1.1 Introduction

In Medical Sciences, Engineering, Agriculture, Industry and related fields everyone is constantly faced with the problem of choosing the "best" among several available "treatments". In the framework of testing hypotheses, the classical procedure attempts to determine whether $k$ parameters associated with $k$ populations all have a common value. For example, if you have $k$ different drugs for treatment of a certain disease, the problem of concern is whether these drugs differ in their therapeutic value. If therapeutic value is measured by some parameter $\mu_j$ for the $j_{th}$ drug, $j=1,\ldots,k$, the classical procedure permits us to decide about the following null hypothesis, sometimes called the homogeneity hypothesis:

$$H_0: \mu_1 = \ldots = \mu_k.$$ 

If the test of homogeneity is the final goal of an investigator or researcher, alternative methods of statistical analysis are not required. However, there are many practical situations in which other types of goals are important. For example, suppose that in Analysis of Variance (ANOVA) the null hypothesis of homogeneity is rejected. An investigator or researcher is seldom satisfied with this conclusion. In particular, he may wish (a) to determine which populations differ from other populations, and in what directions, (b) to see which populations can be considered "best"
in some well defined sense of the term best, and (c) to select populations better than the control population. In case (a) techniques of multiple comparisons or simultaneous inference are frequently appropriate. The methods of multiple comparisons may also provide information that is relevant for the case (b). However, ranking and selection procedures are more appropriate for this purpose. In other words, in ranking and selection procedures one can go one step further for the problem of what to do next when in ANOVA null hypothesis of homogeneity is rejected.

Ranking and Selection procedures were initially developed in the early 1950's. Initially, Paulson (1949) investigated the problem of classifying the given populations into a "superior" and an "inferior" group. Later Paulson (1952) considered the problem of selecting the best of k categories, while comparing (k-1) experimental categories with a standard or control. Bechhofer (1954) for the first time devised the formal framework for ranking populations by considering the problem of ranking means of k normal populations having same known variance, known as indifference-zone approach. Under the same framework Bechhofer and Sobel (1954) proposed selection procedures for ranking k normal populations in terms of their variances. Gupta (1956) devised a different formulation for ranking populations with respect to their means, by selecting a (random sized) nonempty subset S of k populations, which contains the best population, 2
when the bestness of a population was defined in terms of its mean. This approach is referred to as the subset selection approach. A separate field of statistical research known as "Ranking and Selection Procedures" has been developed due to the efforts of Bechhofer(1954), Bechhofer and Sobel(1954) and Gupta(1956). Most of the researchers in last forty years in this field of statistical research have adopted either of these two main approaches of Bechhofer(1954) and Gupta(1956).

Gibbons, Olkin and Sobel(1977) and Gupta and Panchapakesan(1979) provide a comprehensive survey of the developments in the fields of Ranking and Selection. A fairly comprehensive categorized bibliography is provided by Dudewicz and Koo(1982). For a review paper in this field, one may refer to Laan and Verdooren(1989). For some other relevant references one may also refer to a most recent book by Mukhopadhyay and Solanky(1994).

1.2 Definitions and Notations

Consider $k(\geq 2)$ independent populations $\Pi_1, \Pi_2, \ldots, \Pi_k$ such that population $\Pi_i$ is indexed by location parameter $\mu_i (-\infty < \mu_i < \infty)$ and scale parameter $\theta_i (\theta_i > 0)$, $i=1,2,\ldots,k$. Let $x_{1i},\ldots,x_{ni}$ be a random sample of size $n_i$ from the population $\Pi_i$ and let $S_i(T_i)$ be an appropriate estimator of $\mu_i(\theta_i)$, $i=1,2,\ldots,k$, if $\mu_i(\theta_i)$ is unknown. Here $S_i=S_i(x_{i1})$, $T_i=T_i(x_{i1})$, and $x_{i1}=(x_{i1},\ldots,x_{in_i})^t$, $i=1,2,\ldots,k$. Assume that $S_i(T_i)$ has a cumulative distribution function (cdf) $P(x-\mu_i)(F(x/\theta_i))$, $i=1,2,\ldots,k$. Denote the ordered
\( \mu(\theta) \) values by \( \mu_1(\theta), \ldots, \mu_k(\theta) \). Let \( \Pi_i(\Pi_j) \) denote the population associated with \( \mu_1(\theta), \ldots, \mu_k(\theta) \), the \( i \)(th) smallest of \( \mu_1, \ldots, \mu_k(\theta) \). It is assumed that \( \mu_1, \ldots, \mu_k(\theta) \) are unknown, there is no prior knowledge about which of the \( \Pi_1, \ldots, \Pi_k \) is \( \Pi_i(\Pi_j) \), \( i=1,2,\ldots,k \). When the parameters of interest are location (scale) parameters, the populations \( \Pi_1(\Pi_1) \) and \( \Pi_k(\Pi_k) \) associated with \( \mu_1(\theta) \) and \( \mu_k(\theta) \) are defined as "worst" ("best") and "best" ("worst") populations respectively. In case of ties for \( \mu_1(\theta) \) or \( \mu_k(\theta) \), we assume that one of the tied population is arbitrary labelled as the worst (best) or best (worst). Let \( S_1 \leq \ldots \leq S_k \) and \( T_1 \leq \ldots \leq T_k \), denote the ranked values of \( S_i \)'s and \( T_i \)'s, respectively.

Denote

\( R = \{ x : -\infty < x < \infty \} \), the entire real line.

\( R^+ = \{ x : x \geq 0 \} \), nonnegative part of the real line.

\( R^m = \{ x : x = (x_1, \ldots, x_m), -\infty < x_i < \infty, i = 1,2,\ldots,m \} \), \( m \)-dimensional Euclidean space.

\( \Omega = \{ \mu : \mu = (\mu_1, \ldots, \mu_k)^T, \text{ for all } \mu_i \in (-\infty, \infty), i=1,\ldots,k \} \) denote the parametric space in case of location parameters.

\( \mathcal{R}_+^m = \{ x, x = (x_1, \ldots, x_m), x_i \geq 0, i = 1,2,\ldots,m \} \).

\( P_j(\Pi_j \in \mathcal{R} \) is included in the subset \( R \), \( j=1,2,\ldots,k \), for any selection rule \( R \) and any \( \mu \in \Omega \).

**Definition 1.2.1** Let \( \theta \) be a location (scale) parameter of the distribution of random variable (r.v.) \( X \) with probability density...
function (pdf) \( g_x(\cdot; \xi) \), then \( g_x(x; \xi) = g(x - \xi), \) \(-\infty < \xi < \infty \) \((1/\xi) g(x/\xi)\), \(\xi > 0\), where \(g(\cdot)\) is a pdf independent of \(\xi\).

**Definition 1.2.2** For a set \(A\), let \(I_A(z)\) denote the indicator function of set \(A\), that is,

\[
I_A(z) = \begin{cases} 
1, & \text{if } z \in A \\
0, & \text{if } z \notin A.
\end{cases}
\]

**Definition 1.2.3** For a nonempty set \(A\), let \(|A|\) denote the cardinality of the set \(A\) (that is number of elements in \(A\)).

**Definition 1.2.4** A "correct selection" is defined as an event which is in accordance with the goal of the considered problem.

**Definition 1.2.5** A cumulative distribution function (cdf) \(H(\cdot; \xi)\) depending on real parameter \(\xi\) is said to have stochastically increasing property (SIP) if for \(\xi_1 < \xi_2\)

\[
H(x; \xi_1) \leq H(x; \xi_2)
\]

for all \(x\).

Some examples of stochastically increasing distribution function are:

- Any location parameter family (that is \(F(x) = F(x - \mu), \mu \in \mathbb{R}\)).
- Any scale parameter family (that is \(F_\theta(x) = F(x/\theta), x > 0, \theta > 0, \theta \in \mathbb{R}^+\)).
- Any monotone likelihood ratio family (that is \(f_1(x_1)f_2(x_2) \geq f_2(x_1)f_1(x_2)\) for any \(\mu_1 \geq \mu_2\) and \(x_1 \geq x_2\)).

**Definition 1.2.6** A selection procedure \(R\) is said to be unbiased if and only if

\[
\mu_j \geq \mu_i, \quad i = 1, 2, \ldots, k \implies P_j(\mu) \geq P_i(\mu) \quad \text{for } i = 1, 2, \ldots, k \text{ and for all } \mu \in \Omega.
\]
Definition 1.2.7 A selection procedure R is said to be monotone if and only if
\[ \mu_j \geq \mu_i, \quad 1=1,2,\ldots,k \text{ implies that} \]
\[ \frac{P_j}{\mu} \geq \frac{P_i}{\mu} \text{ for all pairs } (j,i) \text{ and for all } \mu \in \Omega. \]

Definition 1.2.8 A selection procedure R is said to be strongly monotone in \( \mu_j \) if and only if
\[ P_j(\mu) \text{ is increasing in } \mu_j \text{ when all other components of } \mu \text{ are fixed, and } P_i(\mu) \text{ is decreasing in } \mu_i (i \neq j) \text{ when all other components of } \mu \text{ are fixed.} \]

Definition 1.2.9 A pdf \( g(x) \) is said to have monotone likelihood ratio (MLR) in \( x \) if for every \( c \in (0,1) \), \( g(xc)/g(x) \) is non-decreasing in \( x \).

Definition 1.2.10 Let \( X \) and \( Y \) be two r.v. with cdf \( F \) and \( G \) respectively, let \( F^{-1} \) and \( G^{-1} \) be the left continuous inverses of \( F \) and \( G \) respectively and assume that
\[ F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \text{ whenever } 0<\alpha<\beta<1, \]
then \( X \) is said to be smaller than \( Y \) in the dispersive order (denoted as \( X \preceq_{disp} Y \)).

Some other definitions and notations, which are used less frequently will be provided at the time of their need.

1.3 Review of Literature

The basic aim of the indifference-zone approach of Bechhofer (1954) is to select one of the populations as the best (or worst) population with probability at least...
\( P^* \left( \frac{1}{k} < P < 1 \right), \) a preassigned constant, whenever the parameter vector \( \mu \) (or \( \theta \)) belongs to a certain subset of \( \mathbb{R}^k \) \( \mathbb{R}_+^k \), called the preference-zone. The region complimentary to the preference-zone is called the indifference-zone. In the subset selection approach of Gupta (1956), the basic goal is to select a nonempty subset \( S \) of \( k \) populations such that the selected subset contains the "best" (or "worst") population with a probability at least \( P^* \left( \frac{1}{k} < P < 1 \right), \) a preassigned constant, for all parametric configurations.

Bechhofer (1954) considered the problem of selecting the best of \( k (\geq 2) \) normal populations with unknown means \( \mu_1, \ldots, \mu_k \) being the parameters of interest. Let \( \mu_1 < \mu_2 < \cdots < \mu_k \) denote the ordered values of \( \mu_1, \ldots, \mu_k \). Let \( S_i, i=1,2,\ldots,k \) denote the sample means based on a random sample of size \( n \) from population \( \Pi_i, i=1,2,\ldots,k \). Assuming population variances \( \theta_1^2, \ldots, \theta_k^2 \) are known, he proposed a selection procedure, which selects the population associated with \( S_{<k} = \max(S_1, \ldots, S_k) \) as the best population and showed that his procedure achieves the goal with probability at least \( P^* \left( \frac{1}{k} < P < 1 \right), \) whenever \( \mu = (\mu_1, \ldots, \mu_k) \) belongs to a part of the parametric space \( \Omega = \{ \mu, \mu = (\mu_1, \ldots, \mu_k)^t, -\infty < \mu_i < \infty, i=1,2,\ldots,k \} \) namely \( \Omega^*_\delta \) in which \( \mu_{[k]} - \mu_{[k-1]} = \delta^* \). The set \( \Omega^*_\delta \) is called the preference zone and its complement \( \Omega^C_{\delta^*} = \Omega - \Omega^*_\delta \) is called the indifference-zone, \( \delta^* \) being a preassigned positive constant. For \( \theta_1^2 = \cdots = \theta_k^2 = \theta^2 \) (say), the common sample size needed to achieve the goal is the smallest nonnegative integer \( n \) satisfying the
following equation
\[ \int_{-\infty}^{\infty} \phi_{k-1}(s + (n_{i-1/2}^1/\delta^*)) \phi(s) ds \geq P^*. \quad \ldots (1.3.1) \]

Values of \( n \) satisfying equation (1.3.1) are tabulated by Bechhofer (1954) for various values of \( P^*, k \) and \( \delta^* \).

Earlier Bahadur (1950) and Bahadur and Goodman (1952) had considered the problem of constructing appropriate decision rules for selecting the normal population associated with the largest mean, when the populations have common known variance. They proved that if the sample means from each population are based on a common sample size, then the "impartial" decision rule always selects the population associated with the largest sample mean. Optimal properties of this decision rule, which selects the population associated with the largest observation have been investigated by Hall (1959), Lehmann (1966), Eaton (1967) and Alam (1973) among others.

Under the subset selection approach of Gupta (1956) the basic goal is to select a nonempty subset of the \( k \) populations so that the best population is contained in the selected subset \( S \) with a minimum probability \( P^*((1/k) < P^* < 1) \), where \( P^* \) is a preassigned constant. Under this approach the size of the selected subset is not determined in advance but it is determined by the data themselves.

Let \( S_i(T_i) \) be the appropriate statistics based on a random sample from population \( \Pi_i \), \( i=1,2,\ldots,k \), with cumulative distribution function \( F(x-\mu_i)(F(x/\theta_i)) \) in the location (scale)
case, where $\mu_i$’s ($\theta_i$’s) are the location (scale) parameters to be ranked. Selection of any subset consistent with the goal (that is including the best population) is a correct selection (CS). The probability of correct selection using a selection procedure $R$ shall be denoted as $P_{\mu}(CS/R)$. The selection procedures proposed by Gupta (1956, 1965) are $R_{01}$ (location case) and $R_{02}$ (scale case), given below

$$R_{01} : \text{Select } \Pi_i \text{ if and only if } S_i = \max_{1 \leq r \leq k} S_r - d \quad \text{(1.3.2)}$$

and

$$R_{02} : \text{Select } \Pi_i \text{ if and only if } T_i = c \max_{1 \leq r \leq k} T_r \quad \text{(1.3.3)}$$

where the selection constants $d=d(k,\star)>0$ and $c=c(k,\star)$, $c\in(0,1)$ are to be determined so that probability of correct selection (CS) is at least $P^*$, that is

$$P_{\mu}(CS|R_{01}) \geq P^*, \quad \mu \in \mathbb{R}^k, \text{ (location case)} \quad \text{(1.3.4)}$$

$$P_{\theta}(CS|R_{02}) \geq P^*, \quad \theta \in \mathbb{R}^k, \text{ (scale case).} \quad \text{(1.3.5)}$$

Gupta (1965) has also discussed some desirable properties of the rules $R_{01}$ and $R_{02}$.

The results of Gupta (1965) have been used by Berger and Gupta (1980) for examining the minimaxity of several rules for normal means problem when the variances are known but not necessarily equal for unequal sample sizes. Many procedures have been developed using subset selection formulation among others by
Gupta and Sobel (1960, 1962), Gupta (1963a), Rizvi and Sobel (1967),
Studden (1967), Barlow and Gupta (1969), Gupta and Deverman (1969),
Gupta and McDonald (1970, 1972), Gupta and Nagel (1971),
Govindarajulu and Gore (1971), Gupta and Panchapakesan (1972),
Ofosu (1974), Patel (1976), Hsu (1980, 1981a), Lam (1986), Gill and

In addition to these two main formulations, some other
formulations have also been considered by several authors.
Sobel (1956), Bechhofer (1958), Bechhofer and Blumenthal (1962),
Paulson (1967), Robbins, Sobel and Starr (1968) have proposed
sequential procedures for selecting the best of \( k \) populations.
Berger (1979) proposed minimax subset selection procedures where
the loss is measured by subset size and later Berger (1981)
proposed the same selection procedures for the multinominal
distribution. Berger and Gupta (1980) obtained minimax rules in
the class of non-randomized, just, and translation invariant
rules when the risk is measured by the maximum probability of
including a nonbest population. Bjornstad (1981) considered a
decision theoretic approach for selecting a non-empty subset of
populations out of \( k \) populations. Goel and Rubin (1977) have
developed some Bayes selection procedures, that is, using a
priori probability distribution on the parameter space and a loss
function. Randles and Hollander (1971) have proposed \( \alpha \)-minimax
selection procedures for selecting treatment populations that
have larger translation parameters than that of a control
population assuming that the density function of underlying population possess the monotone likelihood ratio (MLR) property. Haller (1968) has suggested some optimal k-sample rank order procedures for selection and tests against slippage and ordered alternatives using the test statistics (based on the generalized k-sample U-statistics) proposed by Deshpande (1965) for the k-sample location problem. Santner (1975) formulated restricted subset selection procedures in which the size of the selected subset containing the "best" population from a set of k populations is bounded by $m(1 \leq m \leq k)$, where $m$ is a predetermined integer. Isotonic selection procedures have been investigated by Gupta and Huang (1982) in the case of binomial population; Gupta and Lee (1983) in case of two-parameters exponential population and Gupta and Yang (1984) in case of normal means (common variance $\sigma^2$ may be known or unknown) for selecting populations better than a standard one. Lam (1986) developed a selection procedure for selecting good normal populations. Gill and Sharma (1993) extended the results of Lam (1986) for the exponential set up. Laan (1991a, 1991b, 1992a) considered subset selection procedures for an almost best population for location probability model. Here the goal was to select a subset containing at least one $\Delta$-best population with confidence $P$. A treatment was called $\Delta$-best if it is at a distance less than or equal to $\Delta(\Delta > 0)$ from the best population, whereas the best population was to be associated with the largest location parameter. For this goal he
proposed General Subset Selection rule and the selection constant 'd' was determined so that the selected subset will have at least one $\Lambda$-best population with probability at least $P^*$. For comparing the generalised selection procedures for a $\Lambda$-best treatment with the subset selection procedures for the best treatment, he introduced the concept of relative efficiency. For normal probability model, the values of relative efficiency were computed by Laan(1991b) for various choices of $k$, $P^*$, and $\Lambda$. Laan(1992b) also obtained explicit expressions for expectation and variance of subset size using subset selection procedure. As already mentioned some books have appeared exclusively dealing with ranking and selection procedures. Of these the monograph by Bechhofer, Kiefer and Sobel(1968) deals with sequential procedures with special emphasis on Koopman-Darmois family.

Gupta and Sobel(1960) considered the problem of selecting a subset containing the best of several binomial populations, with $\theta$ as the probability of one of two possible outcomes. Gupta(1963a) has considered the selection problem for gamma populations when the parameters of interest are scale parameters. For selection from multivariate normal populations several criteria were considered, such as Mahalanobis distance function (Alam and Rizvi(1966), Gupta(1966), Gupta and Studden(1970)), and generalized variance (Gnanedesikan and Gupta(1970)). Selection procedures for multinomial distribution have been developed by Gupta and Nagel(1967). Selection procedures for the normal
populations have been proposed by Bechhofer (1954), Bechhofer and Sobel (1954) and Gupta and Huang (1976), among others. Gupta and Sobel (1962) have proposed subset selection procedures based on sample variances to select a subset of \( k \)-normal populations which includes the normal population associated with the smallest variance. Gupta and Huang (1976) have proposed subset selection procedures for normal means and variances under unequal sample sizes. In the same paper, they have provided the lower bound for the PCS for selecting a subset which includes the population associated with the smallest variance. Chen, Dudewicz and Lee (1976) have discussed the modified forms of selection procedures for normal means under unequal sample sizes case. Lam (1986) proposed a selection procedure for selecting a subset of \( k \)-normal populations (having same variance) which contains the population associated with the largest mean. There the author has obtained the infimum of the PCS directly using the distribution of the range of \( k \) standard normal random variables. By defining the bestness in terms of larger population mean, Sobel (1987) proposed a selection procedure \( R \) for selecting the \( t \) best out of \( k \) normal populations when the \( t \)th and \((t+1)\)th largest sample means are close together. In such situations he suggested to select two disjoint subsets \( S_G \) and \( S_I \) instead of one subset \( S \), such that \( S = S_G \cup S_I \), when \( S \) contains the \( t \) best populations. The common sample size \( n \) and the constants of the procedure \( R \) have been defined so that \( P \)-condition are satisfied; one for the
preference zone and other for the indifference-zone. Berger (1981) proposed subset selection procedures for selecting all the treatments populations with means larger than a control population when treatments and control population were assumed to have multivariate normal distributions.

Many researchers, mentioned above, have used observed values of a quantitative characteristic for proposing selection procedures and as such these procedures are essentially parametric procedures. Selection procedures based on ranks for selecting the best population, that is, the one having the largest location parameter were initiated by Lehmann (1963b). Puri and Puri (1969) considered the problem of selecting out of k populations (i) the t best ones without regard to order, (ii) the t best with regard to order and (iii) finding a subset which includes all populations better than a standard one. In these procedures the bestness of a population is characterised by its location parameter, and so on. Puri and Puri (1968) proposed selection procedures based on ranks for the above problems (i), (ii) and (iii) with the difference that the bestness of a population is characterised by its scale parameter. On the basis of a random sample $X_{i1}, \ldots, X_{in}$, $i=1,2,\ldots,k$ of size n from ith population, they proposed selection procedures for the problems (i), (ii) and (iii) in the scale parameter case using the statistics
\[ T_i = \frac{1}{n} \sum_{j=1}^{n} E_{N,j} Z_{N,j}^{(i)}, \quad i=1,2,\ldots,k, \]

Where, 
\[ Z_{N,j}^{(i)} = \begin{cases} 
1, & \text{if ith smallest of } N=nk \text{ absolute values } |x_{ij}|, \\
0, & \text{otherwise}
\end{cases} \]

and \( E_{N,j} \) is the expected value of the square of the \( j \)th order statistic of a sample of size \( N \) from a given continuous distribution function \( F_0 \). Gupta and McDonald (1970) proposed selection procedure based on joint ranking while assuming that populations differ either in their location or scale parameters. They have also given bounds for the infimum of PCS. Some other related papers include McDonald (1972, 74), Blumenthal and Patterson (1969) and Bartlett and Govindarajulu (1967). Randles (1970) used indifference-zone approach in proposing selection procedures for selecting a population with largest location parameter from the given \( k \) populations (differing only in location parameters) on the basis of two samples Hodges & Lehmann (1963) estimators of location difference between two populations. These procedures possess more desirable finite sample properties and equivalent asymptotic properties as compared to the procedures considered by Lehmann (1963b), Bartlett & Govindarajulu (1967) and Puri & Puri (1969). Ghosh (1973) on the basis of indifference-zone formulation proposed selection procedures for location parameters of \( k \) independent symmetric populations using the one sample Hodges & Lehmann (1963) estimators. Most of the above mentioned selection procedures
based on ranks make use of one sample rank statistics. All these selection procedures based on ranks assumed that the data arise from a one way ANOVA type setting. Lee and Dudewicz (1986) proposed selection procedures based on ranks in situations when the data arose from the two way ANOVA model, when the blocks effects enters; namely $P[X_{ij} \leq x] = F(x-\eta_j-\mu_1)$, where $\eta_j$ is a nuisance location parameter of the jth block ($i=1,2,...,k; j=1,2,...,n$). They proposed to prefer ranks within each block as compared to joint ranks to devise selection procedures using indifference-zone approach under the slippage parametric configuration (SPC). However, it was pointed out by Rizvi & Woodworth (1970), that the procedures based on ranking the samples from all the populations jointly cannot control the PCS over the entire parametric space. Hsu (1980, 81a) and Gill & Mehta (1989) proposed subset selection procedures based on the set of all possible two-sample linear rank statistics and have shown that these procedures control the PCS over the entire parametric space. Subset selection procedures, satisfying $P^*$-condition, based on the set of all possible two-sample U-Statistics which can be expressed in terms of ranks under pairwise ranking have been proposed by Gill & Mehta (1991), and Kumar et al (1992, 1994b). Based on random samples $X_{ij}$ ($i=1,2,...,k; j=1,2,...,n_1$), while comparing the k populations in terms of either location or scale parameters, Koziol & Reid (1977) have shown the asymptotic equivalence of two ranking methods of sample observations namely
(i) the ranking of k sample pairwise and (ii) the joint ranking of k samples. The ranking method (i) was proposed by Steel (1960) & ranking method (ii) was proposed by Dunn (1964). Motivated with this asymptotic equivalence of two ranking methods and the fact that selection procedures based on the set of all possible two-samples ranks statistics, derived under the pairwise ranking, control the PCS over the entire parametric space while the procedures based on joint ranking fail to do so, Hsu (1980, 1981a), Gill and Mehta (1989, 1991, 1994), and Kumar et al (1992, 1994b) used pairwise ranking to propose subset selection procedures. Using the asymptotic equivalence of two ranking methods and the fact that PCS is controlled over the entire parametric space under pairwise ranking, we propose some more subset selection procedures using two sample U-statistics, expressible in terms of ranks under pairwise ranking, in this dissertation. The performance of the members in each class, in terms of estimated expected subset size has been assessed through simulation study. The simulation results have also helped us to specify the common sample size required for the implementation of the asymptotic results.

of increasing failure rate (IFR) distribution functions. Kingston and Patel (1980) proposed a selection procedure for selecting the best one of the several Weibull populations. Park and Saxena (1987) considered a non-parametric approach for selecting the least NBU (new better than used) or NWU (new worse than used) population.

Motivated by the results of Park and Saxena (1987), we propose a useful class of selection procedures, based on two-stage technique, for selecting the least dispersive distribution from the k available distributions. This problem finds application in reliability and engineering. In engineering, for example, the goal of the experimentor is to select a firm whose components have least dispersive distribution from the available set of competing firms manufacturing the components of the desired specifications meant for the same purpose.

Simultaneous confidence intervals that can be derived from selection statements without decreasing the nominal confidence level were initiated by Hsu (1981b). Bofinger (1983) considered the problem of selecting the best of k populations and subsequent prediction for this population versus 'the rest'. Under exponential set up, simultaneous confidence intervals for all ratios to the best were developed by Kumar et al (1994c). For some other references in multiple comparisons, one may refer to Hochberg & Tamhane (1987).

Gutmann & Maymin (1987) considered the problem of whether the
selected population is the best? In this dissertation we make an attempt to extend the results of Gutmann & Maymin (1987), derived for location parameters under the normality probability model, to the general scale probability model, in order to attach the best possible p-value to the inference: the selected population is the best.

1.4 Outline of the thesis

In the present dissertation, we continue the study carried out by various researchers mentioned above by developing some selection procedures based on two basic formulations. The proposed selection procedures have been demonstrated by taking real life examples and their performance has been assessed through simulation study.

In chapter-II of this dissertation we have considered $k$ populations such that population $\Pi_i$, $i=1,2,...,k$ is indexed by the location parameter $\mu_i$, $i=1,2,...,k$. The problem is to select a subset of the $k$ populations containing the population associated with the largest location parameter. Any such selection will be termed as correct selection (CS).

Consider $k$ independent populations $\Pi_1, \Pi_2, \ldots, \Pi_k$ and let $F_i(x) = F(x - \mu_i)$ be the absolutely continuous cumulative distribution function (cdf) of the $i^{th}$ population characterised by the location parameter $\mu_i$, $i=1,...,k$. Let $\Omega = \{\mu : \mu = (\mu_1, \mu_2, \ldots, \mu_k)^t, -\infty < \mu_i < \infty, i=1,...,k\}$ is used to denote the parametric space. For any two populations $\Pi_i$ and $\Pi_j$, $\Pi_j$ is...
considered to be better than \( \pi_i \) if \( \mu_j \preceq \mu_i \). For any \( \mu = (\mu_1, \ldots, \mu_k)^t \), we shall denote by \( \mu^{[k]} \) the unique component of \( \mu \) corresponding to the best population. The goal is to select a subset of \( k \) populations containing the best population, the one with the largest location parameter \( \mu^{[k]} \). Then the problem is to find a rule \( R \) such that for a preassigned probability \( P^*(1/k) < P^* < 1 \), this satisfies the probability requirement:

\[
P^* \{ \text{CS} | R \} \geq P^* \text{ for all } \mu \in \Omega.
\]

Let \( X_{i1}, X_{i2}, \ldots, X_{in_i} \) be a random sample of size \( n \) from the \( i \)th population, \( i = 1, \ldots, k \). Let \( X_i = (X_{i1}, X_{i2}, \ldots, X_{in_i})^t \) be a vector of observations from the \( i \)th population and let \( X = (X_{11}, \ldots, X_{1n_1}, \ldots, X_{kn_k})^t \) be the vector of all the observations. Denote \( n_{[1]} \leq \cdots \leq n_{[k]} \) as the ordered values of \( n_{1}, \ldots, n_{k} \).

The selection procedures are based on two sample U-statistics given below. Fix \( i \) in the discussion below and define

\[
H(x) = F_i(x) = F(x - \mu_i),
\]

where \( H(. . .) \) is any continuous distribution function. Now

\[
P_{ij}(x) = F(x - \mu_j) = H(x - \Delta_{ij})
\]

where \( \Delta_{ij} = \mu_j - \mu_i \), \( j \neq i, j = 1, \ldots, k \).

Let \( c \) be a fixed odd integer such that \( 1 \leq c \leq \min(n_{[1]}, \ldots, n_{[k]}) \).
Define the kernel $\phi(.)$ as

$$
\phi(x_1, \ldots, x_{ic}; x_{j1}, \ldots, x_{jc}) =
\begin{cases}
1 \text{ if } \text{Median}(x_{i1}, \ldots, x_{ic}) < \text{Median}(x_{j1}, \ldots, x_{jc}) \\
0 \text{ otherwise,}
\end{cases}
$$

and the U-statistic based on this kernel is

$$
U_{ij}^C = \frac{1}{\binom{n_i}{c} \binom{n_j}{c}} \sum_{(a_1, \ldots, a_c) \in \binom{1, \ldots, n_i}{c}} \sum_{(\beta_1, \ldots, \beta_c) \in \binom{1, \ldots, n_j}{c}} \phi(x_{ia_1}, \ldots, x_{ic}; x_{j\beta_1}, \ldots, x_{j\beta_c}),
$$

where the summation is extended over all combinations of $c$ integers $(a_1, \ldots, a_c)$ chosen without replacement from $(1, \ldots, n_i)$ and all combinations of $c$ integers $(\beta_1, \ldots, \beta_c)$ chosen without replacement from $(1, \ldots, n_j)$.

The expected value of $U_{ij}^C$ under $\mu_i = \mu_j$ is

$$
E[U_{ij}^C] = \frac{c!}{((c-1)/2)!!} \sum_{t=(c+1)/2}^C \binom{c}{t} \beta\left[\frac{c+1}{2} + t, \frac{3c+1}{2} - t\right]
$$

(where $\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$)

$$
= \frac{1}{2} \text{ for all } c.
$$

Define

$$
V_{ij}^C = U_{ij}^C - \frac{1}{2}.
$$

The proposed class of selection procedures based on the statistics $V_{ij}^C$ is:

$$
R_C : \text{Select population } \Pi_j \text{ in the subset if and only if }
$$

$$
V_{ij}^C > b \left[c^2 K C (n_j^{-1} + n_i^{-1})\right]^{1/2}, \forall i, i\neq j.
$$

Here $K_C$ is a constant which depends upon the choice of $c$ (see expression 2.5.4) of chapter II and the constant $b$ is chosen such that
\[ P_0[ V_{ij}^c \geq b \left( c^2 K_c (n^{-1}_j + n^{-1}_i) \right)^{1/2} \text{ for all } i, i \neq j ] \geq p^*, \]

where \( P_0 \) indicates that the probability is computed under \( \mu_1 = \ldots = \mu_k \).

Let \( A \) be the action space of the subset selection problem which is the set of all nonempty subsets of \( \{1, \ldots, k\} \), where taking action \( a \in A \) means the selection of those populations whose indices are in \( a \). For any \( a \in A \), we have shown that for all \( \mu \in \Omega \),

\[ P_\mu[CS|R] = E \left[ E_{\mu} \left[ CS(\mu, a) Z_{R_c}(X,a) \right] \right], \]

where

\[ CS(\mu, a) = \begin{cases} 1 & \text{if } \mu_{[k]} \in \{ \mu_i, i \in a \} \\ 0 & \text{otherwise} \end{cases} \]

and \( Z_{R_c}(X,a) \) is the probability assigned to \( a \) by \( R_c \) having observed \( X \).

Suppose \( S \) denote the size of the selected subset. Then the expected subset size is seen to be

\[ E_{\mu}[S|R_c] = E \left[ \sum_{\mu} |a| Z_{R_c}(X,a) \right], \]

where \( |a| \) denotes the cardinality of the set \( a \).

It has been shown that family of distributions of \( V_{ij}^c \) is stochastically increasing. Hence by using a Lemma 4.2 of Mahamunulu(1967), it has been shown that procedures \( R_c \) satisfy \( P^* \)-condition for every choices of \( c \). Again by using a result of Mahamunulu(1967), the strong monotonicity property of \( R_c \) has been established. The rank representation of \( V_{ij}^c \) is also given.

The proposed selection procedures are compared with existing
procedures in the sense of Pitman asymptotic relative efficiency (ARE). A modification of the proposed selection procedures to select populations better than the unknown control population is also given. Simulation study has been carried out to assess the performance of the proposed selection procedures.

In chapter III a new class of selection procedures based on sample quantiles is proposed. The proposed selection procedures are based on two sample U-statistics. This class contains the procedures proposed in chapter II as a special case. The rank representation of the proposed selection procedures is given and some properties of these selection procedures are also discussed.

Let there be $k$ independent populations such that the distributions associated with them differ only in terms of location parameters and our goal is to select a subset of the given $k$ populations which contains the population associated with the distribution having largest location parameter. The population associated with the largest location parameter is labelled as the best population.

Let $X_{i1}, X_{i2}, \ldots, X_{in_i}$ be a random sample of size $n_i$ from the $i$th population, $i=1, \ldots, k$. Let $X_i = (X_{i1}, X_{i2}, \ldots, X_{in_i})^\top$ be a vector of observations from the $i$th population and let $X = (X_{11}, \ldots, X_{1n_1}, \ldots, X_{kn_k})^\top$ be the vector of all the observations. Denote $n_{[1]} \leq \ldots \leq n_{[k]}$ as the ordered values of $n_1, \ldots, n_k$. 

23
Consider a fixed integer $c$ such that $1 \leq c \leq \min(n_1, \ldots, n_k)$. Define the kernel $\phi(.)$ as

$$
\phi(x_{i1}, \ldots, x_{ic}; x_{j1}, \ldots, x_{jc}) = \begin{cases} 
1 & \text{if } m^{th} \text{ order statistic of } (x_{i1}, \ldots, x_{ic}) \\
0 & \text{otherwise}
\end{cases}
$$

where $1 \leq m \leq c$.

The U-statistic based on this kernel is

$$
U_{ij}^{m,c} = \frac{1}{\binom{n_i}{c} \binom{n_j}{c}} \sum \phi(x_{i\alpha_1}, \ldots, x_{i\alpha_c}; x_{j\beta_1}, \ldots, x_{j\beta_c}),
$$

where the summation is extended over all combinations of $c$ integers $(\alpha_1, \ldots, \alpha_c)$ chosen without replacement from $(1, \ldots, n_i)$ and all combinations of $c$ integers $(\beta_1, \ldots, \beta_c)$ chosen without replacement from $(1, \ldots, n_j)$.

The expected value of $U_{ij}^{m,c}$ under $\mu_i = \mu_j$ is

$$
E[U_{ij}^{m,c}] = \binom{c}{m} \sum_{t=m}^{c} \binom{c}{t} \beta [m + t, 2c-m+1-t] = 1/2 \text{ for all } c \text{ and } m.
$$

Define

$$
W_{ij}^{m,c} = U_{ij}^{m,c} - 1/2.
$$

The proposed selection procedures based on the statistics $W_{ij}^{m,c}$ are:

- $R_{m,c}$: Select population $\Pi_j$ in the subset if and only if

$$
W_{ij}^{m,c} > b [c K_{m,c}^2 (n_j^{-1} + n_i^{-1})]^{1/2} \quad \text{for all } i, i \neq j.
$$

Here $K_{m,c}$ is a constant which depends upon the choice of $m$ and $c$.
(see expression (3.5.4)) of chapter III and the constant \( b \) is chosen such that

\[
P_0\left[ \frac{W_{ij}^{m,c}}{b} \leq c \frac{K^2_{m,c}(n^{-1}_j + n^{-1}_i)}{n} \right]^{1/2} \quad \text{for all } i, i \neq j \neq \ast,
\]

where \( P_0 \) indicates that the probability is computed under \( \mu_1 = \ldots = \mu_k \).

The proposed selection procedures are compared with existing procedures in the sense of Pitman asymptotic relative efficiency (ARE). A modification of the proposed selection procedures to select a subset containing the populations better than an unknown control population is given. Simulation study has been carried out to assess the asymptotic performance of the proposed selection procedures in terms of expected subset sizes for particular parametric configurations of underlying populations.

In chapter IV we propose procedures which are generalisation of the selection procedures proposed earlier by Kumar et al (1992) for location model based on two sample U-statistics using maxima and minima of sub-samples. The proposed general class of selection procedures for the location parameters are based on sample quantiles, which include the sub-sample extrema as special cases. The class of procedures is based on two sample U-statistics. The rank representation of the two sample U-statistics is given and some properties of these selection procedures are also discussed.

Suppose there are \( k \) independent populations such that the
distributions associated with them differ only in terms of location parameters and our aim is to select a subset of the given k populations which contains the population associated with the distribution having largest location parameter. The population with the largest location parameter is termed as the best population.

Let $X_{i1}, X_{i2}, \ldots, X_{in_i}$ constitutes a random sample of size $n_i$ from the $i^{th}$ population, $i=1, \ldots, k$. Let $X_i = (X_{i1}, X_{i2}, \ldots, X_{in_i})^t$ be a vector of observations from the $i^{th}$ population and let $X = (X_{11}, \ldots, X_{in_1}, \ldots, X_{k1}, \ldots, X_{kn_k})^t$ be the vector of all the observations. Denote $n_{[1]} \ldots n_{[k]}$ as the ordered values of $n_1, \ldots, n_k$.

Consider a fix integer $c$ such that $1 \leq c \leq \min(n_1, \ldots, n_k)$. Define the following kernels $\phi^{(1)}(.)$ and $\phi^{(2)}(.)$ as:

$$\phi^{(1)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc}) = \begin{cases} 1 & \text{if the } m^{th} \text{ order statistic(o.s.) of } (X_{i1}, \ldots, X_{ic}) < \text{ the } m^{th} \text{ o.s. of } (X_{j1}, \ldots, X_{jc}) \\ 0 & \text{otherwise} \end{cases}$$

$$\phi^{(2)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc}) = \begin{cases} 1 & \text{if the } (c-m+1)^{th} \text{ o.s. of } (X_{i1}, \ldots, X_{ic}) < \text{ the } (c-m+1)^{th} \text{ o.s. of } (X_{j1}, \ldots, X_{jc}) \\ 0 & \text{otherwise} \end{cases}$$
and let

$$\phi(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc}) = \phi^{(1)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc}) + \phi^{(2)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc})$$

2 if the $m$th order statistic of $(X_{i1}, \ldots, X_{ic})$ is less than the $m$th order statistic of $(X_{j1}, \ldots, X_{jc})$ and the $(c-m+1)$th order statistic of $(X_{i1}, \ldots, X_{ic})$ is less than the $(c-m+1)$th order statistic of $(X_{j1}, \ldots, X_{jc})$

1 if either the $m$th order statistic of $(X_{i1}, \ldots, X_{ic})$ is less than the $m$th order statistic of $(X_{j1}, \ldots, X_{jc})$ or the $(c-m+1)$th order statistic of $(X_{i1}, \ldots, X_{ic})$ is less than the $(c-m+1)$th order statistic of $(X_{j1}, \ldots, X_{jc})$ but not both

0 otherwise.

The U-statistic based on this kernel is

$$U_{ij}^{m,c} = \binom{n_i}{r_c} \binom{n_j}{s_c}^{-1} \sum \phi(X_{ir_1}, \ldots, X_{ir_c}; X_{js_1}, \ldots, X_{js_c}),$$

where the summation is extended over all combinations of $c$ integers $(r_1, \ldots, r_c)$ chosen without replacement from $(1, \ldots, n_i)$ and all combinations of $c$ integers $(s_1, \ldots, s_c)$ chosen without replacement from $(1, \ldots, n_j)$.

The U-statistics $U_{ij}^{m,c}$ can also be expressed as

$$U_{ij}^{m,c} = U_{ij}^{m,c(1)} + U_{ij}^{m,c(2)},$$

where $U_{ij}^{m,c(1)}$ and $U_{ij}^{m,c(2)}$ are the U-statistics associated with the kernels $\phi^{(1)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc})$ and $\phi^{(2)}(X_{i1}, \ldots, X_{ic}; X_{j1}, \ldots, X_{jc})$, respectively.
The expected value of $U_{ij}^{m,c}$ under $\mu_i = \mu_j$ is

$$E[U_{ij}^{m,c}] = \binom{C}{m} \sum_{t=m}^{c} \binom{C}{t} \beta [m + t, 2c-m+1-t]$$

$$+ \sum_{t=c-m+1}^{c} \binom{C}{t} \beta [c - m + 1 + t, c + m - t]$$

$$= 1 \text{ for all } c \text{ and } m.$$

Define

$$T_{ij}^{m,c} = U_{ij}^{m,c} - 1.$$

The proposed selection procedures based on the statistics $T_{ij}^{m,c}$ are:

$$R_{ij}^{m,c}: \text{Select population } n_j \text{ in the subset if and only if}$$

$$T_{ij}^{m,c} \geq b \left[ C \rho_{m,c} (n_j^{-1} + n_{[1]}^{-1}) \right]^{1/2} \text{ for all } i, i \neq j.$$

Here $\rho_{m,c}$ is a constant which depends upon the choice of $m$ and $c$ (see expression (4.5.4)) of chapter IV and the constant $b$ is chosen such that

$$P_0 \left[ T_{ij}^{m,c} \geq b \left[ C \rho_{m,c} (n_j^{-1} + n_{[1]}^{-1}) \right]^{1/2} \text{ for all } i, i \neq j \right] \geq \alpha,$$

where $P_0$ indicates that the probability is computed under $\mu_1 = \ldots = \mu_k$.

The asymptotic relative efficiency (ARE) of our procedures relative to some of the existing procedures is derived. It is shown that the proposed selection procedures can be implemented approximately with the help of existing tables and a practical example has been considered to illustrate the proposed selection procedures. A modification of the proposed procedures to select a subset containing the populations better than an unknown control
population is given. Simulation study has been carried out to assess the asymptotic performance of the proposed selection procedures in terms of expected subset sizes for particular parametric configurations of underlying populations.

In chapter V, we have developed a class of selection procedures for selecting the least dispersive distribution from the k available distributions. In the field of reliability and engineering, it is always useful to select the population which is least dispersive.

Let \( \Pi_1, \Pi_2, \ldots, \Pi_k \) be k independent populations with unknown continuous distribution functions \( F_1, F_2, \ldots, F_k \), respectively.

Marzec and Marzec (1991) defined

\[
\Delta_a(F) = \int_0^1 t_a(x) F^{-1}(x) \, dx, \quad 0 < a \leq 1/2, \quad \quad \quad \ldots (1.4.1)
\]

where

\[
t_a(x) = \begin{cases} 
  x[x/(2a(1-a)] - 1 & \text{if } 0 \leq x \leq a \\
  (2x-1)a/[2(1-a)] & \text{if } a < x \leq 1-a \\
  (1-x)[x/(2a(1-a)]+1-1/[2a(1-a)] & \text{if } 1-a < x \leq 1 
\end{cases}
\]

as a measure of dispersive ordering and have used for testing \( H_0 : \text{disp disp} F_i = F_j \), \( i \neq j \), against the alternative \( H : F_i \text{ disp } F_j \). 

**Definition 1.4.1:** A distribution function \( F_i \) is said to be less dispersive than a distribution function \( F_j \) if \( \Delta_a(F_1) \leq \Delta_a(F_2) \), \( 0 < a \leq 1/2 \).

We use the value of \( \Delta_a(F) \) as the criterion for selecting the least dispersive distribution function from k distribution functions. The population with the smallest \( \Delta_a \) value, termed as
the least dispersive, is labelled as the "best" population.

The proposed selection procedure is based on the L-statistic (see Serfling(1980)), an estimator of $\hat{\lambda}_a(F)$, defined as

\[
L_{a,n}(F) = \frac{1}{n} \sum_{i=1}^{n} t_{a} \left( \frac{i}{n} \right) X_{(i)},
\]

where $X_{(i)}$ denotes the $i$th order statistic corresponding to the sample $X_1, \ldots, X_n$ of size $n$ from distribution $F$.

Asymptotic normality of the statistic $L_{a,n}(F)$ for selecting the least dispersive distribution function from $k$ distribution functions has been used. The asymptotic variances of statistics $L_{a,n}(F), i=1,\ldots,k$ are unknown and unequal. Thus, with large samples, this selection problem is to be considered as a selection problem for the means of $k$ normal distributions with unknown and unequal variances. Due to non-existence of single stage procedure in such circumstances (see Dudewicz(1971) and Dudewicz and Dalal(1975)), a two stage procedure has been proposed.

The performance of the proposed selection procedures has been assessed through simulation study. Some applications of the proposed selection procedures has been discussed by taking exponential, gamma and Lehmann type distributions.

In chapter VI we have considered $k$ populations differing with respect to their scale parameters. We extend the results of Gutmann and Maymin(1987), derived for location parameters under
the normal probability model to the general scale probability model, in order to attach the best possible p-value to the inference: the selected population is the best.

Consider \( k \) independent populations \( \Pi_1, \Pi_2, \ldots, \Pi_k \) such that the population \( \Pi_i \) is characterized by an unknown scale parameter \( \theta_i (>0), \ i=1, \ldots, k. \) Let the sufficient statistic \( T_i \) for \( \theta_i \), based on a random sample of size \( n \) from population \( \Pi_i \), have the probability density function (pdf) \( f_{\theta_i}(x) = (1/\theta_i) f(x/\theta_i) \), with corresponding cumulative distribution function (cdf) \( F_{\theta_i}(x) = F(x/\theta_i), \ x > 0, \ \theta_i > 0, i=1, \ldots, k. \) \( F(.) \) is an arbitrary absolutely continuous cdf with pdf \( f(.) \). Call the population associated with the smallest \( \theta_i \) the "best" population and if \( T_S = \min(T_1, \ldots, T_k) \), we call \( \Pi_S \) the "selected" population. The goal is to attach the best possible p-value to the inference: the selected population is the best.

Let \( S \) be a random variable such that \( T_S = \min(T_1, \ldots, T_k) \). Suppose we infer that the selected population is the best, i.e., \( \theta_S < \min_{i \neq S} \theta_i \), whenever \( (T_1, \ldots, T_k) \in T \), for some \( k \)-dimensional set to be specified below. Then the probability of an error is

\[
P_\theta (\{T_1, \ldots, T_k\} \in T, \ \theta_S \geq \min_{i \neq S} \theta_i),
\]

and the probability of correct inference is

\[
P_\theta (\{T_1, \ldots, T_k\} \in T, \ \theta_S < \min_{i \neq S} \theta_i),
\]

where \( \theta = (\theta_1, \ldots, \theta_k)^t \).

We want to maximize (1.4.3) subject to the probability
defined in (1.4.2) as $\alpha$ (probability of type I error) for all $\theta$.

If the inference $\hat{\theta}_S < \min_{i \neq S} \theta_i$ is made whenever $T_S / \min_{i \neq S} T_i < c$ (here $c \in (0,1)$), then we have shown that the probability of a false inference is maximized when two $\theta_i$'s are equal and the rest are infinite.

An optimality property for the procedure is discussed. A reinterpretation of the results in terms of a conditional hypothesis testing set-up is also given. Application of the results to the exponential and normal probability models are discussed. In case of the exponential distribution, Type-II censored data is also considered.