.

Burr, I. W. (1973). Parameters for a general system of distributions to match a grid of α_3 and α_4. Comm. Statist., 2, 1-21.

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), 685.

Desu, M. M., and Sobel, M. (1971). Nonparametric procedures for selecting fixed-size subsets. Statistical Decision Theory and Related Topics (S. S. Gupta and J. Yackel, eds.), Academic Press, New York, 255-273.

Doksum, K. (1969). Starshaped transformations and the power of rank tests. Ann. Math. Statist. 40, 1167-1176.

Dudewicz, E. J., and Dalal, S. R. (1975). Allocation of observations in ranking and selection with unequal variances. Sankhyā Ser. B 37, 28-78.

Dudewicz, E. J. and Koo, J. O. (1982). The Complete Categorized Guide to Statistical Selection and Ranking Procedures. Series in Mathematical and Management Sciences, Vol. 6, American Sciences Press, Columbus, Ohio.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50, 1096-1121.

Edwards, H. P. (1984a). RANKSEL: A ranking and selection package. The American Statistician, 38, 158-159.

Edwards, H. P. (1984b). RANKSEL: An interactive computer package of ranking and selection procedures. The Frontiers of Modern Statistical Inference Procedures (E. J. Dudewicz, ed.), Vol. 10, Series in Mathematical and Management Sciences, American Sciences Press, Columbus, Ohio, to appear.

Filliben, J. J. (1969). Simple and robust linear estimation of the location parameters of a symmetric distribution. Ph.D. thesis, Princeton University.

Finney, D. J. (1947). The principles of biological assay. J. Royal Statist. Society, Series B, 9, 46-91.

Ghosh, M. (1973). Nonparametric selection procedure for symmetric location parameter populations. Ann. Statist. 1, 773-779.

Gibbons, J. D., Olkin, I., and Sobel, M. (1977). Selecting and Ordering Populations. John Wiley, New York.

Goel, P. K. (1974). On the distribution of standardized mean of samples from logistic population. Technical Report No. 380, Dept. of Statistics, Purdue University, West Lafayette, Indiana.

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

Gupta, S. S. (1963). Probability integrals of the multivariate normal and multivariate t. Ann. Math. Statist. 34, 792-828.

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

Gupta, S. S. and Hsiao, P. C. (1981). On Γ-minimax, minimax, and Bayes procedures for selecting populations close to a control. Sankhya Series B, 43, 291-318.

Gupta, S. S. and Hsiao, P. (1983). Empirical Bayes rules for selecting good populations. J. Statist. Plan. Infer., 8, 87-101.

Gupta, S. S. and Hsu, J. C. (1984a). A computer package for ranking, selection, and multiple comparisons with best. Proceedings of the 1984 Winter Simulation Conference (S. Sheppard, U. Pooch and D. Pegden, eds.), North-Holland Publishing Company, 251-257.

Gupta, S. S. and Hsu, J. C. (1984b). User's Guide to RS-MCB: A Computer Program for Ranking, Selection and Multiple Comparison with the Best, Version X1. Correspondence to be addressed to Professor Jason C. Hsu, Department of Statistics, The Ohio State University, Columbus, Ohio 43210.

Gupta, S. S., and Huang, D.-Y. (1975a). On subset selection procedures for Poisson populations and some applications to the multinomial selection problems. Applied Statistics (R. P. Gupta, ed.), North-Holland, Amsterdam, 97-109.

Gupta, S. S., and Huang, D.-Y. (1975b). On some parametric and nonparametric sequential subset selection procedures. Statistical Inference and Related Topics, Vol. 2 (M. L. Puri, ed.), Academic Press, New York, 101-128.

Gupta, S. S., and Huang, D.-Y. (1976). Selection procedures for the means and variances of normal populations: unequal samples sizes case. Sankhyā Ser. B 38, 112-128.

Gupta, S. S. and Huang, D.-Y. (1980). An essentially complete class of multiple decision procedures. Journal of Statistical Planning and Inference, 4, 115-121.

Gupta, S. S. and Huang, D.-Y. (1981). Multiple Decision Theory: Recent Developments. Lecture Notes in Statistics, Vol. 6, Springer-Verlag, New York.

Gupta, S. S., Huang, D.-Y. and Nagel, K. (1979). Locally optimal subset selection procedures based on ranks. Optimizing Methods in Statistics (J. S. Rustagi, ed.), Academic Press, New York, 251-260.

Gupta, S. S. and Huang, W.-T. (1983). On isotonic selection rules for binomial populations better than a control. Proceedings of the First Saudi Symposium on Statistics and Its Applications.

Gupta, S. S. and Kim, W.-C. (1980). Γ-minimax and minimax decision rules for comparison of treatments with a control. Recent Developments in Statistical Inference and Data Analysis (K. Matusita, ed.), North Holland, 55-71.

Gupta, S. S. and Kim, W.-C. (1984). A two-stage elimination type procedure for selecting the largest of several normal means with a common unknown variance. Design of Experiments: Ranking and Selection (T. J. Santner and A. C. Tamhane, eds.), Marcel Dekker, New York, 77-93.

Gupta, S. S. and Leong, Y.-K. (1979). Some results on subset selection procedures for double exponential populations. Decision Information (C. P. Tsokos and R. M. Thrall, eds.), Academic Press, New York, 277-305.

Gupta, S. S. and Leu, L.-Y. (1983a). Nonparametric selection procedures for a two-way layout problem. Technical Report No. 83-41, Department of Statistics, Purdue University, West Lafayette, Indiana.

Gupta, S. S. and Leu, L.-Y. (1983b). Isotonic procedures for selecting populations better than a standard: Two-parameter exponential distributions. Technical Report No. 83-46, Department of Statistics, Purdue University, West Lafayette, Indiana.

Gupta, S. S. and Liang, T.-C. (1984). Locally optimal subset selection rules based on ranks under joint type II censoring. Technical Report No. 84-33, Department of Statistics, Purdue University, West Lafayette, Indiana.

Gupta, S. S., and McDonald, G. C. (1970). On some classes of selection procedures based on ranks. Nonparametric Techniques in Statistical Inference (M. L. Puri, ed.), Cambridge University Press, London, 491-514.

Gupta, S. S. and Miescke, K. J. (1982). On the problem of finding a best population with respect to a control in two stages. Statistical Decision Theory and Related Topics III (S. S. Gupta and J. O. Berger, eds.), Vol. 1, Academic Press, New York, 473-496.

Gupta, S. S. and Miescke, K. J. (1983). An essentially complete class of two-stage procedures with screening at the first stage. Statist. and Decisions 1(4/5), 427-440.

Gupta, S. S. and Miescke, K. J. (1984). On two-stage Bayes selection procedures. Sankhyā Series B, 46, 123-134.

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

Gupta, S. S., Nagel, K., and Panchapakesan, S. (1973). On the order statistics from equally correlated normal random variables. Biometrika 60, 403-413.

Gupta, S. S., and Panchapakesan, S. (1972). On a class of subset selection procedures. Ann. Math. Statist. 43, 814-822.

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

Gupta, S. S., Panchapakesan, S. and Sohn, J. (1985). On the distribution of the studentized maximum of equally correlated normal random variables. Comm. Statist.-Simulation and Computation, 14(1), 103-135.

Gupta, S. S. and Singh, A. K. (1979). On selection rules for treatments versus control problems. Proceedings of the 42nd Session of the ISI, Manila, Philippines, 229-232.

Gupta, S. S. and Singh, A. K. (1980). On rules based on sample medians for selection of the largest location parameter. Communications in Statistics - Theory and Methods, A9, 1277-1298.

Gupta, S. S., and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist. 29, 235-244.

Gupta, S. S., and Studden, W. J. (1970). On some selection and ranking procedures with applications to multivariate populations. Essays in Probability and Statistics (R. C. Bose et al., eds.), University of North Carolina Press, Chapel Hill, 327-338.

Gupta, S. S. and Yang, H.-M. (1984). Isotonic procedures for selecting populations better than a control under ordering prior. Statistics: Applications and New Directions, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (J. K. Ghosh and J. Roy, eds.), 279-312.

Hahn, G. J. and Shapiro, S. S. (1967). Statistical Models in Engineering. John Wiley & Sons, Inc., New York.

Hogg, R. V., Fisher, D. M. and Randles, R. H. (1972). On the selection of the underlying distribution and adaptive estimation. J. Amer. Statist. Assoc. 67, 597-600.

Hsu, J. C. (1978). Robust and nonparametric subset selection procedures. Tech. Report No. 156, Ohio State Univ., Columbus.

Hsu, J. C. (1981). A class of nonparametric subset selection procedures. Sankhyā Series B, 43, 235-244.

Huang, D.-Y. (1974). On some optimal subset selection procedures for Model I and Model II in treatments versus control problems. Mimeo. Ser. No. 373, Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Huang, D.-Y. and Panchapakesan, S. (1982). Some locally optimal subset selection rules based on ranks. Statistical Decision Theory and Related Topics-III, Vol. 2 (S. S. Gupta and J. O. Berger, eds.), Academic Press, New York, 1-14.

Huang, W.-T. (1984). Nonparametric Isotonic Selection Rules under A Prior Ordering. Design of Experiments: Ranking and Selection (T. J. Santner and A. C. Tamhane, eds.), Marcel Dekker, New York, 95-112.

Joiner, B. L. and Rosenblatt, J. R. (1971). Some properties of the range in samples from Tukey's symmetric lambda distribution. J. Amer. Statist. Assoc., 66, 394-399.

Lehmann, E. L. (1957). A theory of some multiple decision problems. II. Ann. Math. Statist., 28, 547-572.

Lorenzen, T. J. and McDonald, G. C. (1981). Selecting logistic populations using the sample medians. Commun. Statist., A10(2), 101-124.

McDonald, G. 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.

McDonald, G. C. (1972). Some multiple comparison selection procedures based on ranks. Sankhyā Ser. A, 34, 53-64.

McDonald, G. C. (1973). The distribution of some rank statistics with applications in block design selection problems. Sankhyā Ser. A, 35, 187-203.

McDonald, G. C. (1975). Characteristics of three selection rules based on ranks in the 3×2 exponential case. Sankhyā Ser. B, 36, 261-266.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. Ann. Math. Statist. 38, 1079-1091.

Matsui, T. (1984). Moments of a rank vector with applications to selection and ranking. Technical Report No. 84-47, Department of Statistics, Purdue University, West Lafayette, Indiana.

Miescke, K. J. (1980). On two-stage procedures for finding a population better than a control. Technical Report No. 80-28, Dept. of Statistics, Purdue University, West Lafayette, Indiana.

Miescke, K. J. (1983). Two-stage selection procedures based on tests. Design of Experiments: Ranking and Selection (T. J. Santner and A. C. Tamhane, eds.), Dekker, New York, 165-178.

Moberg, T. F., Ramberg, J. S. and Randles, R. H. (1978). An adaptive M-estimator and its application to a selection problem. Technometrics, 20, 255-263.

Mosteller, F. (1948). A k-sample slippage test for an extreme population. Ann. Math. Statist. 19, 58-65.

Mukhopadhyay, N. (1980). Stein's two-stage procedure and exact consistency. Abstract, Bull. Inst. Math. Statist. 9, 206.

Mykytka, E. F. and Ramberg, J. S. (1979). Fitting a distribution to data using an alternative to moments. Proceedings of the IEEE 1979 Winter Simulation Conference, 361-374.

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

Naik, U. D. (1975). Some selection rules for comparing p processes with a standard. Comm. Statist. 4, 519-535.

Öztürk, A. and Dale, R. F. (1985). Least squares estimation of the parameters of the generalized lambda distribution. Technometrics, 27, 81-84.

Paulson, E. (1949). A multiple decision procedure for certain problems in analysis of variance. Ann. Math. Statist. 20, 95-98.

Paulson, E. (1967). Sequential procedures for selecting the best one of several binomial populations. Ann. Math. Statist. 38, 117-123.

Ramberg, J. S. and Schmeiser, B. W. (1972). An approximate method for generating symmetric random variables, Comm. ACM, 15, 987-990.

Ramberg, J. S. and Schmeiser, B. W. (1974). An approximate method for generating asymmetric random variables. Comm. ACM, 17, 78-82.

Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J., and Mykytka, E. F. (1979). A probability distribution and its uses in fitting data. Technometrics, 21, 201-214.

Randles, R. H. (1970). Some robust selection procedures. Ann. Math. Statist. 41, 1640-1645.

Rizvi, M. H., and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest α-quantile. Ann. Math. Statist. 38, 1788-1803.

Rizvi, M. H., Sobel, M., and Woodworth, G. G. (1968). Nonparametric ranking procedures for comparison with a control. Ann. Math. Statist. 39, 2075-2093.

Rizvi, M. H. and Woodworth, G. G. (1970). On selection procedures based on ranks: counterexamples concerning the least favorable configurations. Ann. Math. Statist., 41, 1942-1951.

Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problems. Ann. Statist. 3, 334-349.

Santner, T. J. (1976). A two-stage procedure for selecting δ^*-optimal means in the normal model, Commun. Statist.-Theory and Method, A5, 283-292.

Sobel, M. (1967). Nonparametric procedures for selecting the t populations with the largest α-quantiles. Ann. Math. Statist. 38, 1804-1816.

Tamhane, A. C. (1975). A minimax two-stage permanent elimination type procedure for selecting the smallest normal variance. Technical Report No. 260. Dept. of Operations Research, Cornell Univ., Ithaca, New York.

Tamhane, A. C., and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist.-Theor. Method, A6, 1003-1033.

Tamhane, A. C., and Bechhofer, R. E. (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): an improved PCS lower bound and associated tables. Commun. Statist.-Theor. Method, A8, 337-358.

Tukey, J. W. (1960). The Practical Relationship Between the Common Transformations of Percentages of Counts and of Amounts. Technical Report 36, Statistical Techniques Research Group, Princeton University.

Turnbull, B. W. (1976). Multiple decision rules for compar · .. several populations with a fixed known standard. Commun. Statist.-Theor. Method., A5, 1225-1244.

AD-A159153

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report 85-20	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Multiple Decision Procedures for Tukey's Generalized Lambda Distributions	5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER: Technical Report #85-20	
7. AUTHOR(s): Joong K. Sohn	8. CONTRACT OR GRANT NUMBER(s): N00014-84-C-0167
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, West Lafayette, Indiana 47907	10. PROGRAM ELEMENT, PROJECT, TASK, AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, D.C.	12. REPORT DATE: August 1985
	13. NUMBER OF PAGES: 125
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

Tukey's Generalized lambda distribution; kurtosis; multiple selection procedures; isotonic selection procedures; sample medians; nonparametric procedures; score functions; elimination-type procedures; HPD credible regions

stable

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

Selection and ranking (more broadly multiple decision) problems arise in many practical siutations since it is now well-recognized that the classical tests of homogeneity usually do not provide the answers the experimenter wants. In this thesis ~~we study~~ Tukey's lambda distributions as the underlying model for selection and ranking problems. It is known that the family of Tukey's generalized lambda distributions is very broad and contains most well-known distributions as special cases. Chapter 1 deals with selection and ranking problems based on sample medians for the symmetric lambda distributions and gives →

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

applications of the lambda family of distributions. We investigate some properties of the lambda family of distributions. We also propose some selection procedures and study the properties of these procedures. An application of the lambda distribution for approximating some constants used in the selection and ranking procedures for other symmetric distributions is made.

In Chapter 2, the problems of isotonic selection procedures for the family of lambda distributions and for logistic distributions are considered. Some isotonic procedures are proposed and studied. The approximation of constants used in the proposed procedure is investigated. It is shown that the isotonic procedures are better than some classical procedures in terms of reducing the expected number of bad populations in the selected subsets.

Chapter 3 deals with the problem of choosing the optimal score function for different nonparametric procedures proposed by Nagel (1970) and Gupta and McDonald (1970). A Monte Carlo study is carried out. It indicates that the score function based on uniform distribution is optimal and robust against possible deviations from the underlying distributions.

In Chapter 4, a two-stage elimination-type procedure under the Bayesian setting is proposed and its properties are studied. In particular, we use a stopping rule to construct a $100(1-2\alpha)\%$ Highest Posterior Density Credible region with a common width $2d$ for the unknown means of selected populations at stage 1. A Monte Carlo study is carried out to examine the performance of the proposed procedure.

alpha

END

FILMED

11-85

DTIC