Selection and Ranking Procedures in Reliability Models*

by

Shanti S. Gupta and S. Panchapakesan
Purdue University and Southern Illinois University

Technical Report #85-16

Department of Statistics
Purdue University

July 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.

ABSTRACT

Several procedures appropriate for selecting the best from $k(\geq 2)$ life length distributions are discussed. The selection problem is considered (Section 2) under the classical formulation using both the indifference zone and subset selection approaches, but mostly the latter. Under parametric models (Section 3), procedures are described for selection from gamma, Weibull, and exponential (one- and two-parameter) distributions. Nonparametric procedures and distribution-free procedures are discussed (Section 4) for selection in terms of a specified quantile and reliability at an arbitrarily chosen point in time, respectively. Procedures for selection from restricted families are discussed (Section 6) with special reference to IFR and IFRA families. The procedures described in various cases mentioned above illustrate several modifications and generalizations of the basic goal.

Key words and phrases: Indifference zone, subset selection, restricted subset, gamma, exponential, Weibull, binomial, nonparametric and distribution-free procedures, restricted families, IFR and IFRA families, convex and tail ordering, quantile selection, total life statistic, comparison with control.

AMS 1980 subject classifications: Primary 62F07; Secondary 62N05, 62-02.

SELECTION AND RANKING PROCEDURES IN RELIABILITY MODELS*

Shanti S. Gupta and S. Panchapakesan
Purdue University and Southern Illinois University

1. Introduction

Situations abound in practice where the aim of the statistical analyst is to compare two or more populations in some fashion with a view to rank them or select the best one(s) among them. For example, a purchasing firm may want to determine which one of several competing suppliers of components for a certain computer is producing the highest quality product. Typically, the populations that are compared will be life length distributions of the components from the competing manufacturers. The best population could be defined as the one with the largest mean life or with the largest quantile (percentile) of a given order. In such situations, the classical tests of homogeneity are not designed to answer efficiently several possible questions of interest to the experimenter. Selection and ranking procedures were initially devised in the early 1950's to provide the analyst appropriate tools to answer these questions. Most of the investigations in the last thirty odd years have adopted one or the other of two basic formulations. One of them is the so-called <u>indifference zone</u> (IZ) <u>formulation</u> of Bechhofer (1954) and the other is the <u>subset selection</u> (SS) appraoch of Gupta (1956).

Our main purpose in this paper is to describe some important selection procedures that are relevant to reliability models. Selection procedures are available in the literature for various parametric families of distributions. Many of these distributions serve as appropriate models for the life length of a unit. However, we will be concerned with only a few of these such as

*This reserach 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.

exponential, gamma, and Weibull distributions. Besides some nonparametric and distribution-free procedures, we emphasize selection procedures for restricted families of distributions such as the increasing failure rate (IFR) and increasing failure rate on the average (IFRA) families which are of importance in reliability problems. In dealing with these procedures, we mainly use the SS approach.

In the last three decades and more, the literature on selection and ranking procedures has grown enormously. Several books have appeared exclusively dealing with selection and ranking procedures. Of these, the monograph of Bechhofer, Kiefer and Sobel (1968) deals with sequential procedures with special emphasis on Koopman-Darmois family. Gibbons, Olkin and Sobel (1977) deal with methods and techniques mostly under the IZ formulation. Gupta and Panchapakesan (1979) provide a comprehensive survey of the developments in the field of ranking and selection, with a special chapter on Guide to Tables. They deal with all aspects of the problem and provide an extensive bibliography. Büringer, Martin and Schriever (1980) and, Gupta and Huang (1981) have discussed some specific aspects of the problem. A fairly comprehensive categorized bibliography is provided by Dudewicz and Koo (1982). For a critical review and an assessment of developments in subset selection theory and techniques, reference may be made to Gupta and Panchapakesan (1985).

Section 2 discusses the formulation of the basic problem of selecting the best population using the IZ and SS approaches. Section 3 deals with selection from gamma, exponential and Weibull populations. Procedures for different generalized goals are discussed using both IZ and SS approaches. Nonparametric procedures are discussed in Section 4 for selecting in terms of α-quantiles. This section also discusses procedures for Bernoulli distributions. These serve as distribution-free procedures for selecting from life distributions in terms of reliability at an arbitrarily chosen time. Procedures for selection from restricted families of distributions are described in Section 5. These include procedures for IFR and IFRA families in particular. A brief discussion of selection in comparison with a standard or control follows in Section 6.

2. Selection and Ranking Procedures

Let $\pi_1,\dots,\pi_k$ be k given populations where π_i has the associated distribution function F_{θ_i}, $i = 1,\dots,k$. The θ_i are real-valued parameters taking values in the set Θ. It is assumed that the θ_i are unknown. The ordered θ_i are denoted by $\theta_{[1]} \leq \theta_{[2]} \leq \dots \leq \theta_{[k]}$ and the (unknown) population π_i associated with $\theta_{[i]}$ by $\pi_{(i)}$, $i = 1,\dots,k$. The populations are ranked according to their θ-values. To be specific, $\pi_{(j)}$ is defined to be better than $\pi_{(i)}$ if $i < j$. No prior information is assumed regarding the true pairing between $(\theta_1,\dots,\theta_k)$ and $(\theta_{[1]},\dots,\theta_{[k]})$.

2.1 Indifference Zone (IZ) Formulation

The goal in the basic problem in the IZ approach is to select the best population, namely, the one associated with $\theta_{[k]}$. A procedure is required to choose one of the populations. A correct selection (CS) is a selection of population(s) satisfying the goal. Here CS corresponds to choosing the best population. Any selection procedure is required to guarantee a minimum probability of a correct selection (PCS). In the IZ formulation, this requirement is that, for any rule R,

$$(2.1) \qquad P(CS|R) \geq P^* \quad \text{whenever} \quad \delta(\theta_{[k]},\theta_{[k-1]}) \geq \delta^*,$$

where P(CS|R) denotes the PCS using R, and $\delta(\theta_{[k]},\theta_{[k-1]})$ is an appropriate measure of separation of the best population $\pi_{(k)}$ from the next best $\pi_{(k-1)}$. The constants P* and δ^* are specified by the experimenter in advance. The statistical problem is to define a selection rule which really consists of a sampling rule, a stopping rule for sampling, and a decision rule. If we consider taking a single sample of fixed size n from each population, then the minimum value of n is determined subject to (2.1). A crucial step involved in this is to evaluate the infimum of the PCS over $\Omega_{\delta^*}=\{\theta=(\theta_1,\ldots,\theta_k): \delta(\theta_{[k]},\theta_{[k-1]}) \geq \delta^*\}$. Any configuration of $\underset{\sim}{\theta}$ where this infimum is attained is called a least favorable configuration (LFC). Between two valid (i.e. satisfying (2.1)) single sample procedures, the sample size n is an obvious criterion for efficiency comparison. The region Ω_{δ^*} is called the preference zone. No requirement is made regarding the PCS when $\underset{\sim}{\theta}$ belongs to the complement of Ω_{δ^*} which, in fact, is the indifference zone.

2.2 Subset Selection (SS) Approach

In the SS approach for selecting the best population, the goal is to select a nonempty subset for the k populations which includes the best population. The size of the selected subset is not fixed in advance; it is rather determined by the data themselves. Selection of any subset consistent with the goal (i.e. including the best population) is a correct selection. It is required that, for any rule R,

$$(2.2) \qquad P(CS|R) \geq P^* \qquad \text{for all } \underset{\sim}{\theta} \in \Omega$$

where $\Omega = \{\underset{\sim}{\theta}\}$ is the whole parameter space. It should be noted that there is no indifference zone specification in this formulation. As is to be expected, a crucial step is the evaluation of the infimum of the PCS over Ω. Any subset selection rule that satisfies (2.2) meets the criterion of validity. Denoting the selected subset by S and its size by $|S|$, the expected value of $|S|$ serves as a reasonable measure for efficiency comparison between valid procedures. Besides $E(|S|)$, possible performance characteristics include $E(|S|)$ - PCS and $E(|S|)/PCS$. The former one represents the expected number of nonbest populations included in the selected subset. As an overall measure, one can also consider the supremum of $E(|S|)$ over Ω.

2.3 Some General Remarks

The probability requirement, (2.1) or (2.2) as the case may be, is usually referred to as the basic probability requirement, or the P*-requirement, or the P*-condition. There are several modifications and generalizations of the basic goal and requirements on the procedures in both IZ and SS formulations. These will be described as the necessity arises during our discussion of several procedures. For details on these aspects of the problem, reference may be made to Gupta and Panchapakesan (1979).

Suppose that the best population is the one associated with the largest θ_i. A procedure R is said to be monotone if the probability of selecting π_i is at least as large as that of selecting π_j whenever $\theta_i > \theta_j$.

2.4 Two Types of Subset Selection Rules

Let T_i be the statistic associated with the sample from π_i $(i = 1,\ldots,k)$ with distribution function $F(x,\theta_i)$; the θ_i are the parameters to be ranked. Most of the rules that have been studied in the literature are of one of the following types:

R_1: Select π_i if and only if

(2.3) $$T_i \geq \max_{1\leq j\leq k} T_j - d$$

and

R_2: Select π_i if and only if

(2.4) $$T_i \geq c \max_{1\leq j\leq k} T_j$$

where $d>0$ and $c \in (0,1)$ are to be determined so that the P*-requirement is satisfied.

These rules R_1 and R_2 have been typically proposed when θ_i is a location and a scale parameter, respectively. When θ_i is neither a location nor a scale parameter (e.g. a noncentrality parameter), usually one of these two rules has been proposed depending on the nature of the support of T_i. Most of the rules that are discussed in this paper come under one of these two types. Treatment of R_1 and R_2 in the location and scale case, respectively, is given in Gupta (1965). The following properties hold for R_1 in the location case and R_2 in the scale case.

(1) The procedure is monotone (Gupta, 1965).

(2) If the distribution $F(x,\theta)$ possesses a density $f(x,\theta)$ having a monotone likelihood ratio (MLR) in x, then $E(|S|)$ is maximized when $\theta_1 = \ldots = \theta_k$ and this maximum is kP* (Gupta, 1965).

(3) Under the MLR assumption, the rule is minimax when the loss is measured by $|S|$ or the number of non-best populations selected (Berger, 1979).

(4) In a fairly large class of rules, the procedure is minimax when the loss is measured by the maximum probability of including a non-best population (Berger and Gupta, 1980).

A comprehensive unified Theory is due to Gupta and Panchapakesan (1972), who have considered a class of rules which includes R_1 and R_2 as special cases; see Gupta and Panchapakesan (1979, Section 11.2). Gupta and Huang (1980) have obtained an optimal rule in the class of rules for which the PCS is at least γ by minimizing the supremum of $E(|S|)$.

3. Selection from Parametric Families.

Numerous parametric models are employed in the analysis of life length data and in problems connected with the modeling of aging or failure processes. Among univariate models, a few particular distributions, namely, the exponential, Weibull, and gamma, stand out in view of their proven usefulness in a wide range of situations. Of course, these distributions are related to each other. In this section, we will discuss a few typical procedures for selection from these populations.

3.1 Selection from Gamma Populations

Let $\pi_1,\dots,\pi_k$ denote k given gamma populations with density functions

$$(3.1) \qquad f(x,\theta_i) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\theta_i^{\alpha}} \exp(-x/\theta_i),\ x > 0,\ \theta_i,\ \alpha > 0,\quad i = 1,\dots,k,$$

with a common known shape parameter α. For the goal of selecting a subset containing the best population, namely, the one associated with $\theta_{[k]}$, Gupta (1963a) proposed a rule based on the sample means $\bar{X}_i$, i = 1,...,k, arising from n independent observations from each population. The rule of Gupta (1963a) is

R_3: Select π_i if and only if

(3.2) $$\overline{X}_i \geq c \max_{1 \leq j \leq k} \overline{X}_j$$

where c is the largest number with $0 < c < 1$ for which the P*-requirement is met. The LFC is given by $\theta_1 = \ldots = \theta_k$ and the constant c is determined by

(3.3) $$\int_0^\infty G_\nu^{k-1} (x/c)\, g_\nu(x)\, dx = P^*,$$

where $\nu = n\alpha$ and, G_ν and g_ν are the cdf and the density, respectively, of a standardized gamma random variable (i.e. with $\theta = 1$) with shape parameter ν. Gupta (1963a) has tabulated the values of c for $\nu = 1(1)25$, $k = 2(1)11$, and $P^* = .75, .90, .95, .99$. Tables 1A and 1B (at the end of this paper) are excerpted from the tables of Gupta (1963a) and they provide c-values for $k = 2(1)11$, $\nu = 1(1)20$, and $P^* = 0.90$ and 0.95, respectively.

Depending on the physical nature of the problem, we may be interested in selecting the population associated with $\theta_{[1]}$, which is the best population now. In this case, the procedure analogous to R_3 is

R_4: Select π_i if and only if

(3.4) $$\overline{X}_i \leq \frac{1}{c'} \min_{1 \leq j \leq k} \overline{X}_j$$

where $0 < c' < 1$ is the largest number for which the P*-condition is met. The constant c' is given by

(3.4) $$\int_0^\infty [1 - G_\nu (c'x)]^{k-1}\, g_\nu(x)\, dx = P^*$$

where $\nu = n\alpha$. The values of the constant c' have been tabulated for $\nu = 1(1)25$, $k = 2(1)11$, and $P^* = .75, .90, .95, .99$ by Gupta and Sobel (1962b) who have studied rule R_4 in the context of selecting from k normal populations the one with the smallest variance in a companion paper (1962a).

It is known that the gamma family $\{F(x,\theta)\}$, with common parameter r, is stochastically increasing in θ, i.e., $F(x,\theta_i)$ and $F(x,\theta_j)$ are distinct for $\theta_i \neq \theta_j$, and $F(x,\theta_i) \geq F(x,\theta_j)$ for all x when $\theta_i < \theta_j$. This implies that ranking them in terms of θ is equivalent to ranking in terms of α-quantile for any $0 < \alpha < 1$.

3.2 Selection from Exponential (One-Parameter) Populations

We first note that this is a special case of gamma populations with densities $f(x,\theta_i)$ in (3.1) with $\alpha = 1$. Thus the rules R_3 and R_4 are applicable. Now consider a life testing situation where a sample of n items from each population is put on test and the sample is censored (type-II) at the r-th failure. Let $X_{i1} < X_{i2} < \ldots < X_{ir}$ denote the r completed lives in the sample from π_i, $i = 1,\ldots,k$. Define

$$(3.5) \qquad T_i = \sum_{j=1}^{r} X_{ij} + (n-r) X_{ir}, \qquad i = 1,\ldots,k.$$

The T_i are the so-called total life statistics. It is well-known that $\frac{2T_i}{\theta_i}$ has a chi-square distribution with 2r degrees of freedom. In other words, T_i has a gamma distribution with scale parameter θ_i and shape parameter r. Thus for selecting the population with the largest mean life θ_i, the procedure R_3 (stated in terms of the T_i) will be

R_3: Select π_i if and only if

(3.6) $$T_i \geq c \max_{1\leq j\leq k} T_j$$

where c is given by (3.3) with $\nu = r$.

3.3 Selection from Two-Parameter Exponential Distributions

Let π_i have density

(3.7) $$f(x,\theta_i,\sigma) = \frac{1}{\sigma} \exp -\left\{\frac{(x-\theta_i)}{\sigma}\right\}, \quad x>\theta_i, \ \theta_i,\sigma > 0, \ i = 1,\ldots,k.$$

The density (3.7) provides a model for life length data when we assume a minimum guaranteed life θ_i, which is here a location parameter. It is assumed that all the k populations have a common scale parameter σ. The θ_i are unknown and our interest is in selecting the population associated with the largest θ_i. We will discuss some procedures under the IZ formulation. Consider the generalized goal of selecting a subset of fixed size s so that the t best populations ($1\leq t\leq s<k$) are included in the selected subset. This generalized goal was introduced by Desu and Sobel (1968). The special case of t=s, namely, that of choosing t populations so that they are the t best, was considered originally by Bechhofer (1954). When s=t=1, we get the basic goal of selecting the best population. The probability requirement is that

(3.8) $$PCS \geq P^* \text{ whenever } \theta_{[k-t+1]} - \theta_{[k-t]} \geq \theta^* > 0$$

where θ^* and P^* are specified in advance and a correct selection occurs when a subset of s populations is selected consistent with the goal. Also, for a meaningful problem, we should have $\frac{1}{\binom{k}{t}} < P^* < 1$. In describing several procedures, we will adopt either the generalized goal or one of its special cases. We will consider the two cases of known and unknown σ separately.

Case A: Known σ. We can assume without loss of generality that $\sigma = 1$. Let X_{ij}, $j = 1,\ldots,n$, denote a sample of n observations from π_i, $i = 1,\ldots,k$. Define $Y_i = \min_{1 \le j \le n} X_{ij}$, $i = 1,\ldots,k$.

Raghavachari and Starr (1970) considered the goal of selecting the t best populations (i.e. $1 \le s = t$ k) and they studied the 'natural' rule

(3.9) R_5: Select the t populations associated with $Y_{[k-t+1]},\ldots,Y_{[k]}$.

The LFC for this rule is given by

$$(3.10) \qquad \begin{cases} \theta_{[1]} = \cdots = \theta_{[k-t]}; \\ \theta_{[k-t+1]} = \cdots = \theta_{[k]}; \\ \theta_{[k-t+1]} - \theta_{[k-t]} = \theta^*. \end{cases}$$

The minimum sample size required to satisfy (3.8) is the smallest integer n for which

$$(3.11) \qquad (1-e^{-n\theta^*})^{k-t} + (k-t)e^{n\theta^* t} I(e^{-n\theta^*}, t+1, k-t) \ge P^*$$

where

$$(3.12)\qquad I(z;\alpha,\beta) = \int_0^z u^{\alpha-1}(1-u)^{\beta-1}\,du,\ \alpha,\beta>0,\quad 0 \leq z \leq 1.$$

Equivalently, we need the smallest integer n such that

$$(3.13)\qquad n\theta^* \geq -\log\nu,$$

where ν $(0<\nu<1)$ is the solution of the equation

$$(3.14)\qquad (1-\nu)^{k-t} + (k-t)\nu^{-t}\, I(\nu,t+1,k-t) = P^*.$$

Raghavachari and Starr (1970) have tabulated the ν-values for k=2(1)15, t=1(1)k-1, and P* = .90, .95, .975, .99.

In particular, for selecting the best population, the equation (3.14) reduces to

$$(3.15)\qquad (\nu k)^{-1}\,[1-(1-\nu)^k] = P^*.$$

For the generalized goal, Desu and Sobel (1968) studied the following rule R_6.

R_6: Select the s populations associated with $Y_{[k-s+1]},\ldots,Y_{[k]}$. Given n, k, t, θ^*, and P*, they have shown that the smallest s for which the probability requirement (3.8) is satisfied is the smallest integer s such that

$$(3.16)\qquad \binom{s}{t} \geq P^* \binom{k}{t} e^{-nt\theta^*}.$$

It should be pointed out that Desu and Sobel (1968) have obtained general results for location parameter family. They have also considered the dual problem of selecting a subset of size $s(s \le t)$ so that all the selected populations are among the t best.

Case B: Unknown σ. In this case, we consider the basic goal of selecting the best population. Since σ is unknown, it is not possible to determine in advance the sample size needed for a single sample procedure in order to guarantee the P*-condition. This is similar to the situation that arises in selecting the population with the largest mean from several normal populations with a common unknown variance. For this latter problem, Bechhofer, Dunnett and Sobel (1954) proposed a non-elimination type two-stage procedure in which the first stage samples are utilized purely for estimating the variance without eliminating any population from further consideration. A similar procedure was proposed by Desu, Narula and Villarreal (1977) for selecting the best exponential population. Kim and Lee (1985) have studied an elimination type two-stage procedure analogous to that of Gupta and Kim (1984) for the normal means problem. In their procedure, the first stage is used not only to estimate σ but also to possibly eliminate non-contenders. Their Monte Carlo study shows that, when $\theta_{[k]} - \theta_{[k-1]}$ is sufficiently large, the elimination type procedure performs better than the other type procedure in terms of the expected total sample size.

The procedure R_7 of Kim and Lee (1985) consists of two stages as follows.

Stage 1: Take n_0 independent observations from each π_i $(1 \le i \le k)$, and compute $Y_i^{(1)} = \min_{1 \le j \le n_0} X_{ij}$, and a pooled estimate $\hat{\sigma}$ of σ, namely,

$$\hat{\sigma} = \sum_{i=1}^{k} \sum_{j=1}^{n_0} (X_{ij}-Y_i^{(1)})/k(n_0-1).$$

Determine a subset I of {1,...,k} defined by

$$I = \{i | Y_i^{(1)} \geq \max_{1\leq j\leq k} Y_j^{(1)} - (2k(n_0-1)\hat{\sigma}h/n_0-\theta^*)^+\},$$

where the symbol a^+ denotes the positive part of a, and h(>0) is a design constant to be determined.

(a) If I has only one element, stop sampling and assert that the population associated with $Y_{[k]}^{(1)}$ as the best.

(b) If I has more than one element, go to the second stage.

<u>Stage 2</u>: Take $N - n_0$ additional observations X_{ij} from each π_i for $i \in I$, where

$$N = \max \{n_0, <2k(n_0-1)\hat{\sigma}h/\theta^*>\},$$

and the symbol $<y>$ denotes the smallest integer equal to greater than y. Then compute, for the overall sample, $Y_i = \max_{1\leq j\leq N} X_{ij}$ and choose the population associated with $\max_{i\in I} Y_i$ as the best.

The constant h used in the procedure R_7 is given by

$$(3.17) \qquad \int_0^\infty \{1-(1-\alpha(x))^k\}^2/\{k^2\alpha^2(x)\} \ f_\nu(x)\ dx = P^*$$

where $\alpha(x) = \exp(-hx)$ and $f_\nu(x)$ is the chi-square density with $\nu = 2k(n_0-1)$ degrees of freedom. The h-values have been tabulated by Kim and Lee (1985) for $P^* = .95$, $k = 2(1)5(5)20$, and $n_0 = 2(1)30$.

3.4 Selection from Weibull Distributions.

Let π_i have a two-parameter Weibull distribution given by the cdf

$$(3.18) \qquad F_i(x) \equiv F(x;\theta_i,c_i) = 1 - \exp\{-(x/\theta_i)^{c_i}\},\ x>0,\ \theta_i,c_i > 0,\ i = 1,\ldots,k.$$

The c_i and θ_i are unknown. Kingston and Patel (1980a,b) have considered the problem of selecting from Weibull distributions in terms of their reliabilities (survival probabilities) at an arbitrary but specified time $L > 0$. The reliability at L for F_i $(i=1,\ldots,k)$ is given by

$$(3.19) \qquad \rho_i = 1 - F_i(L) = \exp\{-(L/\theta_i)^{c_i}\}.$$

We can without loss of generality assume that $L = 1$ because the observed failure times can be scaled so that $L = 1$ time unit. Further, letting $(\theta_i)^{c_i} = \lambda_i$, we get $\rho_i = \exp\{-\lambda_i^{-1}\}$. Obviously, ranking the populations in terms of the ρ_i is equivalent to ranking in terms of the λ_i, and the best population is the one associated with $\lambda_{[k]}$, the largest λ_i. Kingston and Patel (1980a) considered the problem of selecting the best one under the IZ formulation using the natural procedure based on estimates of the λ_i constructed from type II censored samples. They also considered the problem of selecting the best in terms of the α-quantiles for a given $\alpha \in (0,1)$, $\alpha \neq 1-e^{-1}$, in the case where $\theta_1 = \ldots = \theta_k = \theta$ (unknown). The α-quantile of F_i is given by $\xi_i = \theta[-\log(1-\alpha)]^{1/c_i}$ so that ranking in terms of the α-quantiles is equivalent to ranking in terms of the shape parameter. It should be noted that the ranking of the c_i is in the same order as that of the associated ξ_i if $\alpha < 1-e^{-1}$, and is in the reverse order if $\alpha > 1-e^{-1}$. The procedures discussed above are based on maximun likelihood estimators as well as simplified linear estimators (SLE) considered by Bain (1978, p. 265). For further details on these procedures, see Kingston and Patel (1980a).

In another paper, Kingston and Patel (1980b) considered the goal of selecting a subset of restricted size. This formulation, usually referred to as restricted subset selection (RSS) approach, is due to Gupta and Santner (1973) and Santner (1975). In the usual SS approach of Gupta (1956), it is possible that the procedure selects all the k populations. In the RSS approach, we restrict the size of the selected subset by specifying an upper bound $m(1 \leq m \leq k-1)$; the size of the selected subset is still random variable taking on values 1, 2,...,m. Thus it is a generalization of the usual approach (m=k). However, in doing so, an indifference zone is introduced. The selection goal can be more general than selecting the best. We now consider a generalized goal in the RSS approach for selection from Weibull populations, namely, to select a subset of the k given populations not exceeding m in size such that the selected subset contains at least s of the t best populations. As before, the populations are ranked in terms of their λ-values. Note that $1 \leq s \leq \min(t,m) \leq k$. The probability requirement now is that

(3.20) $$PCS \geq P^* \quad \text{whenever} \quad \underset{\sim}{\lambda} = (\lambda_1,\ldots,\lambda_k) \in \Omega_{\lambda^*}$$

where

(3.21) $$\Omega_{\lambda^*} = \{\underset{\sim}{\lambda}: \lambda^* \lambda_{[k-t]} \leq \lambda_{[k-t+1]}, \lambda^* \geq 1\}.$$

When t=s=m and $\lambda^* > 1$, the problem reduces to selecting the t best population using the IZ formulation. When $s=t < m=k$ and $\lambda^* = 1$, the problem reduces to selecting a subset of random size containing the t best population (the usual SS approach). Thus the RSS approach integrates the formulations of Bechhofer (1954), Gupta (1956), and Desu and Sobel (1968). General theory under the RSS approach is given by Santner (1975).

Returning to the Weibull selection problem with the generalized RSS goal, Kingston and Patel (1980b) studied a procedure based on type II censored samples from each population. It is defined in terms of the maximum likelihood estimators (or the SLE estimators) $\hat{\lambda}_i$. This procedure is

R_8: Include π_i in the selected subset if and only if

$$\hat{\lambda}_i \geq \max\{\hat{\lambda}_{[k-m+1]},\ c\,\hat{\lambda}_{[k-\ell+1]}\}, \tag{3.22}$$

where $c \in [0,1]$ is suitably chosen to satisfy (3.20).

Let n denote the common sample size and consider censoring each sample at the rth failure. For given k, r, n, s, t, and m, we have three quantities associated with the procedure R_8, namely, P^*, c, and $\lambda^* > 0$. Given two of these, one can find the third; however, the solution may not be admissible. For example, for some P^* and λ^*, there may not be a constant $c \in [0,1]$ so that (3.20) is satisfied unless m=k. Kingston and Patel (1980b) have given a few tables of λ^*-values for selected values of other constants. Their table values are based on Monte Carlo techniques and the choice of SLE's.

4. Nonparametric and Distribution-Free Procedures

Parametric families of distributions serve as life models in situations where there are strong reasons to select a particular family. For example, the model may fit data on hand well, or there may be a good knowledge of the underlying aging or failure process that indicates the appropriateness of the model. But there are many situations in which it becomes desirable to avoid strong assumptions about the model. Nonparametric or distribution-free procedures are important in this context.

Gupta and McDonald (1982) have surveyed nonparametric selection and ranking procedures applicable to one-way classification, two-way classification, and paired-comparison models. These procedures are based on rank scores and/or robust estimators such as the Hodges-Lehmann estimator. For the usual types of procedures based on ranks, the LFC is not always the one corresponding to identical distributions. Since all these nonparametric procedures are relevant in the context of selection from life length distributions, the reader is best referred to the survey papers of Gupta and McDonald (1982), and Gupta and Panchapakesan (1985), and Chapters 8 and 15 of Gupta and Panchapakesan (1979).

There have been some investigations of subset selection rules based on ranks while still assuming that the distributions associated with the populations are known. This is appealing especially in situations in which the order of the observations is more readily available than the actual measurements themselves due, perhaps, to excessive cost or other physical constraints. Under this setup, Nagel (1970), Gupta, Huang and Nagel (1979), Huang and Panchapakesan (1982), and Gupta and Liang (1984) have investigated locally optimal subset selection rules which satisfy the validity criterion that the infimum of the PCS is P* when the distributions are identical. They have used different optimality criteria in some neighborhood of an equi-parameter point in the parameter space. An account of these rules is given in Gupta and Panchapakesan (1985).

Characterizations of life length distributions are provided in many situations by so-called restricted families of distributions which are defined by partial order relations with respect to known distributions.

Well-known examples of such families are those with increasing (decreasing) failure rate and increasing (decreasing) failure rate average. Selection procedures for such families will be discussed in the next section.

In the remaining part of this section, we will be mainly concerned with nonparametric procedures for selection in terms of a quantile and selection from several Bernoulli distributions. Though the Bernoulli selection problem could have been discussed under parametric model, it is discussed here to emphasize the fact that we can use the Bernoulli selection procedures as distribution-free procedures for selecting from unknown continuous (life) distributions in terms of reliability at any arbitrarily-chosen time point L.

4.1 Selection in terms of Quantiles

Let $\pi_1,\ldots,\pi_k$ be k populations with continuous distributions $F_i(x)$, $i = 1,\ldots,k$, respectively. Given $0 < \alpha < 1$, let $x_\alpha(F)$ denote the αth quantile of F. It is assumed that the α-quantiles of the k populations are unique. The populations are ranked according to their α-quantiles. The population associated with the largest α-quantile is defined to be the best. Rizvi and Sobel (1967) proposed a procedure for selecting a subset containing the best. Let n denote the common size of the samples from the given populations and assume n to be sufficiently large so that $1 \leq (n+1)\alpha \leq n$. Let r be a positive integer such that $r \leq (n+1)\alpha < r+1$. It follows that $1 \leq r \leq n$. Let $Y_{j,i}$ denote the jth order statistic in the sample from π_i, $i = 1,\ldots,k$. The procedure of Rizvi and Sobel (1967) is

R_9: Select π_i if and only if

$$(4.1) \qquad Y_{r,i} \geq \max_{1\leq j\leq k} Y_{r-c,j}$$

where c is the smallest integer with $1 \le c \le r-1$ for which the P*-condition is satisfied.

For the procedure R_9, the infimum of the PCS is attained when the distributions $F_1,\ldots,F_k$ are identical and it is shown by Rizvi and Sobel (1967) that c is the smallest integer with $1 \le c \le r-1$ satisfying

$$(4.2) \qquad \int_0^1 G_{r-c}^{k-1}(u)\, dG_r(u) \ge P^*$$

where

$$(4.3) \qquad G_r(u) = \frac{n!}{(r-1)!(n-r)!} u^{r-1}(1-u)^{n-r-1}, \quad 0 \le u \le 1.$$

Rizvi and Sobel have shown that the maximum permissible value of P* such that a c-value satisfying (4.2) exists is $P_1 = P_1(n,\alpha,k)$ given by

$$(4.4) \qquad P_1 = \binom{n}{r} \sum_{i=0}^{k=1} \frac{(-1)^i \binom{k-1}{i}}{\binom{n(i+1)}{r}} .$$

A short table of P_1-values is given by Rizvi and Sobel for $\alpha = 0.5$ and $k = 2(1)10$. The n-values range from 1 in steps of 2 to a value (depending on k) for which P_1 gets very close to 1. Also given by them is a table of the largest value of r-c for $\alpha = \frac{1}{2}$ (which means that $r = (n+1)/2$), $k = 2(1)10$, $n = 5(10)95(50)495$, and $P^* = 0.75, 0.90, 0.95, 0.975, 0.99$. For the IZ approach to this selection problem, see Sobel (1967).

4.2 Distribution-free Procedures Using Bernoulli Model

Let $\pi_1,\ldots,\pi_k$ be k populations with the associated continuous (life) distributions $F_1,\ldots,F_k$, respectively. The reliability of π_i at L is

$\rho_i = 1 - F_i(L)$. Let X_{ij}, $j = 1,\ldots,n$, be sample observations from π_i, $i = 1,\ldots,k$. Define

$$(4.1) \qquad Y_{ij} = \begin{cases} 1 & \text{if } X_{ij} > L, \\ 0 & \text{otherwise,} \end{cases} \qquad i = 1,\ldots,k; \quad j = 1,\ldots,n.$$

Then $Y_{i1},\ldots,Y_{in}$ are independent and identically distributed Bernoulli random variables with success probability ρ_i, $i = 1,\ldots,k$. We are interested in selecting the population associated with the largest ρ_i.

Gupta and Sobel (1960) proposed a subset selection rule based on $Y_i = \sum_{j=1}^{n} Y_{ij}$, $i = 1,\ldots,k$. Their rule is

R_{10}: Select π_i if and only if

$$(4.2) \qquad Y_i \geq \max_{1 \leq j \leq k} Y_j - D$$

where D is the smallest nonnegative integer for which the P*-requirement is met.

An interesting feature of Procedure R_{10} is that the infimum of the PCS occurs when $\rho_1 = \ldots = \rho_k = \rho$ (say) but it is not independent of their common value ρ. For k=2, Gupta and Sobel (1960) showed that the infimum takes place when $\rho = \frac{1}{2}$. When k>2, the common value of ρ_0 for which the infimum takes place is not known. However, it is known that this common value $\rho_0 \to \frac{1}{2}$ as $n \to \infty$. An improvement in the situation is provided by Gupta, Huang and Huang (1976) who investigated conditional selection rules and, using the conditioning argument, obtained a conservative value of d. Their conditional procedure is

R_{11}: Select π_i if and only if

(4.3) $$Y_i \geq \max_{1\leq j\leq k} Y_j - D(t)$$

given $T = \sum_{i=1}^{k} Y_i = t$, where $D(t) > 0$ is chosen to satisfy the P*-condition. Exact result for the infimum of the PCS is obtained only for k=2; in this case, the infimum is attained when $\rho_1 = \rho_2 = \rho$ and is independent of the common value ρ. For k>2, Gupta, Huang and Huang (1976) obtained a conservative value for D(t) and also for D of Rule R_{10}. They have shown that $\inf P(CS|R_{11}) \geq P^*$ if D(t) is chosen such that

(4.4) $$D(t) = \begin{cases} d(t) & \text{for } k = 2 \\ \max\{d(r): r = 0,1,\ldots,\min(t,2n) & \text{for } k > 2 \end{cases}$$

where d(r) is defined as the smallest value such that

(4.5) $$N(2;d(r),r,n) \geq \begin{cases} P^* \binom{2n}{r} & \text{for } k = 2 \\ [1-(1-P^*)(k-1)^{-1}] \binom{2n}{r} & \text{for } k > 2 \end{cases}$$

and $N(k;d(t),t,n) = \sum \binom{n}{s_1}\cdots\binom{n}{s_k}$, the summation being over the set of all nonnegative integers s_i such that $\sum_{i=1}^{k} s_i = t$ and $s_k \geq \max_{1\leq j\leq k-1} s_j - d(t)$.

A conservative constant d for Procedure R_{10} is given by $d \equiv \max_{0\leq t\leq kn} d(t)$. Gupta, Huang and Huang (1976) have tabulated the smallest value d(t) satisfying (4.5) for k = 2,4(1)10, n = 1(1)10, t = 1(1)20, and P* = 0.75, 0.90, 0.95, 0.99.

They have also tabulated the d-values (conservative) for Procedure R_{10} for $P^* = 0.75, 0.90, 0.95, 0.99$, and $n = 1(1)4$ when $k = 3(1)15$, and $n = 5(1)10$ when $k = 3(1)5$.

Under the IZ formulation, one can use the procedure of Sobel and Huyett (1957) for selecting the population associated with the largest ρ_i which guarantees a minimum PCS P^* whenever $\rho_{[k]} - \rho_{[k-1]} \geq \Delta^* > 0$. Based on samples of size n from each population, their procedure based on the Y_i defined in (4.1) is

(4.6) R_{12}: Select the population associated with the largest Y_i, using randomization to break ties, if any.

The sample size required is the smallest n for which the PCS $\geq P^*$ when $\rho_{[1]} = \cdots = \rho_{[k-1]} = \rho_{[k]} - \Delta^*$, the LFC in this case. Sobel and Huyett (1957) have tabulated the sample sizes (exact and approximate) for $k = 2,3,4,10$; $\Delta^* = 0.05(0.05)0.50$, and $P^* = 0.50, 0.60, 0.75(.05)\ 0.95, 0.99$.

When n is large, the normal approximation to the PCS yields

$$n \approx c^2(1-\Delta^{*2})/4\Delta^{*2} \tag{4.7}$$

where $c = c(k,P^*)$ is the constant satisfying

$$\int_{-\infty}^{\infty} \Phi^{k-1}(x+c)\ \varphi(x)\ dx = P^* \tag{4.8}$$

and, Φ and φ denote correspondingly the cdf and density of the standard normal distribution. The c-value can be obtained from tables of Bechhofer (1954), Gupta (1963b), Milton (1963) and Gupta, Nagel and Panchapakesan (1973) for several selected values of k and P^*.

The Bernoulli selection problem has applications to the drug selection problem and to clinical trials. This fact has spurred lots of research activity involving investigations of selection procedures using sampling procedures such as the play-the-winner (PW) sampling rule (introduced by Robbins, 1952 and 1956) and vector-at-a-time (VT) rule with a variety of stopping rules. One of the main considerations in many of these procedures is to design the sampling rule so as to minimize the expected total number of observations and/or the expected number of observations from the worst population. Some of these procedures suffer from one drawback or another. For excellent review/survey/comprehensive assessment of these (and other) procedures, reference should be made to Bechhofer and Kulkarni (1982), Büringer, Martin and Schriever (1980), Gupta and Panchapakesan (1979, Sections 4.2 through 4.6), and Hoel, Sobel and Weiss (1975). For corresponding developments in subset selection theory, see Gupta and Panchapakesan (1979, Section 13.2).

5. Selection from Restricted Families of Distributions

A restricted family of probability distributions is defined by a partial order relation with respect to a known distribution. As we have pointed out earlier, such families provide characterizations of life length distributions. Selection rules for such restricted families were first considered by Barlow and Gupta (1969). We define below the binary partial order relations ($\underset{\sim}{<}$) that have been used in studying selection procedures. These are partial ordering in the sense that they enjoy only reflexivity and transitivity properties, that is, (1) $F \underset{\sim}{<} F$ for all distributions F, and (2) $F \underset{\sim}{<} G$, $G \underset{\sim}{<} H$ implies $F \underset{\sim}{<} H$. Note that $F \underset{\sim}{<} G$ and $G \underset{\sim}{<} F$ do not necessarily imply $F \equiv G$.

<u>Definitions 5.1</u> (1) F is said to be <u>convex</u> with respect to G ($F \underset{c}{<} G$) if and only if $G^{-1}F(x)$ is convex on the support of F.

(2) F is said to be <u>star-shaped</u> with respect to G ($F \underset{*}{<} G$) if and only if $F(0) = G(0) = 0$, and $G^{-1}F(x)/x$ is increasing in $x \geq 0$ on the support of F.

(3) F is said to be <u>r-ordered</u> with respect to G ($F \underset{r}{<} G$) if and only if $F(0) = G(0) = \frac{1}{2}$ and $G^{-1}F(x)/x$ is increasing (decreasing) in x positive (negative).

(4) F is said to be <u>tail-ordered</u> with respect to G ($F \underset{t}{<} G$) if and only if $F(0) = G(0) = \frac{1}{2}$ and $G^{-1}F(x) - x$ is increasing on the support of F.

It is well-known that convex ordering implies star ordering. Further, when $G(x) = 1-e^{-x}(x \geq 0)$, $F \underset{c}{<} G$ is equivalent to saying that F has an increasing failure rate (IFR) and $F \underset{*}{<} G$ is equivalent to saying that F has an increasing failure on the average (IFRA). Of course, if F is IFR, then it is also IFRA. IFR distributions were first studied in detail by Barlow, Marshall and Proschan (1963) and IFRA distributions by Birnbaum, Esary and Marshall (1966). The r-ordering was investigated by Lawrence (1975). Doksum (1969) used the tail-ordering. The convex ordering ans s-ordering (not defined here) have been studied by van Zwet (1964). Without the assumption of the common median zero, the definition 5.1-(4) has been used by Bickel and Lehmann (1979) to define an <u>ordering by spread</u> with the germinal concept attributed to Brown and Tukey (1946). Saunders and Moran (1978) have also perceived this kind of ordering (called <u>ordering by dispersion</u> by them) in the context of a neurobiological problem.

Gupta and Panchapakesan (1974) have defined a general partial ordering through a class of real-valued functions, which provides a unified way to handle selection problems for star-ordered and tail-ordered families. Their ordering is defined as follows.

Definition 5.2. Let $\mathcal{H} = \{h(x)\}$ be a class of real-valued functions $h(x)$. Let F and G be distributions such that $F(0) = G(0)$. F is said to be $\mathcal{H}$-ordered with respect to G ($F \underset{\mathcal{H}}{<} G$) if $G^{-1}F(h(x)) \geq h(G^{-1}F(x))$ for all $h \in \mathcal{H}$ and all x on the support of F.

It is easy to see that we get star-ordering and tail-ordering as special cases of $\mathcal{H}$-ordering by taking $\mathcal{H} = \{ax, a\geq 1\}$, $F(0) = G(0) = 0$, and $\mathcal{H} = \{x+b, b\geq 0\}$, $F(0) = G(0) = \frac{1}{2}$, respectively. Hooper and Santner (1979) have used a modified definition of $\mathcal{H}$-ordering. For some useful probability inequalities involving $\mathcal{H}$-ordering, see Gupta, Huang and Panchapakesan (1984).

5.1 Selection in terms of Quantiles from Star-Ordered Distributions.

Let $\pi_1,\ldots,\pi_k$ have the associated absolutely continuous distributions $F_1,\ldots,F_k$, respectively. All the F_i are star-shaped with respect to a known continuous distribution G. The population having the largest α-quantile ($0<\alpha<1$) is defined as the best population. It is assumed that the best population is stochastically larger than any of the other populations. Under this setup, Barlow and Gupta (1969) proposed a procedure for selecting a subset containing the best. Let $T_{j,i}$ denote the jth order statistic in a sample of n independent observations from π_i, $i = 1,\ldots,k$, where n is assumed to be large enough so that $j \leq (n+1)\alpha < j+1$ for some j. The Barlow-Gupta procedure is

R_{13}: Select π_i if and only if

$$T_{j,i} \geq c \max_{1\leq r\leq k} T_{j,r} \tag{5.1}$$

where $c = c(k,P^*,n,j)$ is the largest number in (0,1) for which the P^*-condition is satisfied. The constant c is given by

(5.2) $$\int_0^\infty G_j^{k-1}(x/c)\, g_j(x)\, dx = P^*$$

where G_j denotes the cdf of the jth order statistic in a sample of n observations from G, and g_j is the corresponding density function. The values of c satisfying (5.2) are tabulated by Barlow, Gupta and Panchapakesan (1969) in the special case of exponential G, i.e. for selecting from IFRA populations, for P* = 0.75, 0.90, 0.95, 0.99, and the following values of k, n, and j: (i) j = 1, k = 2(1)11 (in this case, c is independent of n), (ii) k = 2(1)6, j = 2(1)n, and n = 5(1)10 or 12 or 15 depending on k. Table 2A (at the end of this paper) is excerpted from the tables of Barlow, Gupta and Panchapakesan (1969). It gives the values of c for P* = 0.90, 0.95, k = 2(1)5, n = 5(1)12, and j such that $j \le (n+1)/2 < j+1$ (i.e. appropriate for selection in terms of median).

For the selection of the population with the smallest α-quantile (assumed to be stochastically smaller than any other F_i) the analogous procedure is

R_{14}: Select π_i if and only if

(5.3) $$dT_{j,i} \le \min_{1\le r\le k} T_{j,r}$$

where d = d(k,P*,n,j) is the largest number in (0,1) satisfying the P*-condition and is given by

(5.4) $$\int_0^\infty [1-G_j(xd)]^{k-1} g_j(x)\, dx = P^*$$

where G_j and g_j are defined as in (5.2). Barlow, Gupta and Panchapakesan (1969) have tabulated the values of d in the case of exponential G for $P^* = 0.75$, 0.90, 0.95, 0.99 and the following values of k, n, and j: (i) $j = 1$, $k = 2(1)11$ (d is independent of n), (ii) $k = 2(1)6$, $j = 2(1)n$, $n = 5(1)12$ for $k = 6$, and $n = 5(1)15$ for other k values. Table 2B (at the end of this paper) is excerpted from the tables of Barlow, Gupta and Panchapakesan (1969). It gives the values of d for $P^* = 0.90, 0.95$, $k = 2(1)5$, $n = 5(1)12$, and j such that $j \leq (n+1)/2 < j + 1$ (i.e. appropriate for selection in terms of median).

Suppose that G is the Weibull distribution with cdf $G(x) = 1 - \exp\{-(\frac{x}{\theta})^{\lambda}\}$, $x \geq 0$, and $\theta, \lambda > 0$. It is assumed that λ is known. Then it is easy to see that the new constant c_1 is given by $c_1 = c^{1/\lambda}$, where c is the constant in the exponential case (λ=1). Another interesting special case of G is the half-normal distribution obtained by folding $N(0,\sigma^2)$ at the origin, where σ is assumed to be known. The class of distributions which are star-shaped with respect to this folded normal is a subclass of IFRA distributions. Selection in terms of quantiles in this case has been considered by Gupta and Panchapakesan (1975), who have tabulated the constant c associated with R_{13} for $k = 2(1)10$, $n = 5(1)10$, $j = 1(1)n$, and $P^* = 0.75$, 0.90, 0.95, 0.99.

5.2 Selection in terms of Medians from Tail-ordered Distributions

Barlow and Gupta (1969) considered also the selection of the population with the largest median (assumed to be stochastically larger than other populations) from a set of distributions F_i, $i = 1,\ldots,k$, which have lighter tails than a specified distribution G with $G(0)=\frac{1}{2}$. This means that, for each i, F_i centered at its median Δ_i is r-ordered with respect to G, and $(d/dx)F_i(x+\Delta_i)|_{x=0} \geq (d/dx)G(x)|_{x=0}$. This definition of F_i having a lighter tail than G used by them implies that F_i centered at Δ_i is tail-ordered with respect to G. The procedure of Barlow and Gupta (1969) has been shown by Gupta and Panchapakesan (1974) to work for this wider class defined using tail-ordering. Actually, Gupta and Panchapakesan have also shown a generalized version of this by considering tail-ordering of F_i and G when both are centered at their respective α-quantiles.

For selection in terms of medians, the procedure of Barlow and Gupta is

R_{15}: Select π_i if and only if

$$T_{j,i} \geq \max_{1 \leq r \leq k} T_{j,r} - D, \qquad j \leq (n+1)/2 < j+1, \tag{5.5}$$

where the $T_{j,r}$ are defined as in the case of the procedure R_{13}, and the appropriate constant $D = D(k,P^*,n) > 0$ is given by

$$\int_{-\infty}^{\infty} G_j^{k-1}(t+D)g_j(t)dt = P^*. \tag{5.6}$$

Here, G_j and g_j are the cdf and the density of the jth order statistic in a sample of n independent observations from G. The values of D are given by Gupta and Panchapakesan (1974) in the special case where G is the logistic distribution, $G(x) = [1+e^{-x}]^{-1}$, for $k = 2(1)10$, $n = 5(2)15$, and $P^* = 0.75$, 0.90, 0.95, 0.99.

Using the $\mathcal{H}$-ordering (Definition 5.2) with the functions h satisfying certain properties, Gupta and Panchapakesan (1974) have discussed a class of procedures for selecting the best (i.e. the one which is stochastically larger than any other, assumed to exist) of k distributions F_i, $i,\ldots,k$, which are $\mathcal{H}$-ordered with respect to G. The procedures R_{13} and R_{15} are special cases of their procedure.

Hooper and Santner (1979) considered selection of good populations in terms of α-quantiles for star- and tail-ordered distributions using the RSS approach. Let π_i have the distribution F_i and let $F_{[i]}$ denote the distribution having the ith smallest α-quantile. Denoting the α-quantile of any distribution F by $x_\alpha(F)$, π_i is called a <u>good</u> population if $x_\alpha(F_i) > c^* x_\alpha(F_{[k-t+1]})$, $0<c^*<1$, in the case of star-ordered families, and if $x_\alpha(F_i) > x_\alpha(F_{[k-t+1]}) - d^*$, $d^* > 0$, in the case of tail-ordered families. The goal of Hooper and Santner (1979) is to select a subset of size not exceeding $m(1\le m<k-1)$ that contains at least one good population. They have also considered the problem of selecting a subset of fixed size s so as to include at least r good populations ($r\le t$, $r\le s<k-t+r$) using the IZ approach.

Selection of one or more good populations as a goal is a relaxation from that of selecting the best population(s). A good population is defined suitably to reflect the fact that it is 'nearly' as good as the best. In some form or other it has been considered by several authors; mention should be made of Fabian (1962), Lehman (1963), Desu (1970), Carroll, Gupta and Huang (1975), and Panchapakesan and Santner (1977). A discussion of this can be found in Gupta and Panchapakesan (1985, Section 4.2).

5.3 Selection from Convex Ordered Distributions

Let $\pi_1,\ldots,\pi_k$ have absolutely continuous distributions $F_1,\ldots,F_k$, respectively, of which one is assumed to be stochastically larger than the rest. This distribution, denoted by $F_{[k]}$, is defined to be the best. It is assumed that $F_{[k]} \underset{c}{\le} G$, where G is a known continuous distribution. All distributions in the context are assumed to have the positive real line as the support. Let $X_{j,n}^{(i)}$ $(Y_{j,n})$ denote the jth order statistic in a random sample of size n from $F_i(G)$. Considering samples of size n from $F_1,\ldots,F_k$ each censored at the rth failure, define

$$T_i = \sum_{j=1}^{r} a_j X_{j,n}^{(i)}, \qquad i = 1,\ldots,k, \tag{5.7}$$

where

$$\begin{cases} a_j = g\,G^{-1}(\frac{j-1}{n}) - g\,G^{-1}(\frac{j}{n}), & j = 1,\ldots,r-1, \\ a_r = g\,G^{-1}(\frac{r-1}{n}), \end{cases} \tag{5.8}$$

and g is the density associated with G.

If $G(y) = 1-e^{-y}$, $y \geq 0$, then $a_1 = \ldots = a_{r-1} = 1/n$, and $a_r = (n-r+1)/n$. Consequently, $n\,T_i = \sum_{j=1}^{r-1} X_{j,n}^{(i)} + (n-r+1)\,X_{r,n}^{(i)}$, the well-known <u>total life statistic until the rth failure from</u> F_i.

Now, for selecting a subset containing $F_{[k]}$, Gupta and Lu (1979) proposed the rule

R_{16}: Select π_i if and only if

$$T_i \geq c \max_{1 \leq j \leq k} T_j, \tag{5.9}$$

where c is the largest number in (0,1) satisfying the P*-condition. They have shown that, if $a_j \geq 0$ for $j = 1,\ldots,r$, $a_r \geq c$, and $g(0) \leq 1$, then

$$\inf_{\Omega} P(CS|R_{16}) = \int_0^{\infty} G_T^{k-1}(y/c)\, dG_T(y), \tag{5.10}$$

where G_T is the distribution of $T = \sum_{j=1}^{r} a_j Y_{j,n}$, and Ω is the space of all k-tuples $(F_1,\ldots,F_k)$ such that there is one among them which is stochastically larger than the others and is convex with respect to G. Thus, the constant $c = \min(a_r, c^*)$ where c^* is the solution for c by equating the right-hand side of (5.10) to P^*.

For the special case of $G(y) = 1-e^{-y}$, $y \geq 0$, we get $c = \min(c^*,(n-r+1)/n)$. This special case is a slight generalization of the results of Patel (1976).

6. Comparison with a Standard or Control

Although the experimenter is generally interested in selecting the best of $k(\geq 2)$ competing categories, in some situations even the best one among them may not be good enough to warrant its selection. Such a situation arises when the goodness of a population is defined in comparison with a standard (known) or a control population. For convenience, we may refer to either one as the control.

Let $\pi_1,\ldots,\pi_k$ be the k (experimental) populations with associated distribution functions $F(x,\theta_i)$, $i = 1,\ldots,k$, respectively. The θ_i are unknown. Let θ_0 be the specified standard or the unknown parameter associated with the control population π_0 whose distribution function is $F(x,\theta_0)$. Several different goals have been considered in the literature. For example, one may want to select the best experimental population (i.e. the one associated with $\theta_{[k]}$, the largest θ_i) provided that it is better than the control (i.e. $\theta_{[k]} > \theta_0$), and not to select any of them otherwise. An alternative goal is to select a subset (of random size) of the k populations which includes all those populations that are better than the control. Some of the early papers dealing with these problems are Paulson (1952), Dunnett (1955), and Gupta and Sobel (1958).

One can define a good population in different ways using comparison with a control. For example, π_i may be called good if $\theta_i > \theta_0+\Delta$, or $|\theta_i-\theta_0|\leq\Delta$ for some $\Delta>0$. Several procedures have been investigated with the goal of selecting good populations or those better than the control and these will not be described here. A good account of these can be had from Gupta and Panchapakesan (1979, Chapter 20). A review of subset selection procedures in this context, including recent developments, is contained in Gupta and Panchapakesan (1985).

An important aspect of the recent developments is the so-called isotonic procedures which become relevant in the situations where it is known that $\theta_1 \leq \theta_2 \leq \dots \leq \theta_k$ although the values of the θ_i are unknown. This is typical, for example, of experiments involving different dose levels of a drug so that the treatment effects will have a known ordering. Suppose that a population π_i is defined to be good if $\theta_i \geq \theta_0$ and bad otherwise. For the goal of selecting all the good populations, any reasonable procedure R should have the property: If R selects π_i then it selects all populations π_j for $j>i$. This is the isotonic behavior of R. Naturally, one would consider procedures based on isotonic estimators of the θ_i. Such procedures have been recently studied by Gupta and Yang (1984) in the case of normal means (common variance σ^2, known or unknown), by Gupta and Huang (1982) in the case of binomial populations with success probabilities θ_i, and by Gupta and Leu (1983) in the case of two-parameter exponential populations with guarantee times (location parameters) θ_i and common (known or unknown)scale parameter. All these papers deal with both cases of known and unknown θ_0.

7. Concluding Remarks

In the preceding sections, we have described several selection procedures that have special significance in reliability studies. However, we have confined our attention to the classical type procedures since they are of common interest to a wide variety of users. We have also generally restricted ourselves to single-stage procedures. There is ample literature on two-stage and sequential procedures. Further, we have not discussed decision-theoretic formulations and Bayes and empirical Bayes procedures. There have been substantial developments in these regards, especially using subset selection approach, in the last ten years. For a comprehensive survey of developments until the late 1970's, we refer to Gupta and Panchapakesan (1979). A critical review of developments in the subset selection theory including very recent developments is given by Gupta and Panchapakesan (1985).

Table 1A. Values of the constant c of Rule R_3 satisfying Equation (3.3)

$P^* = 0.90$

ν \ k	2	3	4	5	6	7	8	9	10	11
1	.111	.072	.059	.052	.047	.044	.041	.039	.038	.036
2	.244	.183	.159	.145	.135	.128	.123	.119	.116	.113
3	.327	.260	.232	.215	.203	.195	.188	.183	.178	.174
4	.386	.317	.286	.268	.255	.246	.239	.232	.228	.223
5	.430	.360	.329	.310	.297	.287	.279	.273	.268	.263
6	.466	.396	.364	.345	.332	.321	.313	.307	.301	.296
7	.494	.426	.394	.374	.361	.350	.342	.336	.330	.325
8	.519	.451	.419	.400	.386	.376	.367	.360	.355	.350
9	.539	.472	.441	.422	.408	.398	.389	.382	.376	.371
10	.558	.492	.460	.441	.428	.417	.409	.402	.396	.391
11	.573	.508	.478	.459	.445	.434	.426	.419	.414	.408
12	.588	.524	.493	.474	.461	.450	.442	.435	.429	.424
13	.600	.537	.507	.488	.475	.465	.456	.450	.444	.439
14	.612	.550	.520	.502	.488	.478	.470	.463	.457	.452
15	.622	.561	.532	.514	.500	.490	.482	.475	.469	.464
16	.632	.572	.543	.525	.511	.501	.493	.486	.481	.476
17	.641	.582	.553	.535	.522	.512	.504	.497	.491	.486
18	.649	.591	.562	.544	.532	.522	.514	.507	.501	.496
19	.657	.599	.571	.553	.540	.531	.523	.516	.510	.506
20	.664	.607	.579	.562	.549	.539	.531	.525	.519	.514

Table 1B. Values of the constant c of Rule R_3 satisfying Equation (3.3)

$P^* = 0.95$

ν \ k	2	3	4	5	6	7	8	9	10	11
1	.053	.035	.028	.025	.023	.021	.020	.019	.018	.018
2	.156	.119	.104	.095	.089	.085	.082	.079	.076	.074
3	.233	.188	.168	.156	.148	.142	.138	.134	.131	.128
4	.291	.242	.220	.206	.197	.190	.184	.180	.176	.173
5	.336	.285	.261	.247	.237	.229	.223	.218	.214	.210
6	.372	.320	.296	.281	.271	.263	.256	.251	.247	.243
7	.403	.350	.326	.310	.300	.291	.285	.279	.275	.271
8	.428	.376	.351	.336	.325	.316	.310	.304	.300	.296
9	.451	.399	.374	.358	.347	.339	.332	.326	.322	.317
10	.471	.419	.394	.378	.367	.359	.352	.346	.341	.337
11	.488	.437	.412	.396	.385	.377	.370	.364	.359	.355
12	.504	.453	.428	.413	.402	.393	.386	.380	.376	.371
13	.518	.468	.443	.428	.417	.408	.401	.395	.390	.386
14	.531	.481	.457	.442	.430	.422	.415	.409	.404	.400
15	.543	.494	.470	.454	.443	.434	.428	.422	.417	.413
16	.554	.505	.481	.466	.455	.446	.439	.434	.429	.424
17	.564	.516	.492	.477	.466	.457	.450	.445	.440	.436
18	.574	.526	.502	.487	.476	.468	.461	.455	.450	.446
19	.582	.535	.512	.497	.486	.477	.470	.465	.460	.456
20	.591	.544	.520	.506	.495	.486	.480	.474	.469	.465

Table 2A. Values of the constant c of Rule R_{13} satisfying Equation (4.2) for selecting the IFRA distribution with the largest median $G(x)=1-e^{-x}$, $x\geq 0$, $j\leq(n+1)/2<j+1$, $P^*=0.90$ (top entry), 0.95 (bottom entry)

n \ k	2	3	4	5
5	.32197 .22871	.25464 .18353	.22607 .16388	.20924 .15215
6	.32397 .23045	.25665 .18521	.22808 .16551	.21123 .15377
7	.38021 .28527	.31045 .23611	.27994 .21406	.26164 .20068
8	.38198 .28692	.31228 .23774	.28179 .21568	.26351 .20229
9	.42434 .32973	.35398 .27855	.32257 .25515	.30353 .24079
10	.42587 .33121	.35559 .28005	.32422 .25665	.30519 .24228
11	.45939 .36592	.38927 .31377	.35750 .28958	.33808 .27461
12	.46071 .36724	.39069 .31512	.35896 .29094	.33956 .27597

Table 2B. Values of the constant d of Rule R_{14} satisfying Equation (5.4) for selecting the IFRA distribution with the smallest median $G(x)=1-e^{-x}$, $x\geq 0$, $j\leq(n+1)/2<j+1$, $P^*=0.90$ (top entry), 0.95 (bottom entry)

n \ k	2	3	4	5
5	.32197 .22871	.23711 .17100	.19983 .14516	.17752 .12953
6	.32397 .23045	.23881 .17244	.20134 .14643	.17891 .13060
7	.38021 .28527	.29477 .22441	.25597 .19623	.23226 .17883
8	.38198 .28692	.29636 .22585	.25744 .19755	.23365 .18007
9	.42434 .32972	.33988 .26775	.30072 .23845	.27650 .22014
10	.42587 .33121	.34131 .26909	.30208 .23971	.27779 .22134
11	.45939 .36592	.37647 .30378	.33748 .27399	.31315 .25521
12	.46071 .36724	.37775 .30501	.33871 .27516	.31433 .25634

References

Bain, L. (1978). Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods. Marcel Dekker, New York.

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

Barlow, R. E., Gupta, S. S. and Panchapakesan, S. (1969). On the distribution of the maximum and minimum of ratios of order statistics. Ann. Math. Statist. 40, 918-934.

Barlow, R. E., Marshall, A. W. and Proschan, F. (1963). Properties of probability distributions with monotone hazard rate. Ann. Math. Statist. 34, 375-389.

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., 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 (with special reference to Koopman-Darmois populations). The University of Chicago Press, Chicago.

Bechhofer, R. E. and Kulkarni, R. V. (1982). Closed adaptive sequential procedures for selecting the best of $k \geq 2$ Bernoulli populations. In: S. S. Gupta and J. O. Berger, eds., Statistical Decision Theory and Related Topics - III, Vol. 1, pp. 61-108, Academic Press, New York.

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

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

Bickel, P. J. and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models IV. Spread. In: Jana Jureckova, ed., Contributions to Statistics: Jaroslav Hajek Memorial Volume, pp. 3-40, D. Reidel Publishing Co., Boston.

Birnbaum, Z. W., Esary, J. D. and Marshall, A. W. (1966). A stochastic characterization of wear-out for components and systems. Ann. Math. Statist. 37, 816-825.

Brown, G. and Tukey, J. W. (1946). Some distributions of sample means. Ann. Math. Statist. 7, 1-12.

Büringer, H., Martin, H. and Schriever, K.-H. (1980). Nonparametric Sequential Selection Procedures. Birkhauser, Boston, Massachusetts.

Carroll, R. J., Gupta, S. S. and Huang, D.-Y. (1975). On selection procedures for the t best populations and some related problems. Comm. Statist. 4, 987-1008.

Desu, M. M. (1970). A selection problem. Ann. Math. Statist. 41, 1596-1603.

Desu, M. M., Narula, S. C. and Villarreal, B. (1977). A two-stage procedure for selecting the best of k exponential distributions. Comm. Statist. A - Theory Methods 6, 1223-1230.

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.

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

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.

Fabian, V. (1962). On multiple decision methods for ranking population means. Ann. Math. Statist. 33, 248-254.

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

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill, North Carolina.

Gupta, S. S. (1963a). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math. 14, 199-216.

Gupta, S. S. (1963b). 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 Huang, D.-Y. (1980). A note on optimal subset selection procedures. Ann. Statist. 8, 1164-1167.

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 Huang, W.-T. (1976). On ranking and selection procedures and tests of homogeneity for binomial populations. In: S. Ikeda, T. Hayakawa, H. Hudimoto, M. Okamoto, M. Siotani and S. Yamamoto, eds., Essays in Probability and Statistics, Shinko Tsusho Co. Ltd., Tokyo, Japan, Chapter 33, pp. 501-533.

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

Gupta, S. S., Huang, D.-Y. and Panchapakesan, S. (1984). On some inequalities and monotonicity results in selection and ranking theory. In: Y. L. Tong, ed. Inequalities in Statistics and Probability, IMS Lecture Notes--Monograph Series, Vol. 5, pp. 211-227.

Gupta, S. S. and Huang, W.-T. (1982). On isotonic selection rules for binomial populations better than a standard. Technical Report 82-22, Department of Statistics, Purdue University. To appear in the Proceedings of the First Saudi Symposium on Statistics and Its Applications, held in May 1983.

Gupta, S. S. and Leu, L.-Y. (1983). 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 Lu, M.-W. (1979). Subset selection procedures for restricted families of probability distributions. Ann. Inst. Statist. Math. 31, 235-252.

Gupta, S. S. and McDonald, G. C. (1982). Nonparametric procedures in multiple decisions (ranking and selection procedures). In: B. V. Gnedenko, M. L. Puri and I. Vincze, eds., Colloquia Mathematica Societatis Janos Bolyai, 32: Nonparametric Statistical Inference, Vol. I, pp. 361-389, North-Holland, Amsterdam.

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. (1974). Inference for restricted families: (a) multiple decision procedures; (b) order statistics inequalities. In: F. Proschan and R. J. Serfling, eds., Reliability and Biometry: Statistical Analysis of Lifelength, pp. 503-596, SIAM, Philadelphia.

Gupta, S. S. and Panchapakesan, S. (1975). On a quantile selection procedure and associated distribution of ratios of order statistics from a restricted family of probability distributions. In: R. E. Barlow, J. B. Fussell and N. D. Singpurwalla, eds., Reliability and Fault Tree Analysis: Theoretical and Applied Aspects of System Reliability and Safety Assessment, pp. 557-576, SIAM, Philadelphia.

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. and Panchapakesan, S. (1985). Subset selection procedures: review and assessment. Amer. J. Management Math. Sci.

Gupta, S. S. and Santner, T. J. (1973). On selection and ranking procedures - a restricted subset selection rule. Proceedings of the 39th Session of the International Statistical Institute, Vol. 45, Book I, pp. 478-486.

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 Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. In: I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow and H. B. Mann, eds., Contributions to Probability and Statistics, Chapter 20, pp. 224-248, Stanford University Press, Stanford, California.

Gupta, S. S. and Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. Biometrika 49, 495-507.

Gupta, S. S. and Sobel, M. (1962b). On the smallest of several correlated F-statistics. Biometrika 49, 509-523.

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

Hoel, D. G., Sobel, M. and Weiss, G. H. (1975). A survey of adaptive sampling for clinical trials. In: R. M. Elashoff, ed., Perspectives in Biometry, pp. 29-61, Academic Press, New York.

Hooper, J. H. and Santner, T. J. (1979). Design of experiments for selection from ordered families of distributions. Ann. Statist. 7, 615-643.

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

Kim, W.-C. and Lee, S.-H. (1985). An elimination type two-stage selection procedure for exponential distributions. Comm. Statist. - Theor. Meth.

Kingston, J. V. and Patel, J. K. (1980a). Selecting the best one of several Weibull populations. Comm. Statist. A - Theory Methods 9, 383-398.

Kingston, J. V. and Patel, J. K. (1980b). A restricted subset selection procedure for Weiball distributions. Comm. Statist. A - Theory Methods 9, 1371-1383.

Lawrence, M. J. (1975). Inequalities for s-ordered distributions. Ann. Statist. 3, 413-428.

Lehmann, E. L. (1963). A class of selection procedures based on ranks. Math. Annalen 150, 268-275.

Milton, R. C. (1963). Tables of equally correlated multivariate normal probability integral. Technical Report No. 27, Department of Statistics, University of Minnesota, Minneapolis.

Nagel, K. (1970). On Subset Selection Rules with Certain Optimality Properties. Ph.D. Thesis (Also Mimeograph Series No. 222), Department of Statistics, Purdue University, West Lafayette, Indiana.

Panchapakesan, S. and Santner, T. J. (1977). Subset selection procedures for Δ_p-superior populations. Comm. Statist. A - Theory Methods 6, 1081-1090.

Patel, J. K. (1976). Ranking and selection of IFR populations based on means. J. Amer. Statist. Assoc. 71, 143-146.

Paulson, E. (1952). On the comparison of several experimental categories with a control. Ann. Math. Statist. 23, 239-246.

Raghavachari, M. and Starr, N. (1970). Selection problems for some terminal distributions. Metron 28, 185-197.

Rizvi, M. H. and Sobel, M. (1965). Nonparametric procedures for selecting a subset containing the population with the largest α-quantile. Ann. Math. Statist. 38, 1788-1803.

Robbins, H. (1952). Some aspects of the sequential design of experiments. Bull. Amer. Math. Soc. 58, 527-535.

Robbins, H. (1956). A sequential design problem with a finite memory. Proc. Nat. Acad. Sci. U.S.A. 42, 920-923.

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

Saunders, I. W. and Moran, P. A. P. (1978). On the quantiles of the gamma and F distributions. J. Appl. Prob. 15, 426-432.

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

Sobel, M. and Huyett, M. J. (1957). Selecting the best one of several binomial populations. Bell System Tech. J. 36, 537-576.

van Zwet, W. R. (1964). Convex Transformations of Random Variables. Mathematical Center, Amsterdam.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report #85-16	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Selection and Ranking Procedures in Reliability Models		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Technical Report #85-16
7. AUTHOR(s) Shanti S. Gupta and S. Panchapakesan		8. CONTRACT OR GRANT NUMBER(s) N00014-84-C-0167
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK, AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE July 1985
		13. NUMBER OF PAGES 44
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)

Indifference zone, subset selection, restricted subset, gamma, exponential, Weibull, binomial, nonparametric and distribution-free procedures, restricted families, IFR and IFRA families, convex and tail ordering, quantile selection, total life statistic, comparison with control

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

Several procedures appropriate for selecting the best from $k(\geq 2)$ life length distributions are discussed. The selection problem is considered (Section 2) under the classical formulation using both the indifference zone and subset selection approaches but mostly the latter. Under parametric models (Section 3), procedures are described for selection from gamma, Weibull, and exponential (one- and two-parameter) distributions. Nonparametric procedures and distribution-free procedures are discussed (Section 4) for selection in terms of a specified quantile and reliability at an arbitrarily chosen point in time, respectively.

Procedures for selection from restricted families are discussed (Section 6) with special reference to IFR and IFRA families. The procedures described in various cases mentioned above illustrate several modifications and generalizations of the basic goal.