324

# Some Multiple Decision and Related Problems with Special Reference to Restricted Families of Distributions and Applications to Reliability Theory

by

Shanti S. Gupta and S. Panchapakesan

Purdue University

and

Southern Illinois University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 324

July 1973

---

This research was supported by the Office of Naval Research Contract N00014-67-A-0226-00014. 

# Some Multiple Decision and Related Problems with Special Reference to Restricted Families of Distributions and Applications to Reliability Theory

Shanti S. Gupta† and S. Panchapakesan††

Abstract. This paper deals with procedures for selecting a subset from k given populations so as to include the "best" with a specified guaranteed minimum probability. Some general results relating to subset selection and specific procedures for important classes of distributions are reviewed with special emphasis on restricted families of probability distributions. Such families are defined through partial order relations and are extensively considered in reliability theory. A selection problem for tail-ordered family of distributions is considered and tables are provided for constants needed to implement the procedure. A general partial order relation is defined through a class of real valued functions and a related selection problem is discussed. These results provide a unified view of earlier known results. The rest of the paper gives a brief survey of some important results pertaining to restricted families of distributions such as the star-ordered and convex-ordered distributions. These results relate to life test sampling plans, inequalities for linear combinations of order statistics,

†Department of Statistics, Purdue University, West Lafayette, Indiana 47907.
††Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901.
This research was supported by the Office of Naval Research under Contract N00014-67-A-0226-00014.

estimation of failure rate function and some tests of hypotheses.

1. Introduction. In practice, there often do arise situations in which the experimenter wishes to choose the best (or the t best) from a group of k populations, where the "bestness" of a population is based on ranking them according to the value of any characteristic of interest. For example, he may want to select from a group of life length distributions each with an increasing failure rate, the distribution which has the largest mean life or the largest quantile of a given order. In such situations, the classical tests of homogeneity are deficient in the sense that they have not been designed to answer several possible questions in which the experimenter may be really interested. The need for more meaningful formulations in order to answer these questions set the stage for the early investigations of multiple decision formulations of the problems which have now come to be commonly known as selection and ranking problems.

In the last two decades of investigations in this area by several researchers, two basic formulations, in general, have been adopted. The first one is the so-called indifference zone formulation due to Bechhofer [19] and the other is the subset selection approach of Gupta [44]. The basic problem can be briefly described as follows.

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k populations with associated distribution functions $F_{\theta_i}$, $i = 1,2,\ldots, k$, respectively, where the $\theta_i$ are unknown parameters whose values belong to $\Theta$, an interval on the real line. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered parameters. It is assumed that there is no prior information regarding the correct pairing of the ordered and the unordered $\theta_i$'s. To be precise, we define the population associated with $\theta_{[k]}$ as the <u>best population</u>. In many specific problems, it could be the population associated with $\theta_{[1]}$ depending on the physical meaning of the parameter and desirability of a large or small value of that parameter. The goal of the basic problem in the formulation of Bechhofer [19] is to choose one of the populations as the best. On the other hand, in the subset selection approach of Gupta [44], the objective is to select a subset of the given populations which will include the best population. Of course, one would want to do this with as small a subset size as possible. Further, it is assumed that in case of a tie, one of the populations associated with $\theta_{[k]}$ is considered to have been tagged as the best. Depending on the approach, any selection of the populations satisfying the relevant goal is defined as a <u>correct selection</u>. In both the approaches, it is required of any decision rule R to guarantee a specified minimum for $P(CS|R)$, the probability of making a correct selection (CS) using the rule R. To be more explicit, it is required that

$$P(CS|R) \geq P^* , \quad 1/k < P^* < 1, \tag{1.1}$$

where $P^*$ is specified by the experimenter. In the indifference zone approach, (1.1) is to be satisfied whenever $\delta(\theta_{[k]}, \theta_{[k-1]})$, the distance (suitably defined) between the best and the next best populations exceeds an amount $\Delta^*$ specified by the experimenter. On the other hand, the subset selection approach requires that (1.1) be satisfied whatever be the true configuration of the $\theta_i$. The condition (1.1) is usually referred to as the basic probability requirement or the P*-condition. The procedures, which are discussed in §§ 2 and 3 are under the subset selection formulation. We have confined ourselves here to a description of the problem in its basic simplicity. Even the pioneering paper of Bechhofer [19] considered different goals. Since the early investigations, several authors have considered various modifications of the goal and different sampling rules. Some efforts have also been made to combine both the indifference zone and subset selection approaches. Recent results in this direction have been obtained by Santner [89] and, Gupta and Santner [60]. A list of published literature under both approaches is given in the monograph by Bechhofer, Kiefer and Sobel [20]. The developments and significant results in the literature prior to 1972 under the subset selection formulation have been surveyed by Gupta and Panchapakesan [58].

In reliability theory, we are largely concerned with estimation and optimization of the probability of survival, mean life, or, more generally, life distribution of components or systems. Other problems of interest include those involving quantities such as the probability of proper functioning of the system at either a specified or an arbitrary time. The probability models and statistical techniques applicable in these problems form an important part of the mathematical theory of reliability. Reliability problems have a structure of their own as could be seen, for example, in the development of concepts like monotone failure rates. As such they have stimulated investigations in other associated statistical problems.

In the above context, selection and ranking procedures are relevant to reliability problems. Of course, many procedures developed for the location and scale parameter families are applicable. However, in reliability problems, we are interested in procedures applicable to large classes of distributions such as those having increasing failure rate. This necessitates developing new procedures to meet these various situations. The investigations of Barlow and Gupta [ 6] form the initial efforts in this direction; however, this area of research remains largely unexplored.

The main objective of this paper is two-fold: (i) to present some important subset selection procedures applicable to reliability problems and (ii) to review some important results concerning inequalities involving order statistics and, estimation and hypothesis testing based on order statistics relative to restricted families of probability distributions. While the former sets the format and techniques of subset selection, the latter presents some useful results of potential value for further investigations.

Section 2 discusses some general theory of subset selection procedures and applications of these procedures to problems concerning specific distributions. Mainly all these procedures are parametric in the sense that the underlying distributions are known but for the values of the parameter(s) involved. A short description of some distribution-free procedures is also given. We also discuss some related formulations of problems such as selecting the populations better than a standard.

The next section considers selection procedures relative to some restricted families of distributions. These are of direct interest in reliability because the families with monotone failure rate or monotone failure rate on the average are examples of such families of practical interest. In this section, some new procedures are discussed for tail-ordered distributions and relevant tables are given. A general partial ordering is defined on the space of probability distributions through a class of real valued functions and a selection procedure is discussed using this ordering. It will be seen that this gives a unified view of some of the known results.

Some life testing sampling plans for IFR (DFR) families are discussed in §4. In §5 some inequalities are given for linear combinations of order statistics from F and G where F is (i) starshaped, (ii) convex, with respect to G. Linear combinations of order statistics can be looked upon as weighted sums of spacings. Total life on test statistic is a special case of interest.

A natural problem of interest is to estimate the failure rate when we do not know the functional form of the underlying distribution but only know that it has a monotone failure rate. This as well as the window estimators for the generalized failure rate function have been reviewed in §6. A brief review of some tests for exponentiality against the alternative that the underlying distribution is a member of some restricted family of distributions is also given in this section.

Most of the results reviewed in this paper were obtained by several investigators in the last ten years; a good many of them in the last five or six years. Research in selection and ranking problems like research in statistical problems of reliability relative to restricted families of probability distributions has been growing very fast in recent years. As such it is not possible and certainly is not our aim to give an exhaustive coverage of important results in the two areas. Our objective is to focus the attention on certain important results and highlight the potential for multiple decision formulation for problems in reliability.

2. Subset selection procedures - parametric and nonparametric cases. In this section we follow the general setup described in §1. We assume that the family of distributions $\{F_\theta\}$, $\theta \in \Theta$, is stochastically increasing (SI) in $\theta$, i.e., for $\theta < \theta'$ in $\Theta$, $F_\theta$ and $F_{\theta'}$ are distinct and $F_\theta(x) \geq F_{\theta'}(x)$ for all x. We take $\Theta$ to be an open interval on the real line and assume, unless stated otherwise, that the distributions, $F_\theta$, $\theta \in \Theta$, are absolutely continuous with densities $f_\theta$ and have the same support denoted by $I(F_\theta)$. We also assume that $F_\theta$ is continuously differentiable in $\theta$. Our goal is to select a subset containing the population associated with $\theta_{[k]}$. For this purpose, we define a class of rules $R_h$ using a class $\mathcal{H} = \{h\}$ of continuous real valued functions defined on the real line. The class $\mathcal{H}$ is assumed to have the

following properties:

$$
(2.1)\quad
\begin{array}{ll}
\text{(i)} & h_o: x \to x \text{ is a member of } \mathscr{H}, \\
\text{(ii)} & h(x) \geq x \text{ for every } h \in \mathscr{H} \text{ and } x \in I(F_\theta), \\
\text{(iii)} & \text{for every } x \in I(F_\theta), \text{ and } h \in \mathscr{H}, \text{ there exists} \\
& \text{a sequence } \{h_m\} \text{ in } \mathscr{H} \text{ such that } \lim_{m\to\infty} h_m(x) = h(x), \\
\text{(iv)} & \text{for every } x \in I(F_\theta), \text{ except perhaps on a set of} \\
& \text{Lebesque measure zero, there is a sequence } \{h_m\} \text{ in} \\
& \mathscr{H} \text{ such that } \lim_{m\to\infty} h_m(x) = \infty.
\end{array}
$$

Gupta and Panchapakesan [57] have defined the class of rules, $R_h$: Include the population $\pi_i$ if and only if

$$(2.2)\qquad h(x_i) \geq \max_{1\leq r\leq k} x_r,$$

where $x_1,\ldots, x_k$ is a set of observations from $\pi_1,\ldots, \pi_k$, respectively. Because of the stochastic ordering of $\{F_\theta\}$,

$$(2.3)\qquad \inf_{\Omega} P(CS|R_h) = \inf_{\theta\varepsilon\Theta} \int F_\theta^{k-1}(h(x))\, dF_\theta(x),$$

where $\Omega$ is the parametric space $\{\underline{\theta}: \underline{\theta} = (\theta_1,\ldots, \theta_k); \theta_i \varepsilon \Theta, i = 1,\ldots, k\}$ and the integral is over the support of $F_\theta$. Let

$$(2.4)\qquad A_h(\theta) = \int F_\theta^{k-1}(h(x))\, dF_\theta(x).$$

Regarding the evaluation of the infimum in (2.3), Gupta and Panchapakesan [57] have proved the following theorem.

<u>Theorem 2.1</u>. <u>Suppose, for every</u> $h \varepsilon \mathscr{H}$ <u>and</u> $\theta \varepsilon \Theta$,

$$(2.5)\qquad f_\theta(x) \frac{\partial}{\partial\theta} F_\theta(h(x)) - h'(x)\, f_\theta(h(x)) \frac{\partial}{\partial\theta} F_\theta(x) \geq 0$$

for all $x \in I(F_\theta)$, where $h'(x) = (d/dx)\, h(x)$. Then, under certain regularity conditions, $A_h(\theta)$ is nondecreasing in $\theta$. The monotonicity of $A_h(\theta)$ is strict, if (2.5) holds with strict inequality.

The above theorem is, in fact, a consequence of a more general theorem proved in [57].

If the condition (2.5) is satisfied, then $\inf_{\theta \in \Theta} A_h(\theta) = \lim_{\theta \to \theta_o^+} A_h(\theta) = A_h(\theta_o^+)$, where $\theta_o$ is the left hand endpoint of the open interval $\Theta$. With no loss of generality, we can assume that $F_{\theta_o^+}$ is a distribution function. Because of the properties stated in (2.1), it can be seen that there exists an $h \in \mathscr{H}$ for which

$$A_h(\theta_o^+) = P^* \; , \; 1/k < P^* < 1. \tag{2.6}$$

Remark. In any particular problem of interest we may have $\theta \in \Theta'$, where $\Theta'$ is the interval $\Theta$ closed at either end or at both ends. In this case, let

$$B(h) = \min[A_h(\theta_o^+), \inf_{\theta \in \Theta' \cap \partial\Theta} A_h(\theta)], \tag{2.7}$$

where $\partial\Theta$ is the boundary of $\Theta$. The properties in (2.1) are sufficient to assure the existence of an $h$ for which $B(h) = P^*$. However, in most practical situations, we have $A_h(\theta_o^+) = A_h(\theta_o)$.

Let $p_i$ denote the probability that the population associated with $\theta_{[i]}$ is included in the subset. Then, a procedure R designed to select the population with the largest $\theta$ is said to be monotone if $p_i \leq p_j$ for $1 \leq i < j \leq k$. If $p_i \leq p_k$ for all i, then we say that R is unbiased.

If $h(x)$ is increasing in x, it can be shown that $R_h$ is monotone.

Suppose we denote the number of populations included in the selected subset using a procedure R by S. Then S, called the subset size, is an integer valued random variable. We are interested in the $E(S) \equiv E(S|R_h)$. It is easy to see that $E(S) = p_1 + \ldots + p_k$. The following theorem relating to the supremum of E(S) over $\Omega$ has been proved in [57].

Theorem 2.2. *For the procedure* $R_h$ *defined by* (2.2), $\sup_{\Omega} E(S|R_h)$ *is attained when* $\theta_1 = \theta_2 = \ldots = \theta_k$ *provided that*

$$(2.8) \qquad \frac{\partial}{\partial\theta_1} F_{\theta_1}(h(x))\, f_{\theta_2}(x) - h'(x) \frac{\partial}{\partial\theta_1} F_{\theta_1}(x)\, f_{\theta_2}(h(x)) \geq 0$$

*for every* $h \in \mathscr{A}$, $\theta_1 \leq \theta_2$ *and all* $x \in I(F_\theta)$.

It should be noted that (2.8) implies (2.5). Many of the procedures investigated in the literature are members of the class $R_h$. In the subsequent parts of this section we discuss some specific procedures of interest.

2.1. Selection in terms of location and scale parameters. Two of the important cases in many statistical investigations are those of location and scale parameters. Many of the parameters of interest fall under one of these cases. We first discuss the location parameter case.

(a) Location parameter. In this case we have $F_{\theta_i}(x) = F(x-\theta_i)$, $\theta_i \in (-\infty,\infty)$ and $I(F_\theta) = (-\infty,\infty)$. The usual choice is $h(x) = x + b$, $b \geq 0$. It is easy to see that $A_h(\theta)$ is constant for $\theta \in \Theta$ and hence we can evaluate the inf $P(CS|R_h)$ at $\theta = 0$. It can also be shown [57] that, for the above choice of $h(x)$, the condition (2.8 ) reduces to monotone likelihood ratio (MLR) in x. Thus the constant b satisfying the $P^*$-condition is given by

$$\int_{-\infty}^{\infty} F^{k-1}(x+b)\, dF(x) = P^* \tag{2.9}$$

and $\sup E(S) = k\, P^*$, if $f(x-\theta)$ has MLR in x.

As an application to a specific problem, consider selection of a subset containing the largest mean from k independent normal populations with unknown means $\mu_1, \ldots, \mu_k$ and a common known variance $\sigma^2$. If $\bar{y}_i$, $i = 1, \ldots, k$, are the sample means based on n observations from each population, the rule $R_h$ selects $\pi_i$ if and only if $\bar{y}_i \geq \max\limits_{1 \leq r \leq k} \bar{y}_r - d\sigma/\sqrt{n}$, $d > 0$. Here $h(x) = x + d\sigma/\sqrt{n}$. Thus, the constant d is given by

$$\int_{-\infty}^{\infty} \Phi^{k-1}(u + d)\, \varphi(u)\, du = P^*, \tag{2.10}$$

where, unless otherwise stated, $\Phi(\cdot)$ and $\varphi(\cdot)$ denote the cdf and the density of the standard normal distribution. If $\sigma^2$ is unknown, one will naturally use $s^2$, the pooled estimator of $\sigma^2$ based on $k(n-1)$ degrees of freedom. In this case it can be shown that d is given by

$$\int_{0}^{\infty} \int_{-\infty}^{\infty} \Phi^{k-1}(u+yd)\, \varphi(u)\, g_\nu(y)\, du dy = P^*, \tag{2.11}$$

where $g_\nu(y)$ is the density of $\chi_\nu/\nu$ with $\nu = k(n-1)$. For the case of known $\sigma^2$, the constant d can be obtained for selected values of k, n and $P^*$ from the table of Gupta [46]. Tables for the case of unknown $\sigma^2$ are given by Gupta and Sobel [61]. These procedures have also been discussed by Gupta [48]. The selection problem with unequal sample sizes has been investigated by Gupta and Huang [51].

(b) Scale parameter case. In this case, we have $F_{\theta_i}(x) = F(x/\theta_i)$, $\theta_i \in (0,\infty)$, $I(F_\theta) = (0,\infty)$. For the choice of $h(x) = ax$, $a \geq 1$, it is easily seen that $A_h(\theta)$ is constant for $\theta \in \Theta$ and hence the inf $P(CS|R_h)$ can be evaluated at $\theta = 1$. Further, sup $E(S)$ is attained when $\theta_1 = \ldots = \theta_k$ if the density $f_\theta(x)$ has MLR in x and in that case sup $E(S) = k\,P^*$. The constant a satisfying the $P^*$-condition is given by

(2.12) $$\int_0^\infty F^{k-1}(ax)\, dF(x) = P^*.$$

A specific example of interest is the selection from k gamma populations with densities

(2.13) $$f_{\theta_i}(x) = \{\Gamma(r)\}^{-1}\, \theta_i^{-r}\, x^{r-1}\, e^{-x/\theta_i}, \quad x > 0,\ \theta_i > 0,\ i = 1,\ldots, k.$$

It is known [95] that the gamma distributions have increasing failure rates when $r > 1$. One may be interested in selecting the population with the largest $\theta_i$. This corresponds to selecting the population with the largest mean $\mu_i = r\,\theta_i$. For this goal, Gupta [47] proposed the rule $R_a$: Select $\pi_i$ if and only if

(2.14) $$a\,\bar{x}_i \geq \max_{1\leq r\leq k} \bar{x}_r, \qquad a \geq 1,$$

where $\bar{x}_i$, $i = 1,\ldots, k$, are the sample means based on n independent observations from each of the populations. The appropriate constant a is given by

(2.15) $$\int_0^\infty G_\nu^{k-1}(ax)\, dG_\nu(x) = P^*,$$

where $G_\nu(x)$ is the cdf of a standardized gamma variate (i.e. with $\theta=1$)

with parameter $\nu/2 = nr$. Gupta [47] has tabulated the values of $a^{-1}$ for $k = 2(1)11, \nu = 2(2)50$ and $P^* = .75, .90, .95$ and $.99$. These tables are also applicable for the case when the common $r$ is unknown. This is discussed in §3. The selection problem for the smallest $\theta_i$ can be treated in an analogous manner. This problem arises in the context of selecting a subset containing the normal population with the smallest variance. The rule in this case is an obvious modification of $R_a$. We select $\pi_i$ if and only if $\bar{x}_i \leq a \min_{1 \leq r \leq k} \bar{x}_r$. This rule has been studied by Gupta and Sobel [64] and the related tables are available in their companion paper [65].

2.2. Selection from multivariate normal populations. Selection problems for multivariate normal populations have been investigated using several measures of ranking. Let $\pi_1, \ldots, \pi_k$ be independent p-variate normal distributions, where $\pi_i$ has mean vector $\underline{\mu}_i$ and covariance matrix $\Sigma_i$, $i = 1, \ldots, k$. Let $\underline{x}_{ij}$, $j = 1, 2, \ldots, n$, be random vector observations from $\pi_i$. Let $S_i = (n-1)^{-1} \sum_{\ell=1}^{n} (\underline{x}_{i\ell} - \bar{\underline{x}}_i)(\underline{x}_{i\ell} - \bar{\underline{x}}_i)'$, denote the sample covariance matrix where $\bar{\underline{x}}_i$ is the sample mean vector.

(a) Selection in terms of generalized variance, $|\Sigma|$. In this case $\underline{\mu}_i$ and $\Sigma_i$ are unknown. The goal is to select a subset containing the population associated with the smallest $|\Sigma_i|$. For this problem, Gnanadesikan and Gupta [40] studied the rule R: Select the population $\pi_i$ if and only if

$$(2.16) \qquad |S_i| \leq c^{-1} \min_{1 \leq r \leq k} |S_r|, \qquad 0 < c \leq 1.$$

It has been shown in [40] that

$$\inf_{\Omega} P(CS|R) = \Pr(Y_1 \le c^{-1} Y_j,\ j = 2,\ldots, k), \tag{2.17}$$

where $Y_i$, $i = 1,\ldots, k$, are independent random variables, each being the product of p independent factors, the rth factor being distributed as a chi-square random variable with (n-r) degrees of freedom. The exact distribution of $Y_i$ is unknown except when $p = 2$.

(b) Selection in terms of distance functions. In some problems, it is of interest to use the Mahalanobis distance function, $\lambda_i = \underline{\mu}_i' \Sigma_i^{-1} \underline{\mu}_i$, to rank the populations. In the case of a univariate population, $1/\lambda_i$ is the square of the coefficient of variation. The Mahalanobis distance function seems to be a good measure in some infra-specific taxonomic problems. Our goal is to select a subset containing the population associated with the largest $\lambda_i$. Let $y_{ij} = \underline{x}_{ij}' \Sigma_i^{-1} \underline{x}_{ij}$, $j = 1,\ldots, n$; $i = 1,2,\ldots, k$. Then $y_i = \sum_{j=1}^{n} y_{ij}$ has the non-central chi-square distribution with np degrees of freedom and non-centrality parameter $\lambda_i' = n\lambda_i$. For the case of $\Sigma_i = \Sigma$(known) for all i, Gupta [49] proposed the rule $R_1$: Select the population $\pi_i$ if and only if

$$y_i \ge c \max_{1\le r\le k} y_r, \qquad 0 < c < 1. \tag{2.18}$$

Because of the stochastic ordering of the non-central chi-square distributions in $\lambda'$, the inf $P(CS|R_1)$ is attained when the distributions have the same non-centrality parameter $\lambda'$. Thus

$$\inf_{\Omega} P(CS|R_1) = \inf_{\lambda'\ge 0} \int_0^{\infty} F_{\lambda'}^{k-1}(x/c)\, dF_{\lambda'}(x), \tag{2.19}$$

where $F_{\lambda'}(x)$ is the cdf of a non-central chi-square variable with np degrees of freedom and non-centrality parameter $\lambda'$. Gupta [49] showed that, for k = 2, the integral on the right hand side of (2.19) is non-decreasing in $\lambda' \geq 0$ and hence the infimum takes place at $\lambda' = 0$. Later, Gupta and Studden [60] established the monotonicity for $k \geq 2$. They proved the following theorem.

Theorem 2.3. *Let $g_j(x)$, $j = 0,1,2,\ldots$ be a sequence of density functions on the interval $[0,\infty)$ and define*

$$f_\lambda(x) = \sum_{j=0}^{\infty} (e^{-\lambda}\lambda^j/j!)\; g_j(x), \qquad x \geq 0. \tag{2.20}$$

*For a fixed integer $k \geq 2$ and $a > 1$, let*

$$I(\lambda) = \int_0^\infty F_\lambda^{k-1}(ax)\; dF_\lambda(x) \quad \text{and} \tag{2.21}$$

$$J(\lambda) = \int_0^\infty [1-F_\lambda(x/a)]^{k-1}\; dF_\lambda(x). \tag{2.22}$$

*Then, $I(\lambda)$ and $J(\lambda)$ are nondecreasing in $\lambda$ provided that, for each $\ell \geq 0$,*

$$\sum_{i=0}^{\ell} \binom{\ell}{i} [\{G_{i+1}(ax)-G_i(ax)\}\; g_{\ell-i}(x) - c\; g_i(cx)\; \{G_{\ell-i+1}(x) - G_{\ell-i}(x)\}] \geq 0. \tag{2.23}$$

*Further, the monotonicity is strict if the condition (2.23) holds with strict inequality for some integer $\ell$.*

It should be pointed out that the monotonicity of $J(\lambda)$ is needed for the procedure defined by Gupta and Studden [60] for selecting the population associated with the smallest $\lambda_i$. Gupta and Studden [60] considered the case of $\Sigma_i$ known but not necessarily equal. Thus, for their procedure, $y_{ij} = \underline{x}_{ij}' \Sigma_i^{-1} \underline{x}_{ij}$ and the procedure is substantially the same as $R_1$ defined by (2.18). They have also considered a procedure for the case where the $\Sigma_i$ are different but all unknown. In this case, let $Z_i = \bar{\underline{x}}_i' S_i^{-1} \bar{\underline{x}}_i$. Then, the procedure studied is $R_2$: Select $\pi_i$ if and only if

$$(2.24) \qquad c' Z_i \geq \max_{1 \leq r \leq k} Z_r, \quad c' \geq 1.$$

In this case, it is shown that inf $P(CS|R_2)$ is attained when $\lambda_1' = \ldots = \lambda_k' = 0$. The distribution of $Z_i$, when the $\lambda_i$ are equal, is the non-central F with degrees of freedom np and noncentrality parameter $\lambda'$. The applicability of Theorem 2.3 in these two cases is due to the fact that the non-central chi-square and the non-central F distributions are of the form (2.20) where $g_j(x)$ are central chi-square and F distributions, respectively, with degrees of freedom depending on j. It should be pointed out that, when $\Sigma_i$ are known, a procedure analogous to $R_2$ with $Z_i = \bar{\underline{x}}_i' \Sigma_i^{-1} \bar{\underline{x}}_i$ is undesirable because the constant c in this case does not depend on n. This difficulty is overcome by $R_1$.

(c) Selection in terms of multiple correlation coefficient. Let $\rho_i \equiv \rho_{1 \cdot 2 \ldots p}$ be the multiple correlation coefficient between the first variable and the rest in the population $\pi_i$. Gupta and Panchapakesan [56] investigated procedures for selecting a subset containing the population associated with the largest $\rho_i$. Let the corresponding sample multiple correlation coefficients be $R_i$, i = 1,..., k. One of the procedures investigated by Gupta and Panchapakesan [56] is $R_3$: Select $\pi_i$ if and only if

$$(2.25)\qquad R^{*2} \geq c_1 \max_{1\leq r\leq k} R_r^{*2}, \qquad 0 < c_1 < 1,$$

where $R^{*2} = R_i^2/(1-R_i^2)$, $i = 1,\ldots, k$. Letting $\lambda_i = \rho_i^2$, the distribution of $R_i^{*2}$ is given by

$$(2.26)\qquad u_\lambda(x) = \sum_{j=0}^{\infty} \frac{\Gamma(q+m+j)\lambda^j}{\Gamma(q+m)j!} (1-\lambda)^{q+m} f_{2(q+j),2m}(x)$$

in the so-called unconditional case and by

$$(2.27)\qquad u_\lambda(x) = \sum_{j=0}^{\infty} \frac{e^{-m\lambda}(m\lambda)^j}{j!} f_{2(q+j),2m}(x)$$

in the conditional case, where $q = (p-1)/2$, $m = (n-p)/2$ and $f_{r,s}(x)$ denotes the density of the F-distribution with r and s degrees of freedom. In this case,

$$(2.28)\qquad \inf_\Omega P(CS|R_3) = \inf_\lambda \int_0^\infty U_\lambda^{k-1}(x/c)\, dU_\lambda(x),$$

where $U_\lambda(x)$ is the cdf corresponding to $u_\lambda(x)$. Gupta and Panchapakesan [56] have obtained a theorem analogous to Theorem 2.3 for the unconditional case, where the $g_j$ are weighted with negative binomial weight functions. The approach is similar to that of Theorem 2.3 which appeared first in 1965 in a technical report. However, in view of the remarks on mixtures of distributions made below, we omit the statement of the theorem of Gupta and Panchapakesan [56].

We first note that, for all procedures $R_1$ through $R_3$ discussed above, we are interested in the monotonicity of the integral

$\int F_\lambda^{k-1}(ax)\, d\, F_\lambda(x)$, $a \geq 1$, where $F_\lambda(x)$ is a mixture of a sequence of distributions $G_j(x)$ with either Poisson weights or negative binomial weights. This is the motivation for the theorem obtained by Gupta and Panchapakesan [57] for their class of procedures $R_h$ discussed earlier in this section. Suppose the distribution $F_\theta$ is of the form

$$(2.29) \qquad F_\theta(x) = \sum_{j=0}^{\infty} w(\theta,j)\, G_j(x),$$

where $G_j(x)$, $j = 1,2,\ldots$, is a sequence of distribution functions and $w(\theta,j)$ are non-negative weights such that $\sum_{j=0}^{\infty} w(\theta,j) = 1$ for $\theta \in [0,\infty)$.

We assume that the weights are given by

$$(2.30) \qquad \begin{aligned} w(\theta,j) &= a_j \theta^j / B(\theta) j!, \quad B(\theta) \geq 0,\ \theta \geq 0, \\ a_{j+1} &= (m+\ell j)aj, \qquad j = 0,1,\ldots;\ \ell,m \geq 0. \end{aligned}$$

It is easy to see that $B(\theta) = a_o(1-\theta\ell)^{-m/\ell}$, provided $\theta\ell < 1$. The following theorem is due to Gupta and Panchapakesan [57].

Theorem 2.4. *The condition* (2.8) *is satisfied if, for* $\alpha = 0,1,\ldots,$ $[i/2]$ *and* $b \geq 1$,

$$(2.31) \qquad \begin{aligned} & b^{i-\alpha}(m+\ell\alpha)[g_{i-\alpha}(x)\, \Delta G_\alpha(h(x)) - h'(x) g_{i-\alpha}(h(x))\, \Delta G_\alpha(x)] \\ & + b^\alpha(m+\ell(i-\alpha))[g_\alpha(x)\, \Delta G_{i-\alpha}(h(x)) - h'(x) g_\alpha(h(x))\, \Delta G_{i-\alpha}(x)] \geq 0, \end{aligned}$$

*where* $\Delta G_\alpha(x) = G_{\alpha+1}(x) - G_\alpha(x)$ *and* $[s]$ *denotes the largest integer* $\leq s$.

As an immediate consequence of Theorems 2.1, 2.2, 2.4 and the remark immediately following Theorem 2.2, we see that, if (2.31) holds for every $h \in \mathscr{H}$, then $A_h(\theta)$ defined by (2.4) is nondecreasing in $\theta$ and sup E(S) is attained when $\theta_1 = \ldots = \theta_k$. If we set $m = 1$, $\ell = 0$, and

$a_o = 1$ in (2.30) we get Poisson weights. The negative binomial weights are obtained by letting $\ell = 1$ and $a_o = 1$.

The values of the constant c defining $R_1$ can be obtained from the tables of Gupta [49]. The constants c' of $R_2$ and $c_1$ of $R_3$ can be obtained from the tables of Gupta and Panchapakesan [56]. It should be pointed out that, though we have described only procedures for selecting the population associated with the largest value of the parametric function chosen, the several papers referred to above have also considered the case of the population associated with the smallest value of the parametric function.

2.3. Selection in terms of discrete distributions. We have discussed so far in this section procedures when the underlying distributions $F_\theta$ are continuous possessing densities. We now consider the case where the underlying distributions are discrete. Subset selection procedures have been investigated in the case of distributions of importance such as binomial, negative binomial and Poisson. However, for the purpose of illustration, we are primarily interested in the binomial case.

Suppose $\pi_1, \ldots, \pi_k$ are k independent binomial populations with unknown probabilities of success on a single trial $\theta_1, \ldots, \theta_k$, respectively, where $0 \leq \theta_i \leq 1$, $i = 1, \ldots, k$. The goal is to select a subset containing the population with the largest $\theta_i$. Towards this end Gupta and Sobel [63] proposed the rule R: Select $\pi_i$ if and only if

(2.32)
$$x_i \geq \max_{1\leq r\leq k} - d,$$

where $x_i$ is the observed number of successes in n observations from $\pi_i$ and $d = d(n,k,P^*)$ is the smallest nonnegative integer that will satisfy the P*-condition. It is known that $P(CS|R)$ is minimized when $\theta_1=\ldots=\theta_k$. Thus, the constant d is the smallest nonnegative integer for which

(2.33)
$$\inf_{0\leq\theta\leq 1} \sum_{\alpha=0}^{n} \binom{n}{\alpha}\theta^{\alpha}(1-\theta)^{n-\alpha}\left[\sum_{j=0}^{\alpha+d}\binom{n}{j}\theta^{j}(1-\theta)^{n-j}\right]^{k-1} \geq P^*.$$

It has been shown by Gupta and Sobel [63] that, for $k = 2$, the infimum in (2.33) is attained for $\theta = 1/2$, and that, for a fixed k, the value $\theta_o$ at which the infimum takes place tends to 1/2 as $n\to\infty$. However, in general, the value of $\theta$ for which the infimum takes place is not known. Gupta and Nagel [54] have proposed a randomized rule R' as an alterva- tive. Let $p_k(x_1,\ldots, x_k)$ denote the probability of selecting $\pi_k$ based on the observations $x_1,\ldots, x_k$. Then the rule R' is defined by

(2.34)
$$p_k(x_1,\ldots, x_k) = \begin{array}{lll} 1 & \text{if} & x_k > C_T, \\ \rho & \text{if} & x_k = C_T, \\ 0 & \text{if} & x_k < C_T, \end{array}$$

where $T = \sum_{i=1}^{k} x_i$ is a sufficient statistic for $\theta$ and, $\rho=\rho(T,P^*,k,n)$ and $C_T(P^*,k,n)$ are determined so as to satisfy the P*-condition. The values of $C_T$ and $\rho$ have been tabulated by Gupta and Nagel [54] for $k = 2,3,5$; $n = 5$, 10 and $P^* = 0.75$, 0.90, 0.95, 0.99 with T in each case going from 0 to nk.

As an application, consider independent continuous distributions

$F_i$ with densities $f_i$, $i = 1,\ldots, k$, associated with lengths of life from k populations. Then $r_i(t) = f_i(t)/[1-F_i(t)]$ is the failure rate function for the population $\pi_i$. We assume that these populations are IFRA, i.e., $R_i(t) = t^{-1} \int_0^t r_i(x)dx$ is increasing in t. Such distributions and several results relating to them are discussed at length in the subsequent sections of this paper. Here we will assume that there is one among the k populations, denoted by $F_{[k]}(x)$, such that the associated $R_{[k]}(t) \leq R_i(t)$, for all $t > 0$, $i=1,\ldots, k$. The goal is to select a subset containing that population. It is easy to see that, if $R_i(t) \leq R_j(t)$, then $F_i(t) \leq F_j(t)$. Thus, the best population here is the one which is stochastically larger than any other population. Suppose we put n items from each population on a life test for a period of time, T. Let $x_1,\ldots, x_k$ be the number of failures. Then, $x_i$ has the binomial distribution with $p_i = F_i(T)$ and we are interested in the population with the smallest $p_i$. So we can use the procedures R and R' with $n-x_i$ in the place of $x_i$.

Gupta and Nagel [54] have studied procedures similar to R' defined by (2.34) for the problem of selection from Poisson and negative binomial distributions. The case of Fisher's logarithmic distributions has been discussed by Nagel [79]. The subset selection rules for selecting the cell with the largest (smallest) probabilities in a multinomial distribution have been investigated by Gupta and Nagel [53], and Panchapakesan [80].

2.4. Distribution-free procedures based on ranks. We now assume our earlier setup where $\{F_\theta\}$ is a stochastically increasing family of absolutely continuous distributions. However, it is now further assumed that the functional form of $F_\theta$ is not known. For selecting a subset containing the population associated with the largest $\theta_i$, Gupta and McDonald [52] investigated three classes of rules based on ranks. Let $x_{ij}$, $j = 1,\ldots, n_i$, be independent observations from $\pi_i$, $i = 1,\ldots,$ $k$. All the $N = n_1 +\ldots+ n_k$ observations are pooled and ordered. Let $Z(1) \leq Z(2) \leq\ldots\leq Z(N)$ denote an ordered sample of size $N$ from a continuous distribution $G$ such that $-\infty < a(r) \equiv E_G(Z(r)) < \infty$, $r = 1,\ldots, N$. With each observations $x_{ij}$ associate the number $a(R_{ij})$ and define

$$H_i = n_i^{-1} \sum_{i=1}^{n_i} a(R_{ij}), \qquad i = 1,\ldots, k, \tag{2.35}$$

where $R_{ij}$ denotes the rank of $X_{ij}$ in the combined sample. The classes of procedures studied by Gupta and McDonald are based on the $H_i$. The difficulty with the usual types of procedures is that the infimum of the probability of a correct selection does not necessarily take place when all the populations are identical unless $k = 2$. A detailed survey of these and other related procedures is given by Gupta and Panchapakesan [58].

<u>2.5. Selecting a subset better than a standard</u>. In many practical situations we may be interested in choosing from k given populations those which compare favorably with a standard or a control population in terms of the characteristic of interest. To be precise, let $\pi_i$, $i = 0,1,\ldots, k$, be $(k+1)$ populations with the associated distribution

functions $F_{\theta_i}$. The parameter $\theta_o$ of the standard population, $\pi_o$, may or may not be known. The other $\theta_i$ are unknown. We say that $\pi_i$ is <u>better than</u> $\pi_o$ if $\theta_i \geq \theta_o$ (or $\theta_i \leq \theta_o$). The goal is to select a subset of the k experimental populations with a minimum guarantee on the probability that all populations better than the standard are included in the selected subset. The normal means problem was investigated by Gupta and Sobel [62]. Later Gupta [48] has discussed the cases of location and scale parameters in general. Nonparametric selection procedures when the comparison is in terms of $\alpha$-quantile ($0 < \alpha < 1$) have been studied by Rizvi, Sobel and Woodworth [88].

In problems of comparison of experimental populations with a standard, mention should be made of the formulation of Lehmann [73]. He considers a population to be good if it is sufficiently better than the standard. In other words, $\pi_i$ is <u>positive</u> (or good) if $\theta_i \geq \theta_o + \Delta$ and <u>negative</u> (or bad) if $\theta_i \leq \theta_o$, where $\Delta$ is a given positive constant. Let $S(\theta,\delta)$ and $R(\theta,\delta)$ denote the expected number of <u>true positives</u> (i.e. good populations included in the subset) and the expected number of <u>false positives</u> (i.e., a bad population included in the subset), respectively, using the procedure $\delta$. The problem of Lehmann is to determine a procedure $\delta$ for which $\sup_{\theta\varepsilon\Omega} R(\theta,\delta)$ is minimum subject to the condition that $\inf_{\theta\varepsilon\Omega'} S(\theta,\delta) \geq \gamma$, where $\Omega$ denotes the whole parameter space and $\Omega'$ denotes the set of parameter points for which there is at least one of the populations is positive.

3. Subset selection procedures for restricted families of probability distributions. In the previous section, we discussed selection procedures when the distributions under consideration are known except for the parameters involved. We also discussed some distribution-free procedures assuming only the stochastic ordering of the underlying distributions. Presently we will consider situations where, even though the functional forms of the distributions are not known, we do have some information about the family to which these distributions belong. This information could be useful in obtaining appropriate bounds for the probabilities in which we are interested. These bounds, as we will see, are in terms of some known distribution. For this reason, we will call these quasi-parametric cases. Many interesting examples of such situations arise in practice. The families of distributions having an increasing failure rate (IFR) or an increasing failure rate on the average (IFRA), which have been considered extensively in reliability theory are examples of this type. Such families of distributions can be described in the general setup of probability distributions, which are partially ordered in some sense with respect to a known distribution and this is described below.

3.1. Partial ordering in the space of probability distributions. A binary ordering relation ($\precsim$) is called a partial ordering in the space of probability distributions if

(a) $F \precsim F$ for all distributions $F$, and

(b) $F \precsim G$, $G \precsim H$ imply $F \precsim H$.

It should be noted that $F \precsim G$ and $G \precsim F$ do <u>not</u> necessarily imply that $F \equiv G$. The above definition has been used by early authors and followed in Barlow and Gupta [ 6]. However, it should be noted that a partial ordering defined as above corresponds to a quasi-ordering in the terminology of Barlow et al [ 2] who consider more general ordering of sets.

Assuming that all our distributions are absolutely continuous, we will now define some of the special order relations of interest to us. F and G denote distribution functions.

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

(ii) F is said to be <u>star-shaped</u> w.r.t. G (written $F \underset{*}{<} G$) 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.

(iii) F is said to be <u>r-ordered</u> w.r.t. 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) for x positive (negative).

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

If $G(x) = 1-e^{-x} (x \geq 0)$, then (i) defines the class of IFR distributions studied by Barlow, Marshall and Proschan [ 9], while (ii) defines the class of IFRA distributions studied by Barlow, Esary and Marshall [ 4]. The r-ordering defined by (iii) has been investigated by Lawrence [72]. Doksum [28] has used the tail-ordering. The convex ordering and s-ordering (not defined here) have been studied by

Van Zwet [87]. It is easy to verify that the above order relations are all partial order relations. One can also easily see [26] that convex ordering implies star-ordering.

3.2. Subset selection problems for quantiles. In this section we assume that the populations $\pi_i$, $i = 1,2,\ldots, k$, with the distributions $F_i$ have unique $\alpha$-quantiles, $\xi_{\alpha i}$, $i = 1,2,\ldots, k$. Let $F_{[i]}$ denote the cdf of the population with the ith smallest $\alpha$-quantile. We assume that

(a) $F_{[i]}(x) \geq F_{[k]}(x)$ for all $x$, $i = 1,2,\ldots, k$;

(3.1) (b) that there exists a distribution G such that

$$F_{[i]} \precsim G\ , \quad i = 1,2,\ldots, k,$$

where $\precsim$ denotes a partial ordering relation in the space of probability distributions. Of course, the correct pairing of the unordered and the ordered $F_i$'s is not known. We denote the space of the k-tuples $(F_1, F_2,\ldots, F_k)$ by $\Omega$. Our goal is to select a subset from the k populations so as to include the population with the largest $\alpha$-quantile. For this goal we define a selection procedure when the partial ordering in (3.1) is star-ordering.

Let $T_{j,i}$, $i = 1,2,\ldots, k$, denote the jth order statistic based on n independent observations from $\pi_i$ where $j \leq (n+1)\alpha < j+1$. Then, for selecting the population with the largest $\alpha$-quantile, the following rule $R_1$ was proposed by Barlow and Gupta [6].

$R_1$: Select $\pi_i$ if and only if

$$T_{j,i} \geq c \max_{1\leq r\leq k} T_{j,r}, \tag{3.2}$$

where $0 < c = c(k,P^*,n,j) < 1$ is determined so as to satisfy the probability requirement

$$\inf_{\Omega} P(CS|R_1) = P^* . \tag{3.3}$$

To enable the determination of the constant c, the following result has been obtained in [ 6].

<u>Theorem 3.1</u>. <u>If</u> $F_{[i]}(0) = G(0) = 0$, $F_{[i]}(x) \geq F_{[k]}(x)$, $x \geq 0$, $i = 1,2,\ldots, k$, <u>and</u> $F_{[k]} \underset{*}{<} G$, <u>then</u>

$$\inf_{\Omega} P(CS|R_1) = \int_0^\infty [G_j(x/c)]^{k-1} dG_j(x), \tag{3.4}$$

<u>where</u> $G_j(x)$ <u>is the</u> cdf <u>of the jth order statistic based on</u> n <u>independent observations from the distribution</u> G.

Thus the constant c which defines the procedure $R_1$ is determined by

$$\int_0^\infty [G_j(x/c)]^{k-1} dG_j(x) = P^*. \tag{3.5}$$

As we have noted earlier, when $G(x) = 1-e^{-x}$ $(x \geq 0)$, the populations have increasing failure rates on the average. For this case, the value of the constant c satisfying (3.5) has been tabulated by Barlow, Gupta and Panchapakesan [ 7] for $P^* = 0.75, 0.90, 0.95, 0.99$ and the following values of k,n and j corresponding to each selected value of $P^*$:

(i) $j = 1$, $k = 2(1)11$ (In this case c is independent of n);

(ii) $k = 2(1)4$, $n = 5(1)15$, $j = 2(1)n$;

(iii) $k = 5$, $n = 5(1)12$, $j = 2(1)n$;

(iv) $k = 6$, $n = 5(1)10$, $j = 2(1)n$.

These tables can be used also when the class of distributions F is star-ordered with respect to Weibull distribution $G_\lambda(x) = 1-\exp(-x^\lambda/\theta)$ for $x \geq 0$ and $\theta, \lambda > 0$. It is known that such a class of distributions is the smallest class of continuous distributions containing the Weibull class of distributions with shape parameter $\lambda$ which is closed under the formation of coherent structures and limits in distribution. For the selection problem in this case we use the rule $R_1$ with constant $c_\lambda = c^{1/\lambda}$, where c is obtained from the table in [ 7].

For the quantile selection problem, Rizvi and Sobel [87] proposed a distribution-free procedure $R_2$ which selects the population $\pi_i$ if and only if

$$T_{j,i} \geq \max_{1 \leq r \leq k} T_{j-a,r}, \tag{3.6}$$

where a is the smallest integer with $1 \leq a \leq j-1$ for which the P*-condition is satisfied. An asymptotic comparison of $R_1$ and $R_2$ has been made in [ 6]. Using the asymptotic theory of order statistics, the sample sizes $n_{R_i}(\varepsilon)$, i = 1,2, are found such that $E(S|R_i)-P(CS|R_i) = \varepsilon$. It should be noted that for any rule R, the expression $E(S|R) - P(CS|R)$ denotes the expected number of non-best populations included in the selected subset. For the slippage configuration $F_{[i]}(x) = F(x/\delta)$, $i = 1,\ldots, k-1$, and $F_{[k]}(x) = F(x)$, $0 < \delta < 1$, the <u>asymptotic relative efficiency</u> $ARE(R_1, R_2;\delta)$ of $R_1$ relative to $R_2$ is defined to be the limit as $\varepsilon \to 0$ of the ratio $n_{R_2}(\varepsilon)/n_{R_1}(\varepsilon)$. It has been shown in [ 6] that $ARE(R_1,R_2;\delta) = 1$.

A similar comparison of $R_1$ and the procedure (which we will call

R' here) for selecting the largest gamma parameter discussed in §2 gives

$$\text{(3.5)} \qquad \mathrm{ARE}(R_1,R';\delta) \geq 2(1-\delta)^2\,\bar{\alpha}^{-2}(-\log\bar{\alpha})^2 \cdot [r\,(\log\delta)^2\,\alpha\bar{\alpha}\,(1+\delta^2)]^{-1}, \quad r \geq 1,$$

where r is the common degrees of freedom of the gamma distributions and $\bar{\alpha} = 1-\alpha$. In particular, for r = 1, we get

$$\text{(3.6)} \qquad \mathrm{ARE}(R_1,R';\delta\uparrow 1) \geq \alpha^{-1}\bar{\alpha}(-\log\bar{\alpha})^2 = .493 \text{ when } \alpha = \tfrac{1}{2}.$$

The problem of selecting the population with the smallest $\alpha$-quantile can be handled in a similar way. Here, of course, we will assume that $F_{[i]}(x) \leq F_{[1]}(x)$, $i = 1,2,\ldots,k$ and all $x \geq 0$, and $F_{[i]} \underset{*}{\prec} G$. The selection procedure proposed in this case is

$R_3$: Select $\pi_i$ if and only if

$$\text{(3.7)} \qquad d\,T_{j,i} \leq \min_{1\leq r\leq k} T_{j,r}, \qquad j \leq (n+1)\alpha < j+1,$$

where $0 < d = d(k,P^*,n,j) < 1$ is a constant chosen so as to satisfy the $P^*$-requirement. It is shown in [ 6] that d is determined by

$$\text{(3.8)} \qquad \int_0^\infty [1-G_j(xd)]^{k-1}\,dG_j(x) = P^*.$$

The constants d are tabulated in [ 7] for $P^* = 0.75, 0.90, 0.95, 0.99$ and the following ranges of k,n and j for each $P^*$ value:

(i) $j = 1$, $k = 2(1)11$ (d is independent of n);

(ii) $k = 2(1)5$, $n = 5(1)15$, $j = 2(1)n$;

(iii) $k = 6$, $n = 5(1)12$, $j = 2(1)n$.

3.3. Selection with respect to the means for the class of IFR distributions. Let $\mu_i$ be the mean of the distribution $F(x;\mu_i)$, $i = 1, 2, \ldots, k$, and assume

$$(3.9)\qquad \begin{aligned} &\text{(a)}\quad F(x;\mu_{[i]}) \geq F(x;\mu_{[k]}) \text{ for all } x,\quad i = 1,2,\ldots, k; \\ &\text{(b)}\quad F(x;\mu_{[i]}) \underset{c}{\prec} G(x) = 1-e^{-x},\quad i = 1,2,\ldots, k. \end{aligned}$$

For convenience, we assume that $F(0;\mu_i) = 0$ for all i. For selecting a subset to include the population with the largest mean, $\mu_{[k]}$, Barlow and Gupta [ 6] have proposed the rule $R_3$: Select $\pi_i$ if and only if

$$(3.10)\qquad \bar{x}_i \geq c' \max_{1\leq r\leq k} \bar{x}_r,$$

where $\bar{x}_r$, $r = 1,2,\ldots, k$, are the sample means based on a random sample of n observations from each population. Let $U_{[i]}(x)$ denote the distribution of the mean of the sample from $F(x;\mu_{[i]})$. Then, as an immediate consequence of (3.9)-(a), we have $U_{[i]}(x) \geq U_{[k]}(x)$ for all x, $i = 1, 2,\ldots, k$. Since, the class of IFR distributions is closed under convolution [ 9], $U_{[i]} \underset{c}{\prec} G$, $i = 1,2,\ldots, k$. Using these properties, it has been shown in [ 6] that

$$(3.11)\qquad P(CS|R_3) \geq \int_0^\infty G^{k-1}(x/c')\, dG(x).$$

An obvious disadvantage of the above procedure is that the bound in (3.11), and hence the constant c' satisfying the P*-condition, is independent of n. However, if we restrict the class of distributions to the gamma family, we can obtain a lower bound for the probability of a correct selection which depends on n. By taking the density of the population $\pi_i$ to be $f(x;\theta_i) = \theta_i^{\alpha}x^{\alpha-1}e^{-\theta_i x}/\Gamma(\alpha)$, $x \geq 0$, $\theta_i > 0$, $i = 1,$

2,..., k, and $\alpha \geq 1$. The actual value of $\alpha$ is unknown. Then, selecting the population with the largest mean is equivalent ot selecting the population with the smallest $\theta$. In this it has been shown in [ 6] that

$$(3.12) \qquad P(CS|R_3) \geq \int_0^\infty [G^{(n)}(x/c')]^{k-1}\, dG^{(n)}(x),$$

where $G^{(\alpha)}$ denotes the distribution of a gamma variate with parameter $\alpha$. Thus the constant $c' = c'(k,P^*,n)$ is determined so as to satisfy the $P^*$-requirement by equating the right hand side of (3.12) to $P^*$. The values of $c'$ are tabulated by Gupta [47]. The problem of selecting the population with the largest $\theta$ can be handled in an analogous manner.

3.4. Selection with respect to the median for distributions r-ordered with respect to a specified distribution G. In this part we consider selection procedures with respect to the median for the distributions $F_i$, $i = 1,2,\ldots, k$, which have lighter tails than a specified distribution G. We say that the distribution $F_i$ has a lighter tail than G with $G(0) = 1/2$ if $F_i$ centered at its median, $\Delta_i$, is r-ordered with respect to G and $(d/dx)F_i(x+\Delta_i)\Big|_{x=0} \geq (d/dx)G(x)\Big|_{x=0}$ (see Doksum [28]). Our goal is to select a subset of the k populations which includes the population with the largest median, $\Delta_{[k]}$, which is assumed to be stochastically larger than any of the other populations. In this case, Barlow and Gupta [ 6] have considered the rule $R_4$: Select $\pi_i$ if and only if

$$(3.13) \qquad T_{j,i} \geq \max_{1\leq r\leq k} T_{j,r} - D, \qquad j \leq (n+1)/2 < j+1,$$

where the $T_{j,r}$ are as defined in §3.2. It has been shown that the constant $D = D(k,P^*,n,j) > 0$ satisfying the $P^*$-condition is given by

$$(3.14) \qquad \int_{-\infty}^{\infty} G_j^{k-1}(t+d)d\,G_j(t) = P^*.$$

It could be easily seen [28] that, if $F_i$ has a lighter tail than $G(G(0) = 1/2)$, then it follows that $F_i$, centered at its median $\Delta_i$, is tailed-ordered with respect to G as defined in §3.1. We can show that the procedure $R_4$ of Barlow and Gupta can, in fact, be used for selecting the population with the largest median, when all the populations $F_i$, centered at their respective medians, are tail-ordered with respect to a specified distribution G. Towards this end, we state the following theroem.

<u>Theorem 3.2</u>. <u>If</u> $F_{[i]}(x) \geq F_{[k]}(x)$ <u>for all</u> x, $i = 1,2,\ldots,k$, $G(0) = 1/2$ <u>and</u> $G^{-1}F_{[k]}(x+\Delta_{[k]})-x$ <u>is nondecreasing in</u> x <u>on</u> $\{x: 0 < F_{[k]}(x+\Delta_{[k]})\} < 1$, <u>then</u>

$$(3.5) \qquad \inf_{\Omega_1} P(CS|R_4) = \int_{-\infty}^{\infty} G_j^{k-1}(t+D)\,dG_j,$$

<u>where</u> $G_j$ <u>is as defined earlier in</u> §3.2 <u>and</u> $\Omega_1$ <u>is the space of the k-tuples</u> $(F_1,\ldots,F_k)$ <u>such that</u> $F_i \underset{t}{\prec} G$, $i = 1,2,\ldots,k$.

We leave out the proof of Theorem 3.2 in view of the next theorem which states a little more general result.

Suppose F and G are two continuous distributions with unique $\alpha$-quantiles, $\xi_\alpha$ and $\eta_\alpha$, respectively. We say that F is $\alpha$-<u>quantile tail-ordered</u> with respect to $G(F \underset{t_\alpha}{\prec} G)$ if $G^{-1}F(x+\xi_\alpha)-x-\eta_\alpha$ is nondecreasing in x on the support of F. If $\alpha = 1/2$ and $\xi_{1/2} = \eta_{1/2} = 0$, we have the tail-ordering defined in §3.1. For the $\alpha$-quantile tail-ordering we prove the following lemma.

Lemma 3.1. *Let X (Y) be a random variable having continuous distribution F (G) with a unique $\alpha$-quantile, $\xi_\alpha(\eta_\alpha)$. If $F \underset{t_\alpha}{\prec} G$, then*

$$\Pr(a+\eta_\alpha \le Y \le b+\eta_\alpha) \le \Pr(a+\xi_\alpha \le X \le b+\xi_\alpha), \tag{3.16}$$

*for every a and b such that $a < 0 < b$.*

Proof. Let F' and G' be the distributions of $X' = X-\xi_\alpha$ and $Y' = Y-\eta_\alpha$, respectively. Then, it is easy to see that $G'^{-1}F'(x)-x$ is non-decreasing in x. Letting $\varphi(x) = G'^{-1}F'(x)$, we have

$\Pr(a+\eta_\alpha \le Y \le b+\eta_\alpha)$

$= \Pr(\varphi(a) \le \varphi(Y') \le \varphi(b)) \ge \Pr(a \le X' \le b)$, since $\varphi(Y') \underset{st}{=} X'$ and $\varphi(a) \le a < 0 < b \le \varphi(b)$. The conclusion of the lemma follows immediately.

Now, consider k populations $\pi_1, \pi_2, \ldots, \pi_k$ with associated absolutely continuous distributions $F_i$, $i = 1,2,\ldots, k$. We assume that these distributions have unique $\alpha$-quantiles, $\xi_{\alpha,i}$. Let $F_{[k]}$ denote the distribution which has the ith smallest $\alpha$-quantile. Let G be a specified distribution with the unique $\alpha$-quantile, $\eta_\alpha$. In order to select a subset including the population having the largest $\alpha$-quantile, $\xi_{\alpha[k]}$, assuming that the distributions $F_i$ belong to a family of distributions which are $\alpha$-quantile tail-ordered with respect to G, we propose the rule $R_5$: Select $\pi_i$ if and only if

(3.17) $$T_{j,i} \geq \max_{1 \leq r \leq k} T_{j,r} - D, \qquad j \leq (n+1)\alpha < j+1,$$

where $T_{j,r}$ is the jth order statistic based on n independent observations from $\pi_r$ and $D > 0$ is a constant to be determined so as to satisfy the P*-requirement. We now state and prove a theorem, which helps us to determine D.

Theorem 3.3. *If $F_{[i]}(x) \geq F_{[k]}(x)$ for all x, $i = 1,2,\ldots, k$, and $F_{[i]} \underset{t_\alpha}{\prec} G$, then*

(3.18) $$\inf_{\Omega'} (CS|R) = \int_{-\infty}^{\infty} G_j^{k-1}(t+D)\, dG_j(t),$$

*where $G_j$ is the cdf of the jth order statistic based on n independent observations from G, and $\Omega'$ is the space of all k-tuples $(F_1,\ldots, F_k)$ satisfying the hypothesis of the theorem.*

Proof. Since the stochastic ordering is preserved by the order statistics, we have

(3.19) $$P(CS|R_5) \geq Pr(X_{j,k} \geq \max_{1 \leq r \leq k} X_{j,r} - D),$$

where the $X_{j,r}$ are independent and identically distributed having the same distribution as that of the jth order statistic based on n independent observations from $F_{[k]}$, which is denoted by $F_{j,k}$. Let U(x) be the cdf of $X_{j,r} - \xi_{\alpha[k]}$ and $\varphi(x) = G_j^{-1}U(x) - \eta_\alpha$. If we now set $Y_{j,r} - \eta_\alpha = \varphi(X_{j,r} - \xi_{\alpha[k]})$, it is easy to see that $Y_{j,r}$ has the distribution $G_j$. It is also easy to verify that $\varphi(x)-x$ is nondecreasing in x. Hence the event $[\varphi\{\max_{1 \leq r \leq k} (X_{j,r} - \xi_{\alpha[k]})\} - \varphi\{X_{j,k} - \xi_{\alpha[k]}\} \leq D]$ implies the event $[\max_{1 \leq r \leq k} (X_{j,r} - \xi_{\alpha[k]}) - (X_{j,k} - \xi_{\alpha[k]}) \leq D]$. Thus, we obtain

$$\Pr\left(\max_{1\le r\le k} (Y_{j,r} - \eta_\alpha) - (Y_{j,k} - \eta_\alpha) \le D\right) \tag{3.20}$$

$$\le \Pr\left(\max_{1\le r\le k} (X_{j,r} - \xi_{\alpha[k]}) - (X_{j,k} - \xi_{\alpha[k]}) \le D\right).$$

From (3.19) and (3.20) we get

$$P(CS|R_5) \ge \Pr(Y_{j,k} \ge \max_{1\le r\le k} Y_{j,r} - D). \tag{3.21}$$

This completes the proof of the theorem.

Thus the value of $D = D(k,P^*,n,j)$ is determined by the equation

$$\int_{-\infty}^{\infty} G_j^{k-1}(t+D)\, dG_j(t) = P^*. \tag{3.22}$$

Values of D are given in Table 1 for $k = 2(1)10$, $n = 5(2)15$ and $P^* = 0.75$, 0.90, 0.95 and 0.99 when G is chosen to be the logistic distribution $G(x) = [1+e^{-x}]^{-1}$. In this case, $F \underset{t_\alpha}{\prec} G$ implies $f(x+\xi_\alpha) \ge F(x+\xi_\alpha)[1-F(x+\xi_\alpha)]$ for all x on the support of F, where $\xi_\alpha$ is the $\alpha$-quantile of F. A brief description of the computational methods used in the evaluation of the integral in (3.22) is given below.

Let $u(t,D) = G_j^{k-1}(t+D)\, g_j(t)$ so that (3.22) becomes

$$A(D) = \int_{-\infty}^{\infty} u(t,D)\, dt = P^*. \tag{3.23}$$

Splitting the region of integration into three parts, we write

$$A(D) = \int_{-\infty}^{u_1} u(t,D)\, dt + \int_{u_1}^{u_2} u(t,D)\, dt + \int_{u_2}^{\infty} u(t,D)\, dt \tag{3.24}$$

$$= A_1(D) + A_2(D) + A_3(D), \text{ say.}$$

The numbers $u_1$ and $u_2$ are chosen such that $|A_1(D)| \le 10^{-13}$ and $|A_3(D)| \le 10^{-13}$ so that the total error in omitting the tail parts does

Table 1. D-values of procedure $R_5$ for selection of quantiles

$P^* = 0.75$

| n | j \ k | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 1 | 1.20074 | 1.73431 | 1.99367 | 2.15934 | 2.27897 | 2.37163 | 2.44676 | 2.50964 | 2.56353 |
| | 2 | .89247 | 1.31954 | 1.53713 | 1.68196 | 1.78534 | 1.86815 | 1.93610 | 1.99354 | 2.04318 |
| | 3 | .83162 | 1.25302 | 1.47471 | 1.62323 | 1.73413 | 1.82224 | 1.89514 | 1.95718 | 2.01111 |
| | 4 | .89247 | 1.37390 | 1.63615 | 1.81544 | 1.95120 | 2.06021 | 2.15117 | 2.22912 | 2.29728 |
| | 5 | 1.20074 | 1.92017 | 2.33447 | 2.62658 | 2.85239 | 3.03649 | 3.19131 | 3.32640 | 3.44432 |
| 7 | 1 | 1.17118 | 1.63640 | 1.93465 | 2.09218 | 2.20534 | 2.29261 | 2.36211 | 2.42192 | 2.47218 |
| | 2 | .83772 | 1.22979 | 1.42628 | 1.55405 | 1.64743 | 1.72040 | 1.77996 | 1.83009 | 1.87325 |
| | 3 | .73526 | 1.03396 | 1.27808 | 1.39954 | 1.48921 | 1.55982 | 1.61781 | 1.66686 | 1.70928 |
| | 4 | .70854 | 1.05690 | 1.25464 | 1.37908 | 1.47330 | 1.54726 | 1.60832 | 1.66019 | 1.70521 |
| | 5 | .73526 | 1.12164 | 1.32843 | 1.46827 | 1.57324 | 1.65702 | 1.72657 | 1.78591 | 1.83761 |
| | 6 | .83772 | 1.31069 | 1.55545 | 1.73949 | 1.86344 | 1.97043 | 2.05985 | 2.13658 | 2.20374 |
| | 7 | 1.17118 | 1.88285 | 2.29453 | 2.58533 | 2.81034 | 2.99391 | 3.14896 | 3.28316 | 3.40146 |
| 9 | 1 | 1.15487 | 1.55019 | 1.90250 | 2.05569 | 2.16541 | 2.24981 | 2.31785 | 2.37451 | 2.42285 |
| | 2 | .80935 | 1.18468 | 1.37091 | 1.49132 | 1.57894 | 1.64718 | 1.70271 | 1.74933 | 1.78938 |
| | 3 | .69220 | 1.02384 | 1.19205 | 1.30214 | 1.38294 | 1.44628 | 1.49810 | 1.54179 | 1.57947 |
| | 4 | .64211 | .95822 | 1.12113 | 1.22875 | 1.30826 | 1.37090 | 1.42235 | 1.46588 | 1.50352 |
| | 5 | .62754 | .94458 | 1.11026 | 1.22164 | 1.30268 | 1.36760 | 1.42113 | 1.46655 | 1.50593 |
| | 6 | .64211 | .97507 | 1.15172 | 1.27044 | 1.35922 | 1.42980 | 1.48823 | 1.53796 | 1.58119 |
| | 7 | .69220 | 1.05281 | 1.26285 | 1.39868 | 1.50098 | 1.58278 | 1.65078 | 1.70888 | 1.75955 |
| | 8 | .80935 | 1.25356 | 1.51451 | 1.68738 | 1.81889 | 1.92485 | 2.01347 | 2.08957 | 2.15622 |
| | 9 | 1.15487 | 1.95232 | 2.27258 | 2.56266 | 2.78724 | 2.97053 | 3.12536 | 3.25941 | 3.37759 |
| 11 | 1 | 1.14455 | 1.64368 | 1.88229 | 2.03279 | 2.14037 | 2.22300 | 2.28951 | 2.34484 | 2.39139 |
| | 2 | .79315 | 1.15753 | 1.33770 | 1.45378 | 1.53801 | 1.60346 | 1.65664 | 1.70120 | 1.73944 |
| | 3 | .66770 | .93423 | 1.14363 | 1.24747 | 1.32342 | 1.36278 | 1.43123 | 1.47201 | 1.50711 |
| | 4 | .60726 | .91200 | 1.05242 | 1.15119 | 1.22384 | 1.28087 | 1.32758 | 1.36700 | 1.40102 |
| | 5 | .57738 | .85423 | 1.01199 | 1.10968 | 1.18186 | 1.23874 | 1.28546 | 1.32439 | 1.35917 |
| | 6 | .56912 | .85643 | 1.00633 | 1.10606 | 1.18009 | 1.23861 | 1.28683 | 1.32771 | 1.36312 |
| | 7 | .57738 | .87559 | 1.03258 | 1.13771 | 1.21611 | 1.27831 | 1.32970 | 1.37337 | 1.41128 |
| | 8 | .60726 | .92701 | 1.09778 | 1.21297 | 1.29932 | 1.36811 | 1.42511 | 1.47370 | 1.51597 |
| | 9 | .66770 | 1.02931 | 1.22547 | 1.35902 | 1.45977 | 1.54042 | 1.60753 | 1.66491 | 1.71498 |
| | 10 | .79315 | 1.24109 | 1.48973 | 1.66129 | 1.79193 | 1.89725 | 1.98539 | 2.06111 | 2.12744 |
| | 11 | 1.14455 | 1.94934 | 2.25870 | 2.54833 | 2.77265 | 2.95575 | 3.11046 | 3.24441 | 3.36252 |
| 13 | 1 | 1.13742 | 1.63232 | 1.86841 | 2.01708 | 2.12321 | 2.20463 | 2.27011 | 2.32453 | 2.37087 |
| | 2 | .78138 | 1.13941 | 1.31557 | 1.42878 | 1.51079 | 1.57442 | 1.62604 | 1.66926 | 1.70630 |
| | 3 | .65136 | .95875 | 1.11257 | 1.21245 | 1.28533 | 1.34218 | 1.38852 | 1.42746 | 1.46034 |
| | 4 | .58592 | .85752 | 1.01041 | 1.10387 | 1.17241 | 1.22610 | 1.26998 | 1.30696 | 1.33891 |
| | 5 | .54924 | .81800 | .95566 | 1.04623 | 1.11292 | 1.16532 | 1.20827 | 1.24453 | 1.27582 |
| | 6 | .53032 | .73391 | .93012 | 1.02019 | 1.08677 | 1.13922 | 1.18231 | 1.21876 | 1.25028 |
| | 7 | .52443 | .78906 | .92696 | 1.01861 | 1.08658 | 1.14029 | 1.18450 | 1.22197 | 1.25441 |
| | 8 | .53032 | .81209 | .94493 | 1.04035 | 1.11138 | 1.16766 | 1.21409 | 1.25351 | 1.28769 |
| | 9 | .54924 | .83552 | .98741 | 1.08944 | 1.16569 | 1.22628 | 1.27640 | 1.31904 | 1.35638 |
| | 10 | .58592 | .93732 | 1.06444 | 1.17744 | 1.26228 | 1.32993 | 1.38605 | 1.43391 | 1.47558 |
| | 11 | .65136 | 1.01765 | 1.20129 | 1.33335 | 1.43309 | 1.51300 | 1.57953 | 1.63644 | 1.68612 |
| | 12 | .78138 | 1.22603 | 1.47312 | 1.64379 | 1.77385 | 1.87875 | 1.96656 | 2.04202 | 2.10814 |
| | 13 | 1.13742 | 1.84038 | 2.24913 | 2.53846 | 2.76259 | 2.94557 | 3.10019 | 3.23407 | 3.35213 |
| 15 | 1 | 1.13221 | 1.62403 | 1.85830 | 2.00564 | 2.11072 | 2.19127 | 2.25600 | 2.30976 | 2.35552 |
| | 2 | .77390 | 1.12644 | 1.29976 | 1.41095 | 1.49139 | 1.55372 | 1.60425 | 1.64652 | 1.68272 |
| | 3 | .64078 | .94038 | 1.09093 | 1.18808 | 1.25895 | 1.31398 | 1.35886 | 1.39654 | 1.42890 |
| | 4 | .57126 | .84418 | .98204 | 1.07196 | 1.13776 | 1.18921 | 1.23122 | 1.26656 | 1.29699 |
| | 5 | .53043 | .73786 | .91904 | 1.00505 | 1.06824 | 1.11778 | 1.15833 | 1.19251 | 1.22198 |
| | 6 | .50638 | .75507 | .88293 | .96714 | 1.02920 | 1.07799 | 1.11798 | 1.15177 | 1.18093 |
| | 7 | .49238 | .73861 | .86562 | .94963 | 1.01172 | 1.06064 | 1.10082 | 1.13481 | 1.16419 |
| | 8 | .48894 | .73541 | .86379 | .94905 | 1.01225 | 1.06215 | 1.10321 | 1.13800 | 1.16811 |
| | 9 | .49238 | .74478 | .87680 | .96483 | 1.03028 | 1.08207 | 1.12476 | 1.16099 | 1.19238 |
| | 10 | .50608 | .75807 | .90646 | .99916 | 1.06829 | 1.12313 | 1.16842 | 1.20692 | 1.24032 |
| | 11 | .53043 | .81924 | .95776 | 1.05775 | 1.13257 | 1.19210 | 1.24137 | 1.28332 | 1.31978 |
| | 12 | .57126 | .87714 | 1.04177 | 1.15326 | 1.23706 | 1.30395 | 1.35946 | 1.40683 | 1.44808 |
| | 13 | .64078 | .93248 | 1.18436 | 1.31537 | 1.41440 | 1.49379 | 1.55990 | 1.61649 | 1.66589 |
| | 14 | .77330 | 1.21523 | 1.46121 | 1.63124 | 1.76088 | 1.86548 | 1.95306 | 2.02833 | 2.09430 |
| | 15 | 1.13221 | 1.83384 | 2.24214 | 2.53124 | 2.75524 | 2.93813 | 3.09269 | 3.22652 | 3.34454 |

Each entry is the D-value satisfying $\Pr(Y_{j,k} \geq \max_{1 \leq r \leq k} Y_{j,r} - D) = P^*$ where $Y_{j,1}, Y_{j,2}, \ldots, Y_{j,k}$ are i.i.d. having the same distribution as the jth order statistic based on n independent observations from the logistic distribution with density $e^{-x}/(1+e^{-x})^2$.

Table 1 (cont'd.)

$P^* = 0.90$

| n | j | k = 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 1 | 2.36716 | 2.94555 | 3.08564 | 3.24123 | 3.35454 | 3.44282 | 3.51469 | 3.57504 | 3.62690 |
| | 2 | 1.71841 | 2.13519 | 2.29289 | 2.42474 | 2.52276 | 2.60035 | 2.66433 | 2.71862 | 2.76558 |
| | 3 | 1.59424 | 1.95984 | 2.17261 | 2.31025 | 2.41387 | 2.49668 | 2.56550 | 2.62428 | 2.67553 |
| | 4 | 1.71341 | 2.13921 | 2.40439 | 2.57360 | 2.71265 | 2.80676 | 2.89393 | 2.96886 | 3.03453 |
| | 5 | 2.36715 | 3.07184 | 3.48118 | 3.77081 | 3.99512 | 4.17822 | 4.33293 | 4.46688 | 4.58439 |
| 7 | 1 | 2.31654 | 2.77981 | 3.01013 | 3.15845 | 3.26585 | 3.34914 | 3.41659 | 3.47324 | 3.52168 |
| | 2 | 1.61628 | 1.95081 | 2.13896 | 2.25657 | 2.34334 | 2.41160 | 2.46759 | 2.51485 | 2.55574 |
| | 3 | 1.48859 | 1.72403 | 1.89055 | 2.00197 | 2.08498 | 2.15976 | 2.20595 | 2.25115 | 2.29114 |
| | 4 | 1.35514 | 1.67262 | 1.84322 | 1.95860 | 2.04522 | 2.11428 | 2.17156 | 2.22040 | 2.26231 |
| | 5 | 1.40859 | 1.75595 | 1.94593 | 2.07584 | 2.17414 | 2.25299 | 2.31870 | 2.37497 | 2.42411 |
| | 6 | 1.61628 | 2.04367 | 2.28292 | 2.44982 | 2.57556 | 2.67799 | 2.76386 | 2.83774 | 2.90255 |
| | 7 | 2.31654 | 3.01759 | 3.42580 | 3.71481 | 3.93375 | 4.12160 | 4.27613 | 4.40995 | 4.52795 |
| 9 | 1 | 2.28324 | 2.74437 | 2.96961 | 3.11402 | 3.21828 | 3.29894 | 3.36422 | 3.41875 | 3.46533 |
| | 2 | 1.56529 | 1.93438 | 2.06321 | 2.17407 | 2.25550 | 2.31934 | 2.37156 | 2.41556 | 2.45348 |
| | 3 | 1.32732 | 1.51764 | 1.76930 | 1.87094 | 1.94468 | 2.00357 | 2.05199 | 2.09298 | 2.12844 |
| | 4 | 1.22727 | 1.53480 | 1.65173 | 1.75016 | 1.82352 | 1.88169 | 1.92970 | 1.97048 | 2.00585 |
| | 5 | 1.19958 | 1.47844 | 1.62840 | 1.72961 | 1.80545 | 1.86584 | 1.91586 | 1.95846 | 1.99551 |
| | 6 | 1.22727 | 1.52397 | 1.68492 | 1.79438 | 1.87686 | 1.94282 | 1.99763 | 2.04446 | 2.08528 |
| | 7 | 1.32732 | 1.55226 | 1.84663 | 1.97314 | 2.06908 | 2.14617 | 2.21050 | 2.26563 | 2.31383 |
| | 8 | 1.56529 | 1.93604 | 2.22248 | 2.38669 | 2.51230 | 2.61389 | 2.69912 | 2.77248 | 2.83686 |
| | 9 | 2.28024 | 2.93839 | 3.39586 | 3.68455 | 3.90829 | 4.09102 | 4.24546 | 4.37920 | 4.49716 |
| 11 | 1 | 2.27295 | 2.72223 | 2.94430 | 3.08631 | 3.18965 | 3.26768 | 3.33155 | 3.38483 | 3.43035 |
| | 2 | 1.53472 | 1.95477 | 2.01817 | 2.12509 | 2.20342 | 2.26469 | 2.31471 | 2.35680 | 2.39312 |
| | 3 | 1.28149 | 1.55813 | 1.70172 | 1.79668 | 1.86679 | 1.92196 | 1.96723 | 2.00547 | 2.03649 |
| | 4 | 1.16129 | 1.41838 | 1.55428 | 1.64441 | 1.71131 | 1.76417 | 1.80759 | 1.84456 | 1.87647 |
| | 5 | 1.10346 | 1.35443 | 1.48746 | 1.57661 | 1.64308 | 1.69578 | 1.73928 | 1.77623 | 1.80827 |
| | 6 | 1.08611 | 1.33904 | 1.47435 | 1.56555 | 1.63381 | 1.68811 | 1.73305 | 1.77130 | 1.80453 |
| | 7 | 1.10346 | 1.35723 | 1.50959 | 1.60608 | 1.67860 | 1.73647 | 1.78449 | 1.82545 | 1.86111 |
| | 8 | 1.16129 | 1.44727 | 1.60322 | 1.70959 | 1.78931 | 1.85423 | 1.90775 | 1.95352 | 1.99345 |
| | 9 | 1.28149 | 1.53952 | 1.79074 | 1.91533 | 2.00993 | 2.08602 | 2.14957 | 2.20406 | 2.25173 |
| | 10 | 1.53472 | 1.95156 | 2.18630 | 2.34952 | 2.47444 | 2.57554 | 2.66038 | 2.73343 | 2.79755 |
| | 11 | 2.27295 | 2.97004 | 3.37712 | 3.66561 | 3.88924 | 4.07189 | 4.22627 | 4.35997 | 4.47789 |
| 13 | 1 | 2.26927 | 2.70708 | 2.92790 | 3.06740 | 3.16842 | 3.24635 | 3.30926 | 3.36170 | 3.40646 |
| | 2 | 1.51436 | 1.92847 | 1.98832 | 2.09266 | 2.16896 | 2.22856 | 2.27715 | 2.31798 | 2.35309 |
| | 3 | 1.25296 | 1.52909 | 1.65861 | 1.74994 | 1.81722 | 1.87007 | 1.91336 | 1.94988 | 1.98139 |
| | 4 | 1.12098 | 1.35671 | 1.49509 | 1.58030 | 1.64337 | 1.69310 | 1.73396 | 1.76853 | 1.79841 |
| | 5 | 1.04979 | 1.28363 | 1.40728 | 1.48978 | 1.55109 | 1.59957 | 1.63950 | 1.67335 | 1.70266 |
| | 6 | 1.01164 | 1.24255 | 1.36512 | 1.44712 | 1.50832 | 1.55685 | 1.59691 | 1.63092 | 1.66042 |
| | 7 | 1.00919 | 1.23275 | 1.35701 | 1.44066 | 1.50323 | 1.55296 | 1.59409 | 1.62908 | 1.65947 |
| | 8 | 1.01164 | 1.25171 | 1.38084 | 1.46817 | 1.53369 | 1.58590 | 1.62918 | 1.65605 | 1.69812 |
| | 9 | 1.04879 | 1.31329 | 1.44126 | 1.53500 | 1.60559 | 1.66199 | 1.70884 | 1.74884 | 1.78368 |
| | 10 | 1.12088 | 1.41031 | 1.55319 | 1.65766 | 1.73665 | 1.79996 | 1.85268 | 1.89779 | 1.93717 |
| | 11 | 1.25296 | 1.57567 | 1.75487 | 1.87822 | 1.97197 | 2.04741 | 2.11045 | 2.16453 | 2.21185 |
| | 12 | 1.51436 | 1.92861 | 2.16223 | 2.32478 | 2.44926 | 2.55002 | 2.63460 | 2.70745 | 2.77140 |
| | 13 | 2.26927 | 2.95747 | 3.36429 | 3.65265 | 3.87620 | 4.05879 | 4.21314 | 4.34681 | 4.46471 |
| 15 | 1 | 2.25169 | 2.63608 | 2.91444 | 3.05366 | 3.15373 | 3.23086 | 3.29308 | 3.34491 | 3.38913 |
| | 2 | 1.49992 | 1.83974 | 1.96709 | 2.06961 | 2.14448 | 2.20289 | 2.25047 | 2.29043 | 2.32476 |
| | 3 | 1.23156 | 1.43366 | 1.62871 | 1.71756 | 1.78230 | 1.83416 | 1.87609 | 1.91144 | 1.94191 |
| | 4 | 1.09357 | 1.33150 | 1.45528 | 1.53723 | 1.59777 | 1.64543 | 1.68454 | 1.71759 | 1.74612 |
| | 5 | 1.01318 | 1.23771 | 1.35540 | 1.43368 | 1.49171 | 1.53752 | 1.57520 | 1.60709 | 1.63467 |
| | 6 | .96547 | 1.13297 | 1.29769 | 1.37432 | 1.43129 | 1.47637 | 1.51351 | 1.54500 | 1.57228 |
| | 7 | .93988 | 1.15492 | 1.26901 | 1.34549 | 1.40251 | 1.44772 | 1.48503 | 1.51672 | 1.54419 |
| | 8 | .93178 | 1.14829 | 1.26380 | 1.34151 | 1.39959 | 1.44574 | 1.48389 | 1.51633 | 1.54449 |
| | 9 | .93998 | 1.15180 | 1.28088 | 1.36128 | 1.42154 | 1.46950 | 1.50922 | 1.54305 | 1.57245 |
| | 10 | .96547 | 1.19748 | 1.32274 | 1.40763 | 1.47143 | 1.52233 | 1.56455 | 1.60056 | 1.63190 |
| | 11 | 1.01318 | 1.25164 | 1.39673 | 1.48868 | 1.55800 | 1.61344 | 1.65952 | 1.69888 | 1.73319 |
| | 12 | 1.09357 | 1.35857 | 1.51937 | 1.62255 | 1.70063 | 1.76325 | 1.81543 | 1.86009 | 1.89909 |
| | 13 | 1.23156 | 1.53209 | 1.72989 | 1.85238 | 1.94552 | 2.02052 | 2.08321 | 2.13700 | 2.18408 |
| | 14 | 1.49992 | 1.91223 | 2.14505 | 2.30713 | 2.43129 | 2.53181 | 2.61622 | 2.68892 | 2.75274 |
| | 15 | 2.25169 | 2.94832 | 3.35496 | 3.64322 | 3.86671 | 4.04927 | 4.20359 | 4.33724 | 4.45512 |

Each entry is the D-value satisfying $Pr(Y_{j,k} \geq \max_{1 \leq r \leq k} Y_{j,r} - D) = P^*$ where $Y_{j,1}, Y_{j,2}, \ldots, Y_{j,k}$ are i.i.d. having the same distribution as the jth order statistic based on n independent observations from the logistic distribution with density $e^{-x}/(1+e^{-x})^2$.

Table 1 (cont'd.)

$P^* = 0.95$

| n | j | k = 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 1 | 3.13964 | 3.60038 | 3.83384 | 3.98609 | 4.09731 | 4.18413 | 4.25492 | 4.31443 | 4.36561 |
| | 2 | 2.23218 | 2.53601 | 2.77439 | 2.90092 | 2.99542 | 3.07046 | 3.13248 | 3.18522 | 3.23101 |
| | 3 | 2.06746 | 2.41493 | 2.60778 | 2.73959 | 2.83926 | 2.91917 | 2.98574 | 3.04272 | 3.09246 |
| | 4 | 2.23218 | 2.65174 | 2.88756 | 3.05138 | 3.17670 | 3.27807 | 3.36712 | 3.43634 | 3.50959 |
| | 5 | 3.13964 | 3.83888 | 4.24639 | 4.53510 | 4.75886 | 4.94159 | 5.09624 | 5.22979 | 5.34774 |
| 7 | 1 | 3.08070 | 3.52751 | 3.75235 | 3.89768 | 4.00324 | 4.08526 | 4.15138 | 4.20770 | 4.25556 |
| | 2 | 2.10350 | 2.42729 | 2.59717 | 2.71014 | 2.79396 | 2.85992 | 2.91424 | 2.96022 | 2.99939 |
| | 3 | 1.82152 | 2.11584 | 2.27347 | 2.37972 | 2.45025 | 2.52250 | 2.57493 | 2.61936 | 2.65804 |
| | 4 | 1.74955 | 2.04593 | 2.20733 | 2.31726 | 2.40016 | 2.46649 | 2.52163 | 2.56875 | 2.60982 |
| | 5 | 1.82152 | 2.14796 | 2.32863 | 2.45295 | 2.54741 | 2.62342 | 2.68631 | 2.74137 | 2.78901 |
| | 6 | 2.10350 | 2.51173 | 2.74250 | 2.90328 | 3.02653 | 3.12637 | 3.21023 | 3.28248 | 3.34594 |
| | 7 | 3.08070 | 3.77794 | 4.18477 | 4.47314 | 4.69669 | 4.87929 | 5.03363 | 5.16731 | 5.28521 |
| 9 | 1 | 3.04936 | 3.48687 | 3.70894 | 3.85051 | 3.95308 | 4.03257 | 4.09699 | 4.15085 | 4.19695 |
| | 2 | 2.04016 | 2.34987 | 2.51105 | 2.61765 | 2.69632 | 2.75817 | 2.80689 | 2.85171 | 2.88656 |
| | 3 | 1.71790 | 1.93853 | 2.13207 | 2.22814 | 2.29965 | 2.35628 | 2.40296 | 2.44256 | 2.47687 |
| | 4 | 1.58390 | 1.84199 | 1.98048 | 2.07401 | 2.14407 | 2.19981 | 2.24594 | 2.28520 | 2.31931 |
| | 5 | 1.54553 | 1.81597 | 1.94740 | 2.04352 | 2.11530 | 2.17371 | 2.22173 | 2.26271 | 2.29841 |
| | 6 | 1.58390 | 1.85094 | 2.01314 | 2.11733 | 2.19619 | 2.25944 | 2.31215 | 2.35726 | 2.39665 |
| | 7 | 1.71790 | 2.03318 | 2.20870 | 2.32984 | 2.42207 | 2.49639 | 2.55855 | 2.61192 | 2.65865 |
| | 8 | 2.04016 | 2.44351 | 2.67137 | 2.83372 | 2.95238 | 3.05208 | 3.13537 | 3.20717 | 3.27024 |
| | 9 | 3.04936 | 3.74529 | 4.15178 | 4.43998 | 4.66343 | 4.84596 | 5.00625 | 5.13390 | 5.25177 |
| 11 | 1 | 3.02934 | 3.45477 | 3.68179 | 3.82120 | 3.92132 | 3.99985 | 4.06290 | 4.11555 | 4.16057 |
| | 2 | 2.00249 | 2.33399 | 2.46018 | 2.56312 | 2.63984 | 2.69826 | 2.74638 | 2.78787 | 2.82319 |
| | 3 | 1.65985 | 1.91794 | 2.05392 | 2.14453 | 2.21175 | 2.26483 | 2.30650 | 2.34546 | 2.37743 |
| | 4 | 1.49921 | 1.73866 | 1.86622 | 1.95185 | 2.01573 | 2.06678 | 2.10819 | 2.14368 | 2.17447 |
| | 5 | 1.42238 | 1.65529 | 1.78050 | 1.86503 | 1.92835 | 1.97874 | 2.02045 | 2.05595 | 2.08679 |
| | 6 | 1.39925 | 1.63422 | 1.76157 | 1.84990 | 1.91391 | 1.96490 | 2.00795 | 2.04467 | 2.07663 |
| | 7 | 1.42238 | 1.65791 | 1.80229 | 1.89365 | 1.96278 | 2.01813 | 2.06417 | 2.10352 | 2.13784 |
| | 8 | 1.49921 | 1.76665 | 1.91418 | 2.01544 | 2.09222 | 2.15390 | 2.20535 | 2.24943 | 2.28794 |
| | 9 | 1.65985 | 1.95922 | 2.14199 | 2.26127 | 2.35226 | 2.42565 | 2.48717 | 2.53984 | 2.58635 |
| | 10 | 2.00249 | 2.41222 | 2.62918 | 2.78768 | 2.90937 | 3.00805 | 3.09130 | 3.16253 | 3.22536 |
| | 11 | 3.02934 | 3.72495 | 4.13124 | 4.41934 | 4.64273 | 4.82521 | 4.97948 | 5.11310 | 5.23095 |
| 13 | 1 | 3.01597 | 3.44573 | 3.66336 | 3.80123 | 3.90069 | 3.97755 | 4.03967 | 4.09151 | 4.13579 |
| | 2 | 1.97752 | 2.27769 | 2.42661 | 2.52712 | 2.60095 | 2.65878 | 2.70614 | 2.74583 | 2.78039 |
| | 3 | 1.62285 | 1.87305 | 2.00471 | 2.09152 | 2.15607 | 2.20695 | 2.24873 | 2.28406 | 2.31458 |
| | 4 | 1.44775 | 1.67617 | 1.79724 | 1.87822 | 1.93946 | 1.98613 | 2.02540 | 2.05869 | 2.08752 |
| | 5 | 1.35215 | 1.55991 | 1.68623 | 1.76444 | 1.82284 | 1.86918 | 1.90746 | 1.93998 | 1.96819 |
| | 6 | 1.30316 | 1.51796 | 1.63214 | 1.70986 | 1.76808 | 1.81441 | 1.85275 | 1.88539 | 1.91374 |
| | 7 | 1.28734 | 1.51356 | 1.62032 | 1.69948 | 1.75898 | 1.80643 | 1.84578 | 1.87933 | 1.90852 |
| | 8 | 1.30316 | 1.52603 | 1.64749 | 1.73018 | 1.79252 | 1.84236 | 1.88377 | 1.91913 | 1.94995 |
| | 9 | 1.35215 | 1.59922 | 1.71927 | 1.80822 | 1.87549 | 1.92941 | 1.97431 | 2.01271 | 2.04624 |
| | 10 | 1.44775 | 1.70934 | 1.85404 | 1.95353 | 2.02904 | 2.08976 | 2.14644 | 2.18388 | 2.22196 |
| | 11 | 1.62285 | 1.92842 | 2.09930 | 2.21756 | 2.30777 | 2.38056 | 2.44151 | 2.49389 | 2.53978 |
| | 12 | 1.97752 | 2.37521 | 2.60125 | 2.75921 | 2.88052 | 2.97892 | 3.06166 | 3.13301 | 3.19571 |
| | 13 | 3.01597 | 3.71108 | 4.11723 | 4.40525 | 4.62860 | 4.81106 | 4.96531 | 5.09892 | 5.21676 |
| 15 | 1 | 3.00609 | 3.43640 | 3.64998 | 3.78674 | 3.88529 | 3.96138 | 4.02283 | 4.07407 | 4.11782 |
| | 2 | 1.95975 | 2.25218 | 2.40279 | 2.50161 | 2.57409 | 2.63081 | 2.67711 | 2.71606 | 2.74957 |
| | 3 | 1.59718 | 1.84198 | 1.97002 | 2.05491 | 2.11764 | 2.16701 | 2.20751 | 2.24172 | 2.27125 |
| | 4 | 1.41311 | 1.63425 | 1.75103 | 1.82895 | 1.88680 | 1.93250 | 1.97010 | 2.00194 | 2.02949 |
| | 5 | 1.30664 | 1.51479 | 1.62551 | 1.69973 | 1.75502 | 1.79883 | 1.83496 | 1.86561 | 1.89217 |
| | 6 | 1.24369 | 1.44510 | 1.55286 | 1.62537 | 1.67356 | 1.72259 | 1.75814 | 1.78835 | 1.81456 |
| | 7 | 1.21090 | 1.40907 | 1.51616 | 1.58848 | 1.64266 | 1.68576 | 1.72144 | 1.75180 | 1.77817 |
| | 8 | 1.19936 | 1.39958 | 1.50830 | 1.58176 | 1.63693 | 1.68092 | 1.71738 | 1.74845 | 1.77547 |
| | 9 | 1.21090 | 1.41578 | 1.52762 | 1.60366 | 1.66030 | 1.70663 | 1.74459 | 1.77699 | 1.80520 |
| | 10 | 1.24369 | 1.45926 | 1.57708 | 1.65746 | 1.71813 | 1.76670 | 1.80709 | 1.84161 | 1.87170 |
| | 11 | 1.30664 | 1.53824 | 1.66563 | 1.75288 | 1.81893 | 1.87192 | 1.91608 | 1.95387 | 1.98687 |
| | 12 | 1.41311 | 1.67078 | 1.81359 | 1.91188 | 1.98654 | 2.04661 | 2.09677 | 2.13979 | 2.17741 |
| | 13 | 1.59718 | 1.91012 | 2.06976 | 2.18726 | 2.27692 | 2.34930 | 2.40993 | 2.46204 | 2.50771 |
| | 14 | 1.95975 | 2.35600 | 2.58141 | 2.73898 | 2.86002 | 2.95823 | 3.04081 | 3.11204 | 3.17464 |
| | 15 | 3.00609 | 3.71100 | 4.10706 | 4.39503 | 4.61835 | 4.80079 | 4.95503 | 5.08862 | 5.20646 |

Each entry is the D-value satisfying $Pr(Y_{j,k} \geq \max_{1 \leq r \leq k} Y_{j,r} - D) = P^*$ where $Y_{j,1}, Y_{j,2}, \ldots, Y_{j,k}$ are i.i.d. having the same distribution as the jth order statistic based on n independent observations from the logistic distribution with density $e^{-x}/(1+e^{-x})^2$.

Table 1 (cont'd.)

$P^* = 0.99$

| n | j | k = 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 1 | 4.81237 | 5.25863 | 5.48659 | 5.63611 | 5.74561 | 5.83124 | 5.90114 | 5.95997 | 6.01060 |
| | 2 | 3.25538 | 3.57839 | 3.75450 | 3.87362 | 3.96317 | 4.03460 | 4.09385 | 4.14436 | 4.18832 |
| | 3 | 2.97576 | 3.23278 | 3.46987 | 3.59224 | 3.68545 | 3.76057 | 3.82340 | 3.87734 | 3.92457 |
| | 4 | 3.25538 | 3.54300 | 3.86424 | 4.01931 | 4.13964 | 4.23557 | 4.31716 | 4.38758 | 4.44952 |
| | 5 | 4.81237 | 5.51682 | 5.91272 | 6.20062 | 6.42339 | 6.61636 | 6.76151 | 6.89409 | 7.01131 |
| 7 | 1 | 4.74559 | 5.17950 | 5.39936 | 5.54286 | 5.64637 | 5.72799 | 5.79366 | 5.84910 | 5.89653 |
| | 2 | 3.08339 | 3.33131 | 3.54048 | 3.64737 | 3.72707 | 3.79023 | 3.84234 | 3.88656 | 3.92439 |
| | 3 | 2.62622 | 2.83934 | 3.03345 | 3.13171 | 3.20581 | 3.26505 | 3.31425 | 3.35630 | 3.39291 |
| | 4 | 2.51140 | 2.77461 | 2.92102 | 3.02189 | 3.09854 | 3.16025 | 3.21169 | 3.25583 | 3.29443 |
| | 5 | 2.52622 | 2.90398 | 3.08419 | 3.19907 | 3.28636 | 3.35804 | 3.41756 | 3.46696 | 3.51396 |
| | 6 | 3.08339 | 3.45274 | 3.68017 | 3.83290 | 3.95059 | 4.04630 | 4.12694 | 4.19658 | 4.25787 |
| | 7 | 4.74559 | 5.43958 | 5.84533 | 6.13315 | 6.35638 | 6.53876 | 6.69295 | 6.62651 | 6.94432 |
| 9 | 1 | 4.71022 | 5.13806 | 5.35391 | 5.49345 | 5.59469 | 5.67326 | 5.73760 | 5.79034 | 5.83603 |
| | 2 | 3.00133 | 3.23770 | 3.43936 | 3.54365 | 3.61574 | 3.67507 | 3.72786 | 3.76515 | 3.80056 |
| | 3 | 2.46221 | 2.72503 | 2.85666 | 2.94576 | 3.01259 | 3.06578 | 3.10981 | 3.14729 | 3.17994 |
| | 4 | 2.27139 | 2.43972 | 2.62510 | 2.71174 | 2.77538 | 2.82711 | 2.87011 | 2.90684 | 2.93884 |
| | 5 | 2.21179 | 2.44126 | 2.56856 | 2.65611 | 2.72255 | 2.77592 | 2.82045 | 2.85859 | 2.89191 |
| | 6 | 2.27139 | 2.51570 | 2.65414 | 2.74927 | 2.82190 | 2.88029 | 2.92924 | 2.97129 | 3.00810 |
| | 7 | 2.48221 | 2.75585 | 2.92655 | 3.03856 | 3.12441 | 3.19392 | 3.25229 | 3.30256 | 3.34668 |
| | 8 | 3.00133 | 3.37731 | 3.59312 | 3.74486 | 3.86137 | 3.95707 | 4.03730 | 4.10661 | 4.16753 |
| | 9 | 4.71022 | 5.41400 | 5.80967 | 6.09746 | 6.32357 | 6.50303 | 6.65721 | 6.79076 | 6.91856 |
| 11 | 1 | 4.68832 | 5.11234 | 5.32545 | 5.46283 | 5.56230 | 5.63936 | 5.70177 | 5.75394 | 5.79857 |
| | 2 | 2.95402 | 3.23312 | 3.38045 | 3.47842 | 3.55092 | 3.60803 | 3.65490 | 3.69456 | 3.72869 |
| | 3 | 2.40357 | 2.63589 | 2.76037 | 2.84526 | 2.90826 | 2.95826 | 2.99555 | 3.03462 | 3.06593 |
| | 4 | 2.15241 | 2.35454 | 2.48015 | 2.55870 | 2.61775 | 2.66485 | 2.70389 | 2.73716 | 2.76615 |
| | 5 | 2.03339 | 2.23936 | 2.35108 | 2.42797 | 2.48504 | 2.53252 | 2.57117 | 2.60418 | 2.63296 |
| | 6 | 1.99841 | 2.21433 | 2.31837 | 2.39669 | 2.45607 | 2.50374 | 2.54347 | 2.57749 | 2.60719 |
| | 7 | 2.03339 | 2.24954 | 2.36984 | 2.45285 | 2.51631 | 2.56686 | 2.60935 | 2.64580 | 2.67768 |
| | 8 | 2.15241 | 2.33929 | 2.52247 | 2.61485 | 2.68539 | 2.74236 | 2.79008 | 2.83110 | 2.86705 |
| | 9 | 2.40357 | 2.63251 | 2.84099 | 2.95152 | 3.03631 | 3.10501 | 3.16273 | 3.21246 | 3.25613 |
| | 10 | 2.95402 | 3.32750 | 3.54243 | 3.69363 | 3.81027 | 3.90519 | 3.98520 | 4.05434 | 4.11520 |
| | 11 | 4.68832 | 5.33137 | 5.78760 | 6.07537 | 6.29356 | 6.48091 | 6.63509 | 6.76864 | 6.88644 |
| 13 | 1 | 4.67342 | 5.03484 | 5.30639 | 5.44201 | 5.54027 | 5.61629 | 5.67781 | 5.72917 | 5.77318 |
| | 2 | 2.92259 | 3.13736 | 3.34187 | 3.43775 | 3.50848 | 3.56414 | 3.60976 | 3.64826 | 3.68146 |
| | 3 | 2.35421 | 2.57991 | 2.70095 | 2.78227 | 2.84290 | 2.89092 | 2.93052 | 2.96414 | 2.99318 |
| | 4 | 2.08124 | 2.23433 | 2.39339 | 2.46843 | 2.52425 | 2.56867 | 2.60542 | 2.63669 | 2.66334 |
| | 5 | 1.93448 | 2.12596 | 2.23062 | 2.30185 | 2.35548 | 2.39829 | 2.43381 | 2.46410 | 2.49045 |
| | 6 | 1.85975 | 2.04693 | 2.14988 | 2.22026 | 2.27341 | 2.31596 | 2.35135 | 2.38157 | 2.41731 |
| | 7 | 1.83656 | 2.02498 | 2.12914 | 2.20061 | 2.25476 | 2.29825 | 2.33440 | 2.36537 | 2.39240 |
| | 8 | 1.85975 | 2.05451 | 2.16298 | 2.23763 | 2.29434 | 2.33993 | 2.37800 | 2.41062 | 2.43913 |
| | 9 | 1.93448 | 2.14256 | 2.25913 | 2.33966 | 2.40102 | 2.45047 | 2.49193 | 2.52734 | 2.55842 |
| | 10 | 2.08124 | 2.31325 | 2.44395 | 2.53474 | 2.60413 | 2.66021 | 2.70721 | 2.74763 | 2.78307 |
| | 11 | 2.35421 | 2.63031 | 2.78728 | 2.89691 | 2.98106 | 3.04927 | 3.10659 | 3.15600 | 3.19940 |
| | 12 | 2.92259 | 3.23488 | 3.50926 | 3.66012 | 3.77652 | 3.87127 | 3.95114 | 4.02017 | 4.08094 |
| | 13 | 4.67342 | 5.35696 | 5.77259 | 6.06034 | 6.28353 | 6.46588 | 6.62005 | 6.75359 | 6.87139 |
| 15 | 1 | 4.66253 | 5.03216 | 5.29206 | 5.42692 | 5.52430 | 5.59958 | 5.66045 | 5.71123 | 5.75462 |
| | 2 | [illegible] | 3.17212 | 3.31463 | 3.40897 | 3.47853 | 3.53317 | 3.57791 | 3.61562 | 3.64813 |
| | 3 | 2.32320 | 2.54149 | 2.65978 | 2.73938 | 2.79839 | 2.84477 | 2.88321 | 2.91578 | 2.94335 |
| | 4 | 2.03395 | 2.23055 | 2.33693 | 2.40965 | 2.46230 | 2.50497 | 2.54022 | 2.57018 | 2.59517 |
| | 5 | 1.87138 | 2.05417 | 2.15394 | 2.22156 | 2.27252 | 2.31305 | 2.34663 | 2.37523 | 2.40008 |
| | 6 | 1.77545 | 1.95168 | 2.04815 | 2.11388 | 2.16341 | 2.20297 | 2.23581 | 2.26383 | 2.28821 |
| | 7 | 1.72455 | 1.83927 | 1.99767 | 2.05893 | 2.10824 | 2.14770 | 2.18052 | 2.20855 | 2.23298 |
| | 8 | 1.70954 | 1.83319 | 1.97967 | 2.04577 | 2.09535 | 2.13601 | 2.16945 | 2.19807 | 2.22303 |
| | 9 | 1.72455 | 1.91386 | 2.00333 | 2.07175 | 2.12366 | 2.16537 | 2.20017 | 2.22997 | 2.25600 |
| | 10 | 1.77545 | 1.95373 | 2.05872 | 2.14115 | 2.19625 | 2.24060 | 2.27765 | 2.30942 | 2.33721 |
| | 11 | 1.87138 | 2.07438 | 2.18845 | 2.26742 | 2.32764 | 2.37622 | 2.41686 | 2.45178 | 2.48235 |
| | 12 | 2.03395 | 2.25269 | 2.39179 | 2.48153 | 2.55218 | 2.60567 | 2.65221 | 2.69224 | 2.72735 |
| | 13 | 2.32120 | 2.59439 | 2.75042 | 2.85345 | 2.94317 | 3.01106 | 3.06812 | 3.11731 | 3.16052 |
| | 14 | [illegible] | 3.27135 | 3.48585 | 3.63649 | 3.75273 | 3.84736 | 3.92714 | 3.99609 | 4.05680 |
| | 15 | 4.66253 | 5.35613 | 5.76172 | 6.04946 | 6.27254 | 6.45499 | 6.60915 | 6.74270 | 6.86049 |

Each entry is the D-value satisfying $Pr(Y_{j,k} \geq \max_{1 \leq r \leq k} Y_{j,r} - D) = P^*$ where $Y_{j,1}, Y_{j,2}, \ldots, Y_{j,k}$ are i.i.d. having the same distribution as the jth order statistic based on n independent observations from the logistic distribution with density $e^{-x}/(1+e^{-x})^2$.

not exceed $2\cdot 10^{-13}$. This was always achieved for $u_1 > -35$ and $u_2 < 15$. For evaluating $A_2(D)$, the interval is subdivided into an even number of intervals of length $h = 0\cdot 125$ and Simpson's rule is used first with interval length h and then 2h so as to yield the evaluations $B_h(D)$ and $B_{2h}(D)$, respectively. Now a corrected value, $I_h(D)$ is obtained by setting $I_h(D) = B_h(D) + \varepsilon_h(D)$, where $\varepsilon_h(D) = [B_h(D) - B_{2h}(D)]/15$. Then $|\varepsilon_h(D)|$ usually gives a good estimate of the error $|A_2(D) - I_h(D)|$ and in our computations $|\varepsilon_h(D)| \lesssim 10^{-10}$. Since the truncation error $\lesssim 2\cdot 10^{-13}$, the total error $= |A(D) - I_h(D)| \lesssim 10^{-10}$.

Now, to solve for D from (3.23), we use a modified secant method. Let $D_{i1}$ and $D_{i2}$ be two trial values at the stage i, $i = 1,2,\ldots$ . Define

$$D_{i3} = D_{i1} - A(D_{i1})\ [D_{i1} - D_{i2}]/[A(D_{i1}) - A(D_{i2})]. \tag{3.25}$$

For the stage i+1, $D_{i3}$ replaces $D_{i1}$ if $|A(D_{i1}) - P^*| > |A(D_{i2}) - P^*|$ and replaces $D_{i2}$ otherwise. This process was continued until the stage m, where m is the smallest positive integer for which either $|D_{m3} - D_{m1}| < 10^{-5}$ or $|D_{m3} - D_{m2}| < 10^{-5}$. $D_{m3}$ is taken as the solution of (3.23). It could be safely said that the table values are correct to five places of decimals.

It should be pointed out that Table 1 will be of interest in other problems such as testing of hypotheses based on order statistics from the logistic distribution. As one can see, the table value D is the upper 100(1-P*) percentage point of the distribution of $\max\limits_{1\le r\le k-1} Z_{r,k}, Z_{r,k} =$

$(Y_{j,r} - Y_{j,k})$, where $Y_{j,1}, \ldots, Y_{j,k}$ are i.i.d. having the distribution of the jth order statistic based on n independent observations from the logistic distribution.

We now discuss the asymptotic evaluation of the probability of a correct selection using the procedure $R_5$. We state and prove the following theorem.

Theorem 3.4. *If $F_{[k]} \underset{t_\alpha}{\prec} G$, $F_{[k]}(G)$ has a differentiable density $f_{[k]}(g)$ in a neighborhood of the $\alpha$-quantile $\xi_{\alpha[k]}(\eta_\alpha)$ and $f_{[k]}(\xi_{\alpha[k]}) \neq 0$ $(g(\eta_\alpha) \neq 0)$, then in our previous notation (see Theorem 3.3)*

$$\lim_{n\to\infty} \Pr(X_{j,k} \geq \max_{1\leq r\leq k} X_{j,r} - D)$$

$$(3.26) \qquad = \int_{-\infty}^{\infty} \Phi^{k-1}[x + D\, f_{[k]}(\xi_{\alpha[k]})(n/\alpha\bar{\alpha})^{1/2}] \; d\Phi(x)$$

$$\geq \int_{-\infty}^{\infty} \Phi^{k-1}[x + D\, g(\eta_\alpha)(n/\alpha\bar{\alpha})^{1/2}] \; d\Phi(x),$$

*where $j/n \to \alpha$ as $n \to \infty$, $\bar{\alpha} = 1-\alpha$ and $\Phi(\cdot)$ is the* cdf *of the standard normal variate.*

Proof. $\Pr(X_{j,k} \geq \max_{1\leq r\leq k} X_{j,r} - D)$

$$(3.27) \qquad = \Pr\{(X_{j,k} - \xi_{\alpha[k]})\, f_{[k]}(\xi_{\alpha[k]})(n/\alpha\bar{\alpha})^{1/2}$$

$$\geq [\max_{1\leq r\leq k} (X_{j,r} - \xi_{\alpha[k]}) - D]\, f_{[k]}(\xi_{\alpha[k]})(n/\alpha\bar{\alpha})^{1/2}\}$$

$$\simeq \int_{-\infty}^{\infty} \Phi^{k-1}[x + D\, f_{[k]}(\xi_{\alpha[k]})(n/\alpha\bar{\alpha})^{1/2}] \; d\Phi(x),$$

since $X_{j,k}$ is asymptotically normally distributed with mean $\xi_{\alpha[k]}$ and variance $\alpha\bar{\alpha}/n\, f^2_{[k]}(\xi_{\alpha[k]})$. (Note that $a_n \simeq b_n$ means

$\lim_{n\to\infty} a_n/b_n = 1$).

The second part of (3.26) follows from the fact that $F_{[k]} \underset{t_\alpha}{<} G$ implies that $f_{[k]}(\xi_{\alpha[k]}) \geq g(\eta_\alpha)$.

For $G(x) = (1+e^{-x})^{-1}$, $g(\eta_\alpha) = \alpha\bar{\alpha}$. Thus, for evaluating approximate value of D, we set

$$(3.28) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}[x + D(n\alpha\bar{\alpha})^{1/2}] \; d\Phi(x) = P^*.$$

The D-value satisfying (3.28) can be obtained from the tables of Gupta, Nagel and Panchapakesan [55], who have tabulated H-value satisfying

$$(3.29) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}[(\rho^{1/2}x + H)/(1-\rho)^{1/2}] \; d\Phi(x) = P^*$$

for selected values of k, $\rho$ and P*. For $\rho = 1/2$, $D = H(2/n\alpha\bar{\alpha})^{1/2}$. The exact and asymptotic values of D are given in Table 2 for n = 25, j = 8, 13, 18 and P* = 0.90, 0.95.

3.5. A general partial order relation and related results. We now define a general partial order relation through a class of real valued functions. The star and the tail orderings can be obtained as special cases. This general order relation also throws additional light on a lemma of Gupta [49].

Let $\mathcal{A} = \{h(x)\}$ be a class of real valued functions h(x) defined on the real line. Let F and G be distributions on the real line such that F(0) = G(0). We say that F is $\mathcal{A}$-ordered with respect to G ($F \underset{\mathcal{A}}{<} G$) if $G^{-1}F(h(x)) \geq h(G^{-1}F(x))$ for all $h \in \mathcal{A}$ and all x on the support of F. It is easy to verify that the order relation is reflexive and transitive. It can be seen immediately from the definition

Table 2. Exact (upper) and approximate (lower) values of D for n = 25.

| | j | k: 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $P^*=0.90$ | 8 | 0.78743 | 0.95314 | 1.05532 | 1.11666 | 1.16214 | 1.19803 | 1.22755 | 1.25254 | 1.27415 |
| | | 0.78540 | 0.96643 | 1.06233 | 1.12656 | 1.17430 | 1.21211 | 1.24330 | 1.26971 | 1.29257 |
| | 13 | 0.72315 | 0.89068 | 0.97981 | 1.03965 | 0.08430 | 1.11973 | 1.14898 | 1.17382 | 1.19537 |
| | | 0.72498 | 0.89209 | 0.98061 | 1.03990 | 1.08397 | 1.11887 | 1.14766 | 1.17204 | 1.19314 |
| | 18 | 0.78749 | 0.97773 | 1.08046 | 1.15006 | 1.20236 | 1.24406 | 1.27865 | 1.30813 | 1.33379 |
| $P^*=0.95$ | 8 | 1.01374 | 1.17603 | 1.26244 | 1.32037 | 1.36354 | 1.39774 | 1.42593 | 1.44985 | 1.47057 |
| | | 1.00804 | 1.17436 | 1.26371 | 1.32389 | 1.36893 | 1.40472 | 1.43426 | 1.45938 | 1.48120 |
| | 13 | 0.92977 | 1.08443 | 1.16782 | 1.22421 | 1.26648 | 1.30014 | 1.32801 | 1.35173 | 1.37234 |
| | | 0.93050 | 1.08402 | 1.16650 | 1.22205 | 1.26363 | 1.29666 | 1.32393 | 1.34712 | 1.36720 |
| | 18 | 1.01373 | 1.19014 | 1.28652 | 1.35226 | 1.40186 | 1.44155 | 1.47455 | 1.50274 | 1.52731 |

The approximate values for j = 18 are same as those for j = 5.

that, if $\mathcal{A} = \{ax,\ a \geq 1\}$ and $F(0) = G(0) = 0$, we get the star-ordering. The tail-ordering is obtained by taking $\mathcal{A} = \{x+b,\ b \geq 0\}$ and $F(0) = G(0) = \frac{1}{2}$. We now prove a lemma which is useful in obtaining a lower bound on the probability of a correct selection using a general class of procedures described below.

Lemma 3.2. *If $F \underset{\mathcal{A}}{\prec} G$, then, for any positive integer $m$,*

$$\int_{-\infty}^{\infty} F^m(h(x)) \quad dF(x) \geq \int_{-\infty}^{\infty} G^m(h(x)) \quad dG(x) \tag{3.30}$$

*for all $h \in \mathcal{A}$.*

Proof. Let $X_1, X_2, \ldots, X_{m+1}$ $(Y_1, Y_2, \ldots, Y_{m+1})$ be independent and identically distributed, each with distribution $F$ $(G)$. Then, the left hand member of (3.30) is equal to

$$\Pr(h(X_{m+1}) \geq X_j,\ j = 1,\ldots,m)$$

$$\geq \Pr(\varphi(h(X_{m+1})) \geq \varphi(X_j),\ j=1,\ldots,m), \text{ since } \varphi \equiv G^{-1}F \text{ is nondecreasing}$$

$$\geq \Pr(h(\varphi(X_{m+1})) \geq \varphi(X_j),\ \ j=1,\ldots,m), \text{ since } F \underset{\mathcal{A}}{\prec} G,$$

$$= \Pr(h(Y_{m+1}) \geq Y_j,\ j=1,\ldots,m), \text{ since } \varphi(X_i) \underset{st}{=} Y_i,\ i = 1,\ldots, m+1,$$

$$= \int_{-\infty}^{\infty} G^m(h(x)) \quad dG(x).$$

Remark. In selection problems we will have $\Pr(h(X_{t+1}) \geq \max_{1 \leq j \leq t+1} X_j)$. In this case, the proof of the lemma is valid if we assume that $h(x) \geq x$ on the support of $F$.

Now, we wish to point out the relevance of a lemma of Gupta [49], which is stated below essentially in its original form.

*Lemma 3.3. Let X be a random variable having the distribution $F_\lambda$. Let $h_b(x)$ be a class of functions and suppose there exists a distribution function F such that $h_b(g_\lambda(x)) \geq g_\lambda(h_b(x))$ for all $\lambda$ and all x, where $g_\lambda(x)$ is defined by $F_\lambda(g_\lambda(x)) = F(x)$ for all x. Then, for any* $t > 0$,

$$\int F_\lambda^t\ (h_b(x))\quad dF_\lambda(x) \geq \int F^t(h_b(x))\quad dF(x). \tag{3.31}$$

Though this lemma was not stated in terms of a partial order relation, it can be seen that the hypothesis of the lemma amounts to saying $F_\lambda \underset{\mathscr{H}}{\gtrsim} F$, where $\mathscr{H} = \{h_b\}$. It should also be pointed out here that Lemma 3.2 can be proved for any $m > 0$ but we have restricted ourselves to a positive integer m which is the case in selection problems.

Now we describe a selection problem for a family of $\mathscr{H}$-ordered distributions. Let $\pi_1, \pi_2, \ldots, \pi_k$ be k populations and $F_i$ be the distribution function associated with $\pi_i$, $i = 1,2,\ldots, k$. We assume that there is one among the k populations, denoted by $F_{[k]}$, which is stochastically larger than any of the others. We also assume that $F_i \underset{\mathscr{H}}{\lesssim} G$, $i = 1,2,\ldots, k$, where G is a specified continuous distribution and $\mathscr{H} = \{h\}$ is a class of continuous distributions satisfying the properties stated in (2.1) for x on the support of G. Let $T_i = T(X_{i1}, X_{i2},\ldots, X_{in})$, $i = 1,2,\ldots, k$, and $T = T(Y_1, Y_2,\ldots, Y_n)$ be statistics based on independent observations $X_{i1}, X_{i2},\ldots, X_{in}$ from $F_i$, $i = 1,2,\ldots, k$, and $Y_1, Y_2,\ldots, Y_n$ from G. It is assumed that the $T_i$ and T preserve the stochastic and $\mathscr{H}$-ordering of the parent distributions. Then, for selecting a subset containing $F_{[k]}$, we propose the selection rule $R_6$: select $\pi_i$, if and only if,

(3.32) $$h(T_i) \geq \max_{1 \leq r \leq k} T_r,$$

where a function $h \in \mathscr{A}$ is chosen so that the P*-condition is satisfied.

<u>Theorem 3.5</u>. <u>Let</u> $F_i(x) \geq F_{[k]}(x)$ <u>and</u> $F_i \underset{\mathscr{A}}{\prec} G$, $i = 1,2,\ldots,k$. <u>Then</u>

(3.33) $$\inf_{\Omega} P(CS|R_6) = \int G_T^{k-1}(h(x))\, d\, G_T(x),$$

<u>where</u> $G_T$ <u>is</u> <u>the</u> cdf <u>of</u> T.

<u>Proof</u>. Since the $T_i$ preserve the stochastic ordering, we have

(3.34) $$P(CS|R_6) \geq \int F_{T[k]}^{k-1}(h(x))\, d\, F_{T[k]}(x),$$

where $F_{T[k]}(x)$ is the cdf of the $T_i$ from $F_{[k]}$. The statement of the theorem follows immediately because of Lemma 3.3 and the $\mathscr{A}$-ordering property of the $T_i$ and T.

In §2 we discussed the class of procedures $R_h$. A natural question that arises is how to choose a class $\mathscr{A}$ for a particular selection problem. Of course, it should depend on any additional knowledge we may have about the family of distributions we are dealing with. In this context a partial answer is provided by the selection problem discussed above.

We now discuss some life testing problems and sampling plans for the IFR (DFR) distributions.

4. Life testing problems and results relating to restricted families of distributions. The life testing problems are closely related to reliability problems. They are not only relevant in problems concerning the length of life of items such as electronic components and ball bearings but also in biological problems concerning the life of individuals such as human beings. Hence such problems are relevant to both reliability and biometry.

Since the life test data are censored long before all items on test have actually failed, the need arose to develop new estimation and decision procedures based on such censored data. Methods of estimation and testing procedures based on order statistics naturally play an important role in this area for the obvious reason that the life test experiments yield observations which are ordered as they arise.

In the early investigations the underlying distribution for the life characteristic was assumed to be the one and the two parameter exponential [35], [36]. While such a model is very nice to handle mathematically and is surprisingly applicable in many situations where the failure rate can be assumed to be constant over the time span of interest, there are several circumstances under which the assumption of a constant failure rate becomes untenable. In some cases [45], [82], the normal and lognormal distributions have been used. Some of the other widely used life length models are the Weibull [69], gamma [24], and the extreme value distributions [43]. In the case of the Weibull distribution in its two-parameter form [$F(x) = 1 - \exp(-(x/\theta)^{\beta})$, $x > 0$,

$\theta,\beta > 0]$, the failure rate is of the form $r(x) = \Lambda\, x^{\beta-1}$, where $\Lambda > 0$. As one can immediately see, the Weibull distribution provides a desirable flexibility in describing a life length model, because, for different choices of $\beta$, we have a failure rate which is constant for $\beta = 1$, increasing for $\beta > 1$ and decreasing for $\beta < 1$. Considerable work has also been done under a quasi-parametric setup. Under such a setup the distribution, though not functionally known, is assumed to have an increasing (decreasing) failure rate or an increasing (decreasing) failure rate on the average. The use of the exponential model is still borne out by the fact that it serves as a "boundary" for the IFR and DFR distributions and hence is employed to obtain useful bounds. In this section we describe some of the important results relating to both parametric and quasi-parametric cases. The general literature on life testing has grown enormously and the reader is referred to the bibliographical information available in [27], [41] and [78].

4.1. Exponential model. The estimation problem for the life test data under the exponential model was first considered by Epstein and Sobel [35]. They considered what is now usually referred to as Type II censoring. A given number, n, of items are put on test and the experiment is terminated at the rth failure. In a subsequent paper [36], the same authors have obtained some theorems relevant to life testing under the two-parameter exponential model with density

$$(4.1)\qquad f(x;\,\theta,A) = \begin{cases} (1/\theta)\exp\{-(x-A)/\theta\}, & x \geq A, \\ 0\qquad\qquad\qquad\qquad\quad, & \text{otherwise}, \end{cases}$$

where $\theta > 0$ is unknown and $A \geq 0$. The parameter $\theta$ is the mean-time-between-failures and A is usually referred to as the guarantee time. Let us denote the r failure times by $X_{1,n} < X_{2,n} < \ldots < X_{r,n}$. Then, the maximum likelihood estimator of $\theta$, assuming A to be known, is shown to be

$$\hat{\theta}_{r,n} = r^{-1}\{\sum_{i=1}^{r} (X_{i,n}-A) + (n-r)(X_{r,n}-A)\}. \tag{4.2}$$

Define

$$\begin{aligned} W_i &= (n-i+1)(X_{i,n}-X_{i-1,n}), \quad i = 1,2,\ldots, r, \\ X_{o,n} &= A, \\ T &= \sum_{i=1}^{r} (X_{i,n}-A) + (n-r)(X_{r,n}-A). \end{aligned} \tag{4.3}$$

The statistic T is called the total life statistic. The most important results relating to the $W_i$ and T are contained in the following theorem.

Theorem 4.1. For $1 \leq r \leq n$, the random variables $W_i$, $i = 1,2,\ldots, r$, are independent and identically distributed with the density function in (4.1) with $A = 0$. Further, $2T/\theta$ has a chi-square distribution with 2r degrees of freedom.

Epstein and Sobel [36] have also obtained several theorems relevant to estimation for various cases. It should be pointed out that the distribution of T readily allows one to construct confidence intervals and tests of significance for $\theta$. For these and other estimation problems the reader is referred, among others, to Epstein [32], [34], Bartholomew [18] and Epstein and Tsao [38]. In the subsequent sections, we discuss some acceptance sampling plans under parametric and quasi-parametric models.

4.2. Life test sampling plans: parametric case. For the exponential model with A = 0, Epstein [31] has discussed sampling plans to meet prescribed requirements on the producer's and consumer's risks. He considered life test data censored after a fixed number r of failures as well as data censored at time $t = \min(X_{r,n}, T_o)$, where $X_{r,n}$ is the time at which the rth failure occurs and $T_o$ is a pre-assigned time.

Suppose we put n units on life test and terminate at the rth failure. If $\theta_o$ is the acceptable (high) mean life and $\theta_1$ is some specified unacceptable (low) mean life, then the lot is accepted or rejected according as $\hat{\theta}_{r,n} > C$ or $\hat{\theta}_{r,n} \leq C$, where $\hat{\theta}_{r,n}$ is the maximum likelihood estimator in (4.2). It is required that the Operating Characteristic (OC) function $L(\theta)$ satisfies

$$L(\theta_o) = 1-\alpha \text{ and } L(\theta_1) \leq \beta. \tag{4.4}$$

The constant $C = (2r)^{-1}\theta_o\, \chi^2_{1-\alpha}(2r)$, where r is the smallest integer such that

$$\chi^2_{1-\alpha}(2r)[\chi^2_{\beta}(2r)]^{-1} \geq \theta_1\theta_o^{-1} \tag{4.5}$$

and $\chi^2_{\gamma}(m)$ is the upper $100\gamma\%$ point of the chi-square distribution with m degrees of freedom.

On the other hand, if we consider terminating the experiment at time $t = \min(X_{r,n}, T_o)$ as explained earlier, we accept or reject the lot according as $t = T_o$ or $t < T_o$. The results concerning this procedure were obtained by Epstein [31] and summarized by him later in [33]. For a pre-assigned $T_o$, it is required to find r and n such that the OC-function satisfies the conditions that $L(\theta_0) = 1-\alpha$ and $L(\theta_1) \leq \beta$. It has been shown that the appropriate r is precisely the same r used in the first case, namely, the smallest integer such that (4.5) is true. A good approximation for n, when $\theta_0 \geq 3\ T_0$, is given by

$$n = [r\{1-\exp(-C^{-1}T_0)\}^{-1}], \tag{4.6}$$

where $C = (2r)^{-1}\theta_0\ \chi^2_{1-\alpha}(2r)$.

In the above cases, we can also use a replacement model in which the failed items are replaced by new ones. Necessary modifications in the results can be made in this case and these will not be discussed here. Also we do not discuss the sequential life tests in the exponential case studied by Epstein and Sobel [37].

Now we consider some sampling plans based on data censored at a pre-assigned time T obtained from a life distribution F. Let $\theta$ denote a parameter of this distribution. It could be, for example, the mean or a given percentile. We consider any value $\theta$ to be acceptable if $\theta \geq \theta_0$, where $\theta_0$ is some specified value. The sampling plan accepts the lot if the number of failures is less than or equal to an acceptance number, c. Given the desired confidence level, P*, the parameter goal, $\theta_0$, and the test time T, we wish to determine the smallest n such that the consumer's risk in adopting the sampling plan $(n,c,T,\theta_0,P^*)$ does not exceed 1-P* whatever $\theta$ may be in the interval $[0,\theta_0]$.

It should be noted that these life tests can be terminated prior to time T with the rejection as the result. In fact, we can terminate the experiment at the smaller of the two times T and $T_{c+1}$, where $T_{c+1}$

is the time to the (c+1)st failure. In this case, we accept the lot at the end of time T if and only if $T_{c+1} > T$. Thus, letting $L_n(p)$ denote the probability of acceptance when n items are put on test and letting $p = F(t;\theta)$, we have

$$(4.7) \qquad \begin{aligned} L_n(p) &= \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i} \\ &= 1 - I_p(c+1,\ n-c), \end{aligned}$$

where $I_p(c+1,\ n-c)$ is the incomplete beta function. If the family of distributions $\{F(t;\theta)\}$ is stochastically increasing in $\theta$, then $F(t;\theta) \geq F(t;\theta_o)$ for $\theta \leq \theta_o$. Since $L_n(p)$ decreases in p, in order to satisfy the requirement of the consumer's risk, we choose the smallest n such that

$$(4.8) \qquad I_{p_o}(c+1,\ n-c) \geq P^*,$$

where $p_o = F(T;\theta_o)$.

The above procedure has been studied by Gupta and Groll [50] in the case of the two-parameter gamma distribution with density

$$(4.9) \qquad f_r(t;\theta) = \{\theta\, \Gamma(r)\}^{-1} \exp\{-t/\theta\}\ (t/\theta)^{r-1}, \quad 0 < t < \infty,\ \theta > 0,\ r > 0.$$

In this case, since $\theta$ is a scale parameter the acceptable quality is stipulated by specifying $T/r\theta_o$ and we want the consumer's risk to be bounded by 1-P* for $\theta \leq \theta_o$. However, (4.8) is still valid. Gupta and Groll have tabulated the minimum sample size n satisfying (4.8) for selected values of P*, r, $T/r\theta_o$ and c. Similar acceptance procedures have been investigated by Gupta [45] for normal and lognormal distributions. Our main interest is the quasi-parameter case which is described below.

4.3. Life test sampling plans: distribution-free and quasi-parameter case. In our discussion of sampling plans based on life test data we have so far assumed a complete knowledge of the functional form $F(t;\theta)$ except for the specific value of the parameter $\theta$. Now we consider truncated life test sampling plans investigated by Barlow and Gupta [ 5]. We first assume that, though the functional form of $F(t;\theta)$ is not known, there are known bounds on $F(t;\theta)$. To be specific, let us assume that $F(t;\theta) \geq b(t;\theta)$ for $t \geq 0$ where $b(t;\theta)$ is a known function, decreasing in $\theta$. Since $L_n(p)$ in (4.7) is a decreasing function of p, we have

$$L_n[F(T;\theta)] \leq L_n[b(T;\theta)] \leq L_n[b(T;\theta_o)] \tag{4.10}$$

for $\theta \leq \theta_o$. Thus the sampling plans of the type described in the parametric case are obtained by choosing the smallest positive integer n satisfying

$$L_n[b(T,\theta_o)] \leq 1\text{-}P^*, \tag{4.11}$$

where $c, \theta_o$ and T are fixed. If $b(t;\theta)$ is a sharp bound on $F(t;\theta)$, then it is nondecreasing in $\theta$ since $F(t;\theta)$ is nondecreasing in $\theta$. If $\theta$ is the mean, then $b(t;\theta) = b(t/\theta;1)$ since the mean is a scale parameter for a positive random variable. In this case, $b(t;\theta)$ is decreasing in $\theta$ and the consumer's risk is controlled for $\theta \geq \theta_o$. A few known examples of the bound $b(t;\theta)$ are given below.

Example 1. (Markov inequality). Let $\theta$ be the mean of a nonnegative random variable. Then

$$(4.12) \qquad F(t;\theta) \geq b(t;\theta) = \begin{cases} 0 , & t < \theta, \\ 1-\theta/t, & t \geq \theta. \end{cases}$$

A sampling plan based on Markov's inequality would afford protection over the class of all distributions on the positive axis. However, as one would expect, the bound is quite wide. Also we will need the test time T to exceed the goal mean life.

Example 2. (Unimodal density). Suppose that the density f of a nonnegative random variable is unimodal with unknown mode and first moment $\theta$. Then, we have from Barlow and Marshall [ 8]

$$(4.13) \qquad F(t;\theta) \geq b(t;\theta) = \begin{cases} 0 , & 0 \leq t \leq \theta, \\ 2-2\theta t^{-1}, & \theta \leq t \leq 3\theta/2, \\ 1-\theta(2t)^{-1}, & t \geq 3\theta/2. \end{cases}$$

As we can readily see, the bound in this case is a slight improvement over Example 1.

Example 3. ($PF_2$ density). Let f be a $PF_2$ (Polya frequency of order 2) density with mean $\theta$. That is to say, log $f(t;\theta)$ is concave in t when $f(t;\theta) > 0$. In this case we have the following bound which has been tabulated by Barlow and Marshall [ 8].

$$(4.14) \qquad F(t;\theta) \geq b(t;\theta) = \begin{cases} 0, & t \leq \theta, \\ 1-\sup\limits_{m \geq t}(1-e^{-at})(1-e^{-mt})^{-1} , & t \geq \theta, \end{cases}$$

where a is determined by

$$\theta = (1-e^{-am})^{-1} \int_0^m ax\, e^{-ax}\, dx.$$

A sampling plan based on this bound would afford less protection than the IFR sampling plans discussed below because of the fact that a $PF_2$ density f implies an IFR distribution F but not conversely.

Sampling plans for the IFR and DFR distributions. In discussing the sampling plans for the IFR and DFR distributions, we will consider the two cases, namely, (i) $\theta = \mu_r$, the rth moment and (ii) $\theta = \xi_\alpha$, the $\alpha$th quantile. First we discuss the IFR distributions.

Case (i): $\theta = \mu_r$. It is known [10, p. 40] that

$$1-F(t) \le \begin{cases} 1, & t < \mu_r^{1/r}, \\ e^{-wt}, & t \ge \mu_r^{1/r}, \end{cases} \qquad r > 0, \tag{4.15}$$

where w is uniquely determined by

$$\mu_r = r \int_0^t e^{-wx}\, x^{r-1}\, dx. \tag{4.16}$$

Hence, the n required in the sampling plan is the least positive integer which satisfies (4.8) with $p_o = b(T;\mu_{ro})$, where $\mu_{ro}$ denotes the desired rth moment. The solution of (4.8) depends on the fact that $T^r > \mu_o$. A special case of interest is that of the mean (r = 1). In this case, (4.16) reduces to

$$\mu_1 w = 1-e^{-wt}. \tag{4.17}$$

The values of n are tabulated in this case by Barlow and Gupta [ 5] for $T\mu_1^{-1} = 1.1(0.1)\,2.0(0.2)3.0$, c = 0(1)10 and P* = 0.90, 0.95. In all these cases, the table also gives the exact value of the consumer's risk.

Sobel and Tichendorff [90] have tables for the exponential case but their tables are for the ratios $T/\mu_{10} \leq 1$. However, some rough comparisons have shown that the IFR sampling plan would require a longer test time. The main difficulty with the IFR test plan lies not so much in its conservative nature as in the nature of the parameter chosen to represent quality, namely, the mean. Despite the great intuitive appeal, it appears that the mean life is not an appropriate parameter to represent its quality except in the exponential case. The quantiles are more appropriate in the case of IFR distributions.

Case (ii): $\theta = \xi_\alpha$. In this case, we have

$$b(t;\xi_\alpha) = \begin{cases} 0 , & t < \xi_\alpha , \\ 1-(1-\alpha)^{t/\xi_\alpha} , & t \geq \xi_\alpha . \end{cases} \tag{4.18}$$

The sample size required to insure a specified quantile life $\xi_{\alpha o}$ is given by the least n satisfying (4.8) with $p_o = b(T;\xi_{\alpha o})$. The minimum sample sizes have been tabulated by Barlow and Gupta [ 5] for $\alpha = 0.1(0.1)0.9$, $T\xi_\alpha^{-1} = 1.0(0.1)2.0(0.2)3.0$, $c = 0,1,2$ and $P^* = 0.90$, $0.95$. Since the bound in (4.18) is the exponential with $\alpha$th quantile $\xi_\alpha$ when $t \geq \xi_\alpha$, the sampling plans obtained for $T \geq \xi_{\alpha o}$ are the same as those for the exponential case. However, the exponential life test sampling plans can also be developed for $T \leq \xi_{\alpha o}$ which is not possible with the IFR assumption alone.

Similar sampling plans can be developed for the DFR distributions when $\theta$ is either the mean or the $\alpha$th quantile. These are discussed by Barlow and Gupta [ 5] with tables for the quantile case. The appropriate inequalities in these cases are given below.

When $\theta = \mu_r$,

$$(4.19) \qquad b(t;\mu_r) = \begin{cases} 1-\exp\{-t\lambda_r^{-1/r}\}, & t \leq r\,\lambda_r^{1/r}, \\ 1-r^r\lambda_r(et)^{-r}, & t \geq r\,\lambda_r^{1/r}, \end{cases}$$

where $\lambda_r = \mu_r\{\Gamma(r+1)\}^{-1}$. When $\theta = \xi_\alpha$, we have

$$(4.20) \qquad b(t;\xi_\alpha) = \begin{cases} 1-(1-\alpha)^{t/\xi_\alpha}, & t \leq \xi_\alpha, \\ \alpha, & t > \xi_\alpha. \end{cases}$$

In all the above cases, Barlow and Gupta have also given bounds on the OC function.

Barlow and Proschan [13] have investigated acceptance sampling plans in the IFR(DFR) case based on a censored sampling. To be specific, we assume that n items from an IFR population are placed on test and the testing is discontinued after the rth ($r < n$) failure. There is no replacement of failed items. The acceptance sampling plan accepts the lot if and only if $\hat{\theta}_{r,n} \geq (\theta_o)$, where $\hat{\theta}_{r,n}$ is as defined in (4.2), $\theta_o$ is the mean life goal and $C(\theta_o) = \theta_o\,\chi^2_\alpha(2r)/2r$. It is shown in [13] that

$$(4.21) \qquad P_F(\hat{\theta}_{r,n} \geq C(\theta_o)|\theta) \geq P_G(\hat{\theta}_{r,n} \geq C(\theta_o)|\theta)$$

for $\theta \geq \theta_o\,\chi^2_\alpha(2r)/2(n-r+1)$, where G is exponential with mean $\theta$.

Similar plans are considered for the case where we wish to establish that the qth quantile, $\xi_q^o$, exceeds some quantile life goal, $\xi_q^o$ with a Type I error $\leq \alpha$. The decision rule proposed in [13] is to accept the lot if and only if $\hat{\theta}_{r,n} \geq C'(\xi_q^o)$, where $C'(\xi_q^o) = \xi_q^o \chi_\alpha^2(2r)/[-2r \log(1-q)]$.

In using exponential life tests based on censored samples, it is shown in [13] how, depending on the objective chosen, either the producer or the consumer, but certainly not both simultaneously, can be protected within the class of IFR(DFR) distributions. Barlow and Proschan [13] have also discussed similar plans when the failed items are replaced.

5. Order statistics in multiple decision problems and reliability theory, and some important results for restricted families of distributions. In the discussion of selection and ranking problems in §§ 2 and 3 we saw that the procedure is usually defined in terms of the largest or the smallest order statistic. With more complex goals of selection problems, the procedure will depend on appropriate functions of the ordered observations. In some cases (not described in this paper) one may use a linear combination of order statistics. It could be a convex combination or a contrast. In selection problems relating to the quantiles of IFR or IFRA populations, the decision involves the sample quantiles. For a survey of some important results concerning order statistics and their role in problems of statistical inference the reader is referred to a recent paper of the authors [49].

As we pointed out earlier, the life testing problems involve order statistics. Moreover, they provide a natural area for the use of order statistics to a great advantage since the observations are ordered as they arise. Estimation and testing problems involve the use of statistics based on ordered observations such as the total life statistic. Some estimation and testing problems relating to families with monotone failure rate have been discussed in §6 below.

Presently we survey some important results concerning order statistics from the restricted families of distributions. These results are mainly about linear combinations of order statistics from a distribution F which is star-shaped or convex with respect to G with special interest in the case of exponential G. We assume that F and G are absolutely continuous. Throughout this section, unless otherwise stated, $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ will denote order statistics based on n independent observations from F and $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ will denote those from G. Unless the context needs more clarity, for the sake of notational convenience, we will leave out n and write $X_{(i)}$ and $Y_{(i)}$ for $X_{i,n}$ and $Y_{i,n}$, respectively.

If we consider any linear combination $\sum_{i=1}^{r} a_i X_{(i)}$, $1 \leq r \leq n$, it is easy to see that it can be written in the form $\sum_{i=1}^{r} \bar{A}_i (X_{(i)} - X_{(i-1)})$, where $X_{(o)} \equiv 0$ and $\bar{A}_i = \sum_{j=1}^{i} a_j$. The successive differences $D_i = X_{(i)} - X_{(i-1)}$, $i=2,\ldots,n$, are called the <u>spacings</u>. A linear combination of order statistics is therefore a weighted sum of the spacings. The theory of spacings, in general, is an interesting area of research. The literature on spacings that had appeared prior to 1965 has been substantially surveyed by Pyke [85]. Some tests for increasing failure rate based on spacings are discussed in §6. Some of the recent developments in the theory of spacings have been surveyed by Pyke [86], who has also set out some open problems in the asymptotic theory of spacings.

5.1. Inequalities for linear combinations of order statistics from $F$ star-shaped with respect to $G$. The important results summarized here are due to Barlow and Proschan [11].

Theorem 5.1. Let $F \underset{*}{<} G$.

(i) Suppose there exists $k$, $1 \le k \le n$, such that $0 \le \bar{A}_1 \le \bar{A}_2 \le \cdots \le \bar{A}_k \le 1$, and when $k < n$, $\bar{A}_{k+1} = \cdots = \bar{A}_n = 0$. Then

$$F\left(\sum_{i=1}^{n} a_i X_{(i)}\right) \underset{st}{\le} G\left(\sum_{i=1}^{n} a_i Y_{(i)}\right). \tag{5.1}$$

(ii) Suppose $a_i \ge 0$ for $i = 1,2,\ldots, n$ and $a_n \ge 1$. Then

$$F\left(\sum_{i=1}^{n} a_i X_{(i)}\right) \underset{st}{\ge} G\left(\sum_{i=1}^{n} a_i Y_{(i)}\right). \tag{5.2}$$

In (5.1) and (5.2), "st" indicates stochastic inequalities. The above results are useful in obtaining conservative lower and upper tolerance limits, respectively (see Barlow and Proschan [12]). The following theorem relates to the expected values of the order statistics.

Theorem 5.2. Let $F \underset{*}{<} G$. Then $E\, X_{i,n}/E\, Y_{i,n}$ is (i) decreasing in $i$, (ii) increasing in $n$, and (iii) $E\, X_{n-i,n}/E\, Y_{n-i,n}$ is decreasing in $n$.

If $F$ is an IFRA distribution (that is, $G$ is exponential), we can see from the above theorem that $E\, X_{i,n} / \sum_{j=1}^{i} (n-j+1)^{-1}$ is decreasing in $i$ and increasing in $n$. The theorem also provides bounds on $E\, X_{i,n}$. If $\theta$ is the common mean of $F$ and $G$, it is easy to show that

$$(5.3) \qquad \theta\, E\, Y_{i,n}/E\, Y_{i,i} \le E\, X_{i,n} \le \theta\, E\, Y_{i,n}/E\, Y_{1,n-i+1}.$$

In order to state a few other interesting results, we need the following definitions.

Definition 5.1. A sequence $\underline{a} = (a_1,\ldots, a_n)$ is said to majorize a sequence $\underline{b} = (b_1,\ldots, b_n)$ (written $\underline{a} \succ \underline{b}$), if $a_1 \ge a_2 \ge \cdots \ge a_n$, $b_1 \ge \cdots \ge b_n$, and $\sum_{i=1}^{r} a_i \ge \sum_{i=1}^{r} b_i$ for $r = 1,2,\ldots, n-1$, while $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$.

The above definition is according to Beckenbach and Bellman [21] and differs slightly from that of Hardy, Littlewood and Pólya [67].

Definition 5.2. If a differentiable function $H(z_1,\ldots, z_n)$ satisfies $(z_i-z_j)(\frac{\partial H}{\partial z_i} - \frac{\partial H}{\partial z_j}) \ge 0$ for all $\underline{z}$, $i,j$, then $H$ is said to satisfy the Schur condition.

Theorem 5.3. Let $F$ and $G$ have a common mean $\theta$, and $F \underset{*}{<} G$. Then

(i) $\sum_{i=1}^{r} E\, Y_{(i)}/\sum_{i=1}^{r} E\, X_{(i)}$

and

$$[\sum_{i=1}^{r} (n-i+1)\ E(Y_{(i)}-Y_{(i-1)})]/[\sum_{i=1}^{r} (n-i+1)\ E(X_{(i)}-X_{(i-1)})]$$

are increasing in $r$, $1 \le r \le n$;

(ii) $(E\, Y_{(n)}, E\, Y_{(n-1)},\ldots, E\, Y_{(1)}) \succ (E\, X_{(n)}, E\, X_{(n-1)},\ldots, E\, X_{(1)})$

and

$$\sum_{i=1}^{r} (n-i+1)\ E(X_{(i)}-X_{(i-1)}) \ge \sum_{i=1}^{r} (n-i+1)\ E(Y_{(i)}-Y_{(i-1)})$$

for $1 \le r \le n$;

(iii) $H(E\,Y_{(n)}, E\,Y_{(n-1)}, \ldots, E\,Y_{(1)}) \geq H(E\,X_{(n)}, E\,X_{(n-1)}, \ldots, E\,X_{(1)})$, *if* H *is a Schur function;*

(iv) $\sum_{i=1}^{n} a_i (n-i+1)\, E(X_{(i)}-X_{(i-1)}) \geq \sum_{i=1}^{n} a_i (n-i+1)\, E(Y_{(i)}-Y_{(i-1)})$,

*if* $a_1 \geq a_2 \geq \ldots \geq a_n$.

When F is IFRA (DFRA), part (i) of the above theorem means that $E\,\hat{\theta}_{r,n}$ is decreasing (increasing) in r so that $E\,\hat{\theta}_{r,n} \geq (\leq)\ \theta$, where $\hat{\theta}_{r,n} = r^{-1}\, T_{r,n}$ and

(5.4) $$T_{r,n} = \sum_{i=1}^{r} (n-i+1)(X_{i,n}-X_{i-1,n})$$

is the total life statistic defined in (4.3) with A = 0. Also, part (iv) yields

(5.5) $$\sum_{i=1}^{n} a_i (n-i+1)\, E(X_{(i)}-X_{(i-1)}) \leq (\geq)\ \theta \sum_{i=1}^{n} a_i,$$

if $a_1 \geq a_2 \geq \ldots \geq a_n$.

Further, if F is IFRA with mean θ, the result (5.3) on bounds for expectation reduces to

(5.6) $$\theta \sum_{j=1}^{i} (n-j+1)^{-1} / \sum_{j=1}^{i} j^{-1} \leq E\,X_{i,n} \leq n\theta \sum_{j=1}^{i} (n-j+1)^{-1}, \quad 1 \leq i < n,$$

and

(5.7) $$\theta \leq E\,X_{n,n} \leq \theta \sum_{j=1}^{n} j^{-1}.$$

The above bounds are non-trivial but only sharp for i = 1 or i = n. On the other hand, if F is DFRA with mean θ, the following bounds can be obtained using Theorem 5.2.

$$0 \le E\, X_{1,n} \le \theta/n,$$

(5.8) $$0 \le E\, X_{i,n} \le i\,\theta \sum_{j=1}^{i} (n-j+1)^{-1} / \sum_{j=1}^{i} j^{-1}, \qquad 1 < i < n,$$

$$\theta \sum_{j=1}^{n} j^{-1} \le E\, X_{n,n} \le n\theta.$$

All lower bounds in (5.8) are sharp. The upper bound is sharp when $i = 1$ and $i = n$.

Let $\bar{X} = n^{-1} \sum_{i=1}^{n} X_{i,n}$ and $\bar{Y} = n^{-1} \sum_{i=1}^{n} Y_{i,n}$. We define $(U_1, \ldots, U_n) \underset{st}{\prec} (V_1, \ldots, V_n)$ to mean $\sum_{i=1}^{j} U_i \underset{st}{\le} \sum_{i=1}^{j} V_i$ for $j = 1, 2, \ldots, n-1$, while $\sum_{i=1}^{n} U_i \underset{st}{=} \sum_{i=1}^{n} V_i$. We state below a theorem due to Marshall, Olkin and Proschan [75].

<u>Theorem 5.4</u>. <u>Let</u> $F \underset{*}{\prec} G$. <u>Then</u>

(i) $(X_{(n)}/n\,\bar{X}, \ldots, X_{(1)}/n\,\bar{X}) \underset{st}{\prec} (Y_{(n)}/n\,\bar{Y}, \ldots, Y_{(1)}/n\,\bar{Y})$;

(ii) $H(X_{(n)}/n\,\bar{X}, \ldots, X_{(1)}/n\,\bar{X}) \underset{st}{\le} H(Y_{(n)}/n\,\bar{Y}, \ldots, Y_{(1)}/n\,\bar{Y})$,

<u>if</u> H <u>is</u> <u>a</u> <u>Schur</u> <u>function</u>;

(iii) $\sum_{i=1}^{r} (n-i+1)(X_{(i)}-X_{(i-1)})/\bar{X} \underset{st}{\ge} \sum_{i=1}^{r} (n-i+1)(Y_{(i)}-Y_{(i-1)})/\bar{Y}$;

(iv) $(n^{-1} \sum_{i=1}^{n} X_{(i)}^2 - \bar{X}^2)/\bar{X}^2 \underset{st}{\le} (n^{-1} \sum_{i=1}^{n} Y_{(i)}^2 - \bar{Y}^2)/\bar{Y}^2$;

(v) <u>if</u>, <u>in</u> <u>addition</u>, $a_1 \ge \ldots \ge a_n$, <u>then</u>

$$\sum_{i=1}^{n} a_i X_{(i)}/\bar{X} \underset{st}{\ge} \sum_{i=1}^{n} a_i Y_{(i)}/\bar{Y}$$

<u>and</u>

$$\sum_{i=1}^{n} a_i (n-i+1)(X_{(i)}-X_{(i-1)})/\bar{X} \underset{st}{\ge} \sum_{i=1}^{n} a_i (n-i+1)(Y_{(i)}-Y_{(i-1)})/\bar{Y}.$$

In the case of an IFRA (DFRA) distribution F, we obtain from part (iii) of the theorem

(5.9) $$\hat{\theta}_{r,n}(\underset{\sim}{X})/\bar{X} \underset{st}{\geq} (\underset{st}{\leq}) \hat{\theta}_{r,n}(\underset{\sim}{Y})/\bar{Y}.$$

Let us consider a life test with n items on test where successive failures are observed until a pre-assigned time $t_o$. For this sample, define

(5.10) $$V(t_o) = \sum_{i=1}^{r} X_{i,n} + (n-r)t_o,$$

where r denotes the number of observations $\leq t_0$. As we can see, $V(t_o)$ denotes the total time on test up to time $t_o$. This statistic has been studied by Epstein and Sobel [37]. For F starshaped with respect to G, we have the following theorem concerning the total time on test due to Barlow and Proschan [11].

Theorem 5.5. Let $F \underset{*}{<} G$ with a common mean $\theta$. Then

(5.11) $$E[\sum_{i=1}^{r} X_{i,n} + (n-r)t_o] \geq E[\sum_{i=1}^{s} Y_{i,n} + (n-s)t_o],$$

where r(s) denotes the number of X (Y) observations $\leq t_o$.

When F is an IFRA distribution with a known mean, we can use (5.11) to obtain a lower bound on the expected total time on test in truncated sampling from F.

5.2. Inequalities in the case of F convex with respect to G. In this part, we consider pairs of distributions F and G such that $F \underset{c}{<} G$. We give a few results due again to Barlow and Proschan [11].

Theorem 5.6. Let $F \underset{c}{<} G$ and $F(0) = G(0) = 0$. Then the following statements are true.

(i) Suppose $0 \leq \bar{A}_i \leq 1$ for $i = 1,2,\ldots, n$. Then

(5.12) $$F\left(\sum_{i=1}^{n} a_i X_{(i)}\right) \underset{st}{\le} G\left(\sum_{i=1}^{n} a_i Y_{(i)}\right).$$

(ii) _Suppose there exists a_ $k$, $0 \le k \le n$, _such that_ $\bar{A}_i \ge 1$, $i = 1, 2, \ldots, k$ _and for_ $k < n$, $\bar{A}_i \le 0$, $i = k+1, \ldots, n$. Then

(5.13) $$F\left(\sum_{i=1}^{n} a_i X_{(i)}\right) \underset{st}{\ge} G\left(\sum_{i=1}^{n} a_i Y_{(i)}\right).$$

(iii) _Let_ $F$ _and_ $G$ _have a common mean_ $\theta$ _and_ $\bar{A}_i \ge 1$, $i = 1,2,\ldots, r$. _Then_

(5.14) $$P_F\left[\sum_{i=1}^{r} \bar{A}_i (X_{(i)} - X_{(i-1)}) \ge x\right] \ge P_G\left[\sum_{i=1}^{r} \bar{A}_i (Y_{(i)} - Y_{(i-1)}) \ge x\right]$$

for $x \le \theta \min(\bar{A}_1, \ldots, \bar{A}_r)$.

(iv) _Let_ $a_i \ge 0$ _for_ $i = 1,2,\ldots, n$, _and_ $\sum_{i=1}^{n} a_i \le 1$. _Then_

(5.15) $$F\left(\sum_{i=1}^{n} a_i \ EX_{(i)}\right) \le G\left(\sum_{i=1}^{n} a_i \ EY_{(i)}\right).$$

One can use (5.12) and (5.13) to obtain conservative lower and upper tolerance limits for distributions $F$ for which $G^{-1}F$ is convex [12]. In the cases of IFR and DFR distributions, we have additional results concerning the total time on test.

_Theorem 5.7_. _Let_ $F$ _be IFR (DFR) and_ $F(0^-) = 0$. _Then the total life statistic_ $T_{r,n}$ _defined in_ (5.4) _is stochastically increasing (decreasing) in_ $n \ge r$.

Another result concerning IFR distribution can be readily obtained from (5.14) by setting $\bar{A}_i = n-i+1$, $i = 1,2,\ldots, r$. This gives the following corollary.

_Corollary 5.7.1_. _Let_ $F$ _be IFR with mean_ $\theta$ _and_ $G(t) = 1-e^{-t/\theta}$. _Then_

(5.16) $$P_F(T_{r,n}(\underset{\sim}{X}) > x) \geq P_G(T_{r,n}(\underset{\sim}{Y}) > x), \quad x \leq (n-r+1)\theta,$$

*where* $T_{r,n}(\underset{\sim}{X})$ *denotes the total life statistic using* X *observations.*

Sharp bounds on expected values of order statistics from an IFR distribution can be given in terms of the pth percentile. A result in this direction is given below.

*Theorem 5.8. Let* F *be IFR with* pth *percentile* $\xi_p$. *Then, for* $1 \leq j \leq n$

(5.17) $$E\, X_{(j)} \leq \max\{\xi_p, \frac{p}{-\log q}(\frac{1}{n} + \ldots + \frac{1}{n-j+1})\},$$

*and*

(5.18) $$E\, X_{(j)} \geq \sum_{i=0}^{j-1} \binom{n}{i} \int_0^{\xi_p} [1-g(x)]^i [g(x)]^{n-i}\, dx,$$

*where* $q = 1-p$ *and* $g(x) = \exp\{\xi_p^{-1} x \log q\}$.

*5.3. Properties preserved in taking order statistics from IFR (DFR) distributions.* Now we point out some differences in the behavior of order statistics and the spacings between IFR and DFR distributions. For an IFR distribution, while the order statistics also have an IFR distribution [10, pp. 38-39], the same is not true for the spacings. In the case of a DFR distribution, it has been shown in [10] that the order statistics are not necessarily DFR, whereas the spacings are. A stronger property than IFR is that the distribution F has a density f such that log f(x) is concave where finite. In this case we say that f is a *Pólya frequency of order 2* ($PF_2$). Barlow and Proschan [11] have proved the following theorem.

*Theorem 5.9. Suppose* f *is* $PF_2$ *with* f(x) *not necessarily zero for*

$x < 0$. Let $f_{in}$ denote the density of $X_{i,n}$ and $h_i$ denote that of the spacing $X_{i,n}-X_{i-1,n}$, $i = 2,\ldots, n$. Then

(i) $f_{in}$ is $PF_2$ for $i = 1,\ldots, n$;

(ii) $h_i$ is $PF_2$ for $i = 2,\ldots, n$;

(iii) $h_1$ is $PF_2$, if $f(x)=0$ for $x<0$, where $h_1$ is the density of $X_{1,n}$.

The fact that the order statistics preserve the IFR property has a physical meaning. If we consider a so-called k-out-of-n structure consisting of n identical components each having an IFR failure distribution, then the life of the structure corresponds to the kth order statistic and hence is also IFR.

Besides these properties relating to order statistics, there are other interesting basic properties of IFR and DFR distributions which are useful in many situations. For example, the IFR property is preserved under convolution where as the DFR property is not. This would mean that a system consisting of a single IFR unit supported by n-1 spares will have the IFR property. These and other properties can be found in the monograph by Barlow and Proschan [10].

6. Some estimation and testing problems for restricted families of distributions. In this section we describe briefly some of the estimation problems and tests of hypotheses concerning mainly distributions with monotone failure rate. We will not deal with these very elaborately here. Our purpose is to include some of the important problems for the sake of completeness of our survey. The details and other references can be obtained from the recent book by Barlow *et*

al [ 2].

6.1. Estimation for distributions with monotone failure rate. If F is the life distribution of an item which is subject to wear-out, the shape of the failure rate function often suggests when to take appropriate maintenance actions. So the estimation of the failure rate function from a sample of n independent identically distributed lifetimes is of practical interest. Experience indicates that, in many situations, the failure rate function could be just monotone or U shaped. We consider here the monotone failure rate functions. The U shaped case can be handled by modifying the monotonic estimators, see for example Bray, Crawford and Proschan [25].

Let F be an unknown distribution in the class $\mathcal{F}$ of IFR distributions with support contained in $[0,\infty)$. We say that $F \in \mathcal{F}$ if $-\log[1-F(x)]$ is convex on the support (i.e., the points of increase) of F, an interval contained in $[0,\infty)$. It can be shown [10, p.26] that F is absolutely continuous except for the possibility of a jump at the right hand endpoint of its support.

We wish to obtain a maximum likelihood estimator (MLE) of the failure rate function $r(x) = f(x)[1-F(x)]^{-1}$ which is well-defined for $F(x) < 1$. For convenience, if F is IFR, we define $r(x) = \infty$ for all x such that $F(x) = 1$. It is not possible to obtain an MLE for $F \in \mathcal{F}$ by directly maximizing $\prod_{i=1}^{n} f(X_{i,n})$ because $f(X_{n,n})$ can be chosen arbitrarily large. To get over this difficulty we adopt the approach of Grenander [42], who first obtained supremum of the likelihood function over the subclass $\mathscr{F}^M$ of distributions $F \in \mathscr{F}$ whose failure rate functions are bounded by M. Using the fact that for any distribution F and any x for which r is finite on $[0,x)$

$$(6.1) \qquad 1-F(x) = \exp[-\int_0^x r(u)\,du],$$

we can write the log likelihood $L = L(F)$ for $F \in \mathscr{F}^M$ in the form

$$(6.2) \qquad L = \sum_{i=1}^{n} \log r(X_{i,n}) - \sum_{i=1}^{n} \int_0^{X_{i,n}} r(u)\,du .$$

Suppose $F \in \mathscr{F}^M$ has failure rate r and let F* be the distribution with failure rate

$$(6.3) \quad r^*(x) = \begin{cases} 0, & x < X_{1,n}, \\ r(X_{i,n}), & X_{i,n} \le x < X_{i+1,n}, \quad i = 1,\dots, n-1, \\ r(X_{n,n}), & x \ge X_{n,n} . \end{cases}$$

Then it is easy to see that $L(F) \le L(F^*)$. This shows that the MLE for r is a step function and that L may be replaced by the function

$$(6.4) \qquad \sum_{i=1}^{n} \log r(X_{i,n}) - \sum_{i=1}^{n-1} (n-i)(X_{i+1,n} - X_{i,n})\, r(X_{i,n}) .$$

So the problem reduces to maximizing (6.4) subject to $r_1 \le r_2 \le \dots \le r_n \le M$, where $r_i = r(X_{i,n})$, $i = 1,\dots, n$. It has been shown by Marshall and Proschan [76] that the solution is given by

$$(6.5) \quad \hat{r}_n^M(x) = \begin{cases} 0, & x < X_{1,n}, \\ \hat{r}_i, & X_{i-1,n} \le x < X_{i,n}, \ i = 2,\dots, n, \\ M, & x \ge X_{n,n}, \end{cases}$$

where

$$(6.6) \qquad \hat{r}_i = \hat{r}_n(X_{i,n}) = \min_{i \le t \le n-1} \max_{1 \le s \le i} \frac{t - s + 1}{\sum_{j=s}^{t} (n-j)(X_{j+1,n} - X_{j,n})}$$

and it is assumed that $M > \max_{1 \le i \le n-1} [(n-i)(X_{i+1,n} - X_{i,n})]^{-1}$. Now the MLE is obtained by letting $M \to \infty$, so that $\hat{r}_n(x) = +\infty$ for $x \ge X_{n,n}$. The solution in (6.5) can be obtained elegantly in the general context of isotonic regression explained in [ 2]. It should also be pointed out that (6.5) is different in form from the one given in [42] but is equivalent to it. The following strong consistency theorem has been proved by Marshall and Proschan [76].

Theorem 6.1. *If F is IFR, then for every* $x_o$,

$$(6.7) \qquad r(x_o^-) \le \underline{\lim}\, \hat{r}_n(x_o) \le \overline{\lim}\, \hat{r}_n(x_o) \le r(x_o^+)$$

*with probability one.*

The above theorem leads to the following corollary.

Corollary 6.1.1. *If r is increasing and continuous on* [a,b], *then for* $t \in [a,b]$

$$(6.8) \qquad \lim_{n \to \infty} |\hat{r}_n(t) - r(t)| = 0$$

*with probability one.*

Now, the MLE of the distribution function F is given by

$$(6.9) \qquad \hat{F}_n(x) = 1 - \exp[- \int_{-\infty}^{x} \hat{r}_n(u)\, du]$$

and if r is continuous, the strong consistency of $\hat{F}_n$ follows as an almost immediate consequence of Theorem 6.1.

An interesting and useful fact is that (6.6) can be expressed in the form

$$\hat{r}_n(X_{i,n}) = \min_{t \geq i+1} \max_{s \leq i} \frac{F_n(X_{t,n}) - F_n(X_{s,n})}{H_n^{-1}[F_n(X_{t,n})] - H_n^{-1}[F_n(X_{s,n})]}, \tag{6.10}$$

where

$$H_n^{-1}\left(\frac{i}{n}\right) = \int_0^{X_{i,n}} [1-F_n(u)]\, du \tag{6.11}$$

and $F_n$ denotes the empirical distribution. The right hand member of (6.11) is equal to $n^{-1}\ T_{n,i}$, where

$$T_{n,i} = \sum_{j=1}^{i} (n-j+1)(X_{j,n} - X_{j-1,n}), \tag{6.12}$$

is the total time on test up to the ith observation defined in (4.3). The total time on test transformation first appeared in [76] but was not made use of by the authors. The role of the transformation in the proof of the consistency of the MLE can be seen in [ 2].

The asymptotic distribution of the MLE of r has been obtained by Prakasa Rao [83], whose approach is to reduce the estimation problem to that of a Weiner process and use convergence theorems for stochastic processes. His result is given below.

Theorem 6.2. *Let F be IFR with failure rate r. Let ξ be such that 0 < F(ξ) < 1. Further suppose that r is differentiable at ξ with non-zero derivative and r(ξ) > 0. Let $\hat{r}_n(\xi)$ denote the MLE of r(ξ) based on n independent observations. Then the asymptotic distribution of*

$$\left[\frac{2n\ r^4(\xi)\ f(\xi)}{r'(\xi)}\right]^{1/3} \left[\frac{1}{\hat{r}_n(\xi)} - \frac{1}{r(\xi)}\right] \tag{6.13}$$

*has density $\frac{1}{2}\ \psi(x/2)$ where ψ is the density of the minimum value of*

$W(t) + t^2$ *and* $W(t)$ *is a two-sided Wiener-Levy process with mean* 0, *variance* 1 *per unit* t *and* $W(0) = 0$.

In the case of a DFR distribution F (i.e., log $[1-F(x)]'$ is convex on the support of F, an interval of the form $[\alpha,\infty)$), the distribution is absolutely continuous except for a jump at $\alpha$. Suppose $\alpha$ is known and $\alpha = X_{1,n} = \cdots = X_{k,n} < X_{k+1,n} < \cdots < X_{n,n}$. If $k = 0$, we define $X_{o,n} = \alpha$. It can be shown that the MLE is

$$(6.14) \qquad \hat{r}_n(x) = \begin{cases} k/n\,, & x = \alpha, \\ \hat{r}_n(X_{i,n}), & X_{i-1,n} < x \le X_{i,n}, \quad i = k+1,\ldots, n, \end{cases}$$

where

$$(6.15) \qquad \hat{r}_n(X_{i,n}) = \max_{t \ge i} \min_{s \le i-1} (t-s)[T_{n,t} - T_{n,s}]^{-1}$$

and $T_{n,i}$ is as defined in (6.12). Contrary to the IFR case, this DFR estimator is not unique; it is determined by the likelihood equation only for $x \le X_{n,n}$, and may be extended beyond $X_{n,n}$ in any manner that preserves the DFR property. In this case, the statement of Theorem 6.2 is valid with DFR and $-r'(\xi)$ substituted for IFR and $r'(\xi)$, respectively. When $\alpha$ is unknown, it is shown in [76] that MLE is found among the DFR distributions with support $[X_{1,n},\infty)$ and thus the problem reduces to the case of known $\alpha$. The relatively slow rate of convergence indicated by (6.13) prompted Barlow and van Zwet [16], [17] to consider the so-called window estimators for the generalized failure rate function discussed below.

As we have pointed out earlier in §3, if F is IFR, then $F \underset{c}{<} G$, where $G(x) = 1-e^{-x}$, $x \ge 0$. Since, in this case, $G^{-1}F(x) = -\log[1-F(x)]$, the failure rate function $r(x)$ can be written as

$$(6.16) \qquad r(x) = (d/dx)\, G^{-1}F(x) = f(x)/g[G^{-1}F(x)],$$

where $g$ is the density of $G$. This idea is used to define the <u>generalized failure rate function</u> for any general $G$ by the relation (6.16).

Let $\mathscr{F}$ be the class of absolutely continuous distribution functions $F$ on $(0,\infty)$ (here $F(0)$ need not be 0) with positive and right (or left) continuous density $f$ on the interval where $0 < F < 1$. $F^{-1}(0)$ and $F^{-1}(1)$ are taken to be equal to the left and right hand endpoints of the support of $F$. We consider $F \underset{c}{\prec} G$, where $F, G \in \mathscr{F}$ and $G$ is specified. Let $X_{1,n} \le X_{2,n} \le \cdots \le X_{n,n}$ denote an ordered sample from $F$. For each $n$, we define a <u>grid</u> on $(-\infty,\infty)$, namely, a finite or infinite sequence $\cdots < t_{n,-1} < t_{n,0} < t_{n,1} < \cdots$. In each <u>window</u> $[t_{n,j}, t_{n,j+1})$ we choose a point $x_{n,j}$ to which a non-negative weight $w(x_{n,j})$ is assigned. We start with an initial or <u>basic estimator</u> $\hat{\rho}_n$ for $r$, the generalized failure rate function. Here we take $\hat{\rho}_n$ to be the <u>naive</u> estimator for $r$ given by

$$(6.17) \quad \hat{\rho}_n(x) = [gG^{-1}F_n(x_i)(t_{n,i+1}-t_{n,i})]^{-1}[F_n(t_{n,i+1})-F_n(t_{n,i})]$$

for $t_{n,i} \le x < t_{n,i+1}$, where $F_n$ is the empirical distribution corresponding to our sample. We choose

$$(6.18) \qquad w(x_{n,j}) = gG^{-1}F_n(x_j)(t_{n,j+1}-t_{n,j}).$$

Now, using $\hat{\rho}_n$ and the weights $w(x_{n,j})$, we define

$$(6.19) \quad \hat{r}_n(x) = \min_{s\ge i+1} \max_{r\le i} [\sum_{j=r}^{s-1} \hat{\rho}_n(x_{n,j})w(x_{n,j})]/\sum_{j=r}^{s-1} w(x_{n,j}), \quad t_{n,i}\le x<t_{n,i+1},$$

which is called the isotonic regression of $\hat{\rho}_n$ with respect to the weights $w(x_{n,j})$. It should be noted that $\hat{\rho}_n(x_{n,j})$ is defined only if $X_{1,n} \leq t_{n,j} < t_{n,j+1} \leq X_{n,n}$. This condition is always assumed in (6.19) though not explicitly stated. By choosing

$$(6.20) \qquad w(x_{n,j}) = gG^{-1}F(x_{n,j})\ (t_{n,j+1}-t_{n,j}),$$

we see that

$$(6.21)\ \hat{r}_n(x) = \min_{s \geq i+1}\ \max_{r \leq i} [F_n(t_{n,s}) - F_n(t_{n,r})][\sum_{j=r}^{s-1} (t_{n,j+1}-t_{n,j}) gG^{-1}F_n(x_{n,j})]^{-1}.$$

Note that $\hat{r}_n$ is a nondecreasing step function and that $\hat{r}_n = \hat{\rho}_n$ whenever $\hat{\rho}_n(x_{n,j})$ is nondecreasing in j.

If we consider the random grid $t_{n,j} = X_{j,n}$, $j = 1,\ldots, n$, determined by the order statistics, (6.21) reduces to

$$(6.22)\ \hat{r}_n(x) = \min_{i+1 \leq s \leq n}\ \max_{1 \leq r \leq i}\ (s-r)[n \sum_{j=r}^{s-1} gG^{-1}(j/n)(X_{j+1,n}-X_{j,n})]^{-1}$$

for $X_{i,n} \leq x < X_{i+1,n}$, $i = 1,2,\ldots, n-1$. This is, in fact, the MLE when G is exponential. In this case

$$n\ gG^{-1}(j/n)(X_{j+1,n}-X_{j,n}) = (n-j)(X_{j+1,n}-X_{j,n})$$

which is the total time on test in the interval $(X_{j,n}, X_{j+1,n}]$. For this reason the quantities $gG^{-1}(j/n)(X_{j+1,n}-X_{j,n})$ are called the total time on test weights for general G.

For the grid consisting of order statistics, the estimator $\hat{\rho}_n$ is not consistent. Under certain regularity conditions, it is shown by Barlow and van Zwet [16], [17] that $\hat{r}_n$ is strongly consistent, provided the grid becomes "dense" in the support of F with probability 1 as $n\to\infty$.

Their results on the associated asymptotic distributions are summarized in the following theorem.

Theorem 6.3. Let $F, G \in \mathcal{F}$. Further assume

(a) the generalized failure rate $r(x)$ is nondecreasing in $x \geq 0$;

(b) $r$ is continuously differentiable and $f''$ exists in a neighborhood of $x$;

(c) $r'(x) > 0$ and

(d) $t_{n,i+1} - t_{n,i} = cn^{-\alpha}$, $0 < \alpha < 1$.

Then, the following statements are true.

(A) If $\frac{1}{5} < \alpha < \frac{1}{3}$, then $[c\ f(x)]^{1/2}\ n^{(1-\alpha)/2}[r(x)]^{-1}[\hat{r}_n(x)-r(x)]$ is asymptotically $N(0,1)$.

(B) If $\frac{1}{7} < \alpha < \frac{1}{5}$, then

$$[c\ f(x)]^{1/2}\ n^{(1-\alpha)/2}[r(x)]^{-1}[\hat{r}_n(x)-r(x)-c^2n^{-2\alpha}f''(x)r(x)\ \{24f(x)\}^{-1}]$$

is asymptotically $N(0,1)$.

(C) If $\frac{1}{3} < \alpha \leq 1$, then $[2nf(x)/r'(x)r^2(x)]^{1/3}[\hat{r}_n(x)-r(x)]$ has density $\frac{1}{2}\,\psi(u/2)$, where $\psi$ is the density of the minimum value of $W(t) + t^2$ and $W(t)$ is a two-sided Wiener-Levy process with mean 0 and variance 1 per unit $t$ and $W(0) = 0$.

The problem of optimal choice of window size for nonparametric estimators of the density and failure rate function have been investigated by Parzen [81], Watson and Leadbetter [93], and Weiss and Wolfowitz [94].

Isotonic estimators for star-ordered families of distributions have been proposed and studied by Barlow and Scheuer [15]. The MLE's, for example, when G is exponential and uniform, are <u>not</u> isotonic estimators. Moreover, surprisingly enough, they are also <u>not</u> consistent. It is shown [2, p. 259] that the isotonic estimators in these cases are consistent. Percentile estimators in the star-ordering case have also been discussed in [2].

<u>6.2. Tests for monotone failure rate</u>. As we have seen elsewhere, the knowledge that the failure rate is increasing is useful in constructing appropriate bounds for survival probability, and obtaining inequalities for linear combinations of order statistics. Also, for practical applications, a knowledge of this fact leads to the maximum likelihood estimate of the failure rate. It also helps, for example, in adopting economical replacement policies obtained under this assumption.

As before, $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ is an ordered sample from the distribution F, with density f, where $f(x) = 0$ for $x < 0$, and failure rate $r(x)$. The interest is in testing the null hypothesis, $H_o: r(x) = \lambda$, $\lambda$ an unknown positive constant, against the alternative hypothesis, $H_1$: $r(x)$ is nondecreasing, but not constant.

Let us define

$$V_{i,j} = \begin{cases} 1 , & \text{if } \bar{D}_{i,n} \geq \bar{D}_{j,n} \text{ for } i,j = 1,2,\ldots, n, \\ 0 , & \text{otherwise,} \end{cases} \tag{6.23}$$

where $\bar{D}_{i,n} = (n-i+1)\, D_{i,n}$ and $D_{i,n} = X_{i,n} - X_{i-1,n}$, $i = 1,2,\ldots, n$. It is assumed that $X_{o,n} \equiv 0$. The $\bar{D}_{i,n}$ are called the <u>normalized spacings</u>.

Proschan and Pyke [84] have proposed a test based on the statistic

(6.24) $$V_n = \sum_{\substack{i,j=1 \\ i<j}}^{n} V_{i,j} \; .$$

The null hypothesis is rejected at the level $\alpha$ of significance if $V_n > v_{n,\alpha}$, where $v_{n,\alpha}$ is determined so that $\Pr(V_n > v_{n,\alpha} | H_o) = \alpha$. The distribution of $V_n$ under $H_o$ can be obtained by a recurrence relation and is given by Kendall [70] and Mann [74], who have also given the values of $P[V_n \leq k]$ for $n \leq 10$. It has been shown by Proschan and Pyke [84] that the test is unbiased and that $V_n$, suitably normalized, is asymptotically normally distributed for a wide class of alternatives. In particular, under mild assumptions when the underlying distribution is IFR, $V_n$ is asymptotically normally distributed. The asymptotic relative efficiency of this test has also been considered in [84] relative to the likelihood-ratio test for Weibull and gamma alternatives.

Bickel and Doksum [23] have considered the statistics $T_1 = \Sigma i R_i$ and $T_2 = -\Sigma\, i \log[1-R_i(n+1)^{-1}]$, where $R_i$ denotes the rank of $\bar{D}_{i,n}$ among $\bar{D}_{1,n},\ldots,\bar{D}_{n,n}$. Then $T_1$ is asymptotically equivalent to the Proschan-Pyke statistic. It has been shown in [23] that the asymptotic normality holds for sequences of alternatives $\{F_{\theta_n}\}$ that approach the null hypothesis distribution, namely, the exponential with failure rate $\lambda$, as $n\to\infty$. It is also shown that, the rank statistics that are asymptotically most powerful in the class of linear rank tests, are nowhere most powerful in the class of all tests, when the scale parameter $\lambda$ is known. If $\lambda$ is unknown, studentizing of the linear normalized spacing tests, which are asymptotically most powerful for $\lambda$ known, leads to procedures which have only the same asymptotic power as the most powerful linear rank tests. Bickel [22] has considered four classes of tests including those based on $T_1$ and $T_2$. He shows that, under certain regularity conditions, each of these classes contains a test asymptotically equivalent to the asymptotically most powerful similar test.

The tests of Bickel and Doksum are in fact unbiased against IFRA alternatives as shown by results in Barlow and Proschan [11]. It was shown by Barlow and Proschan [14] that analogous tests designed to treat incomplete samples of failure data are also unbiased against IFRA alternatives.

Barlow and Doksum [ 3] have considered the problem of testing $H_o$: $F \underset{c}{=} G$ (that is, $G^{-1}F$ is linear on the support of F) against the alternative $H_1$: $F \underset{c}{\prec} G$ and $F \underset{c}{\neq} G$, where G is assumed to be known and $F \underset{c}{\prec} G$. They considered tests based on the <u>cumulative total time on test statistic</u>

$$(6.25) \qquad U_n = n^{-1} \sum_{i=1}^{n-1} [\sum_{j=1}^{i} \bar{D}_{j,n}] / \sum_{i=1}^{n} \bar{D}_{i,n} \, ,$$

where the $\bar{D}_{i,n}$ are the normalized spacings. In terms of $H_n^{-1}$ defined in (6.11) we can also write it as

$$(6.26) \qquad U_n = n^{-1} \sum_{i=1}^{n-1} H_n^{-1}(i/n)/H_n^{-1}(1) \, .$$

The null hypothesis is rejected for large values of the statistic. Asymptotic normality of $U_n$ has been shown in [ 3]. The corresponding result when G is exponential is attributed to an unpublished paper by

Nadler and Eilbott. Asymptotic minimax property of the tests based on $U_n$ has also been established by Barlow and Doksum over a class of alternatives based on the Kolmogorov distance and in each of the classes of statistics considered by Bickel and Doksum [23] when G is exponential.

Barlow [1] has investigated a likelihood ratio test for $H_0$: $F \in \mathscr{F}_0$ against $H_1$: $F \in \mathscr{F} - \mathscr{F}_0$, where $\mathscr{F}_0$ denotes the class of exponential distributions with possible truncation on the right and $\mathscr{F}$ is the class of all IFR distributions. The test is shown to be unbiased. He also considers a likelihood ratio test for $H_0$: $F \in \mathscr{F}_0$ against $H_1$: $F \in \mathscr{F}_1$, where $\mathscr{F}_0$ and $\mathscr{F}_1$ are the classes of IFRA and DFRA distributions, respectively. It is shown by him that the likelihood ratio statistic has a nonincreasing density on (0,1) under the exponential assumption. It should be pointed out that in these tests the usual concept of maximum likelihood does not suffice and the concept used is a generalization of the usual concept proposed by Kiefer and Wolfowitz [71].

7. Concluding remarks. As we have explained elsewhere, we have confined our interest to a few important results relating to subset selection and reliability problems. We have not discussed here several other procedures of interest which have been investigated such as sequential procedures, Bayes procedures and procedures based on paired comparisons. These and other procedures have been reviewed by Gupta and Panchapakesan [58].

One related problem that is not discussed here is the estimation of the ordered parameters. Some work has been done on maximum likelihood and interval estimation among others by Dudewicz [29], Dudewicz and Tong [30], Tong [91], and Tong and Saxena [92]. Many similar problems arise when the distributions belong to some restricted family.

There are other classes of life distributions which have been studied recently. If $F$ is a life distribution ($F(0) = 0$), we say $F$ is <u>new better than used</u> (NBU) if $\bar{F}(x+y) \leq \bar{F}(x)\,\bar{F}(y)$ for all $x,y \geq 0$, where $\bar{F} = 1-F$. If the inequality is reversed, $F$ is said to be <u>new worse than used</u> (NWU). If $F$ is NBU, the obvious interpretation is that the chance $\bar{F}(x)$ that a new unit will survive to age $x$ is greater than the chance $\bar{F}(x+y)/\bar{F}(y)$ that an unfailed unit of age $y$ will survive an additional time $x$. We note that if $G(x) = 1-e^{-x}$, $x \geq 0$, then saying $F$ is NBU is equivalent to the statement that $\varphi \equiv G^{-1}F$ is <u>superadditive</u>, i.e., $\varphi(x+y) \geq \varphi(x) + \varphi(y)$ for all $x$ and $y$, on the support of $F$. If $F$ is NWU, then $\varphi$ is <u>subadditive</u>. It follows from the results of Bruckner and Ostrow [26] that $F$ is IFRA (DFRA) implies that $F$ is NBU(NWU). Some properties of these families have been studied by Esary, Marshall and Proschan [39]. Additional properties of such classes of distributions and their importance in the study of replacement policies have been recently discussed by Marshall and Proschan [77]. In another recent paper, Hollander and Proschan [68] have considered a test based on a U-statistic for testing the hypothesis that $F$ is exponential against the alternative that $F$ is NBU (and not exponential).

<u>Acknowledgment.</u> The authors wish to thank Ms. Britt Marie Alsen

for carrying out the computations for the tables in this paper. Thanks are also due to Professor Herman Rubin for some helpful suggestions regarding these computations.

REFERENCES

[1] R. E. BARLOW, Likelihood ratio tests for restricted families of probability distributions, Ann. Math. Statist., 39 (1968), pp. 547-560.

[2] R. E. BARLOW, D. J. BARTHOLOMEW, J. M. BREMNER and H. D. BRUNK, Statistical Inference under Order Restrictions, John Wiley, New York, 1972.

[3] R. E. BARLOW and K. A. DOKSUM, Isotonic tests for convex orderings, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, University of California Press, Berkeley and Los Angeles, California, 1972, pp. 293-323.

[4] R. E. BARLOW, J. D. ESARY and A. W. MARSHALL, A stochastic characterization of wear-out for components and systems, Ann. Math. Statist., 37 (1966), pp. 816-825.

[5] R. E. BARLOW and S. S. GUPTA, Distribution-free life test sampling plans, Technometrics, 8 (1966), pp. 591-613.

[6] R. E. BARLOW and S. S. GUPTA, Selection procedures for restricted families of probability distributions, Ann. Math. Statist., 40 (1969), pp. 905-917.

[7] R. E. BARLOW, S. S. GUPTA and S. PANCHAPAKESAN, On the distribution of the maximum and minimum of ratios of order statistics,

Ann. Math. Statist., 40 (1969), pp. 918-934.

[8] R. E. BARLOW and A. W. MARSHALL, Tables of bounds for distributions with monotone hazard rate, J. Amer. Statist. Assoc., 60 (1965), pp. 872-890.

[9] R. E. BARLOW, A. W. MARSHALL and F. PROSCHAN, Properties of probabily distributions with monotone hazard rate, Ann. Math. Statist., 34 (1963), pp. 375-389.

[10] R. E. BARLOW and F. PROSCHAN, Mathematical Theory of Reliability, John Wiley, New York, 1965.

[11] R. E. BARLOW and F. PROSCHAN, Inequalities for linear combinations of order statistics from restricted families, Ann. Math. Statist., 37 (1966), pp. 1574-1592.

[12] R. E. BARLOW and F. PROSCHAN, Tolerance and confidence limits for classes of distributions based on failure rate, Ann. Math. Statist., 37 (1966), pp. 1593-1601.

[13] R. E. BARLOW and F. PROSCHAN, Exponential life test procedures when the distribution has monotone failure rate, J. Amer. Statist. Assoc., 62 (1967), pp. 548-560.

[14] R. E. BARLOW and F. PROSCHAN, A note on tests for monotone failure rate based on incomplete data, Ann. Math. Statist., 40 (1969), pp. 595-600.

[15] R. E. BARLOW and E. M. SCHEUER, Estimation from accelerated life tests, Technometrics, 13 (1971), pp. 145-159.

[16] R. E. BARLOW and W. R. VAN ZWET, Asymptotic properties of

isotonic estimators for the generalized failure rate function. Part II: asymptotic distributions, Rep. ORC 69-10, Operations Research Center, Univ. of California, Berkeley, California, 1969.

[17] R. E. BARLOW and W. R. VAN ZWET, Asymptotic properties of isotonic estimators for the generalized failure rate function. Part I: strong consistency, Nonparametric Techniques in Statistical Inference, M. L. Puri, Cambridge University Press, London, 1970, pp. 159-173.

[18] D. J. BARTHOLOMEW, A problem in life testing, J. Amer. Statist. Assoc., 52 (1957), pp. 350-355.

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

[20] R. E. BECHHOFER, J. KIEFER and M. SOBEL, Sequential Identification and Ranking Procedures, Statistical Research Monographs, The University of Chicago Press, Chicago, 1968.

[21] E. F. BECKENBACH and R. BELLMAN, Inequalities, Springer-Verlag, Berlin, 1961.

[22] P. J. BICKEL, Tests for monotone failure rate II, Ann. Math. Statist., 40 (1969), pp. 1250-1260.

[23] P. J. BICKEL and K. A. DOKSUM, Tests for monotone failure rate based on normalized spacings, Ann. Math. Statist. 40 (1969), pp. 1216-1235.

[24] Z. W. BIRNBAUM and S. C. SAUNDERS, A statistical model for life-length of materials, J. Amer. Statist. Assoc., 53 (1958), pp. 151-163.

[25] T. A. BRAY, G. B. CRAWFORD and F. PROSCHAN, Maximum likelihood estimation of a U-shaped failure rate function, Document D1-82-0660, Boeing Scientific Research Laboratories, Seattle, Washington, 1967.

[26] A. M. BRUCKNER and E. OSTROW, Some function classes related to the class of convex functions, Pacific J. Math., 12 (1962), pp. 1203-1215.

[27] W. R. BUCKLAND, Statistical Assessment of the Life Characteristic, Griffin, London, 1964.

[28] K. DOKSUM, Starshaped transformations and the power of rank tests, Ann. Math. Statist., 40 (1969), pp. 1167-1176.

[29] E. J. DUDEWICZ, Maximum likelihood estimators for ranked means, Z. Wahrscheinlichkeitstheorie Verw. Gebiete, 19 (1971), pp. 29-42.

[30] E. J. DUDEWICZ and Y. L. TONG, Optimal confidence intervals for the largest location parameter, Statistical Decision Theory and Related Topics, S. S. GUPTA and J. YACKEL, Academic Press, New York, 1971, pp. 363-376.

[31] B. EPSTEIN, Truncated life tests in the exponential case, Ann. Math. Statist., 25 (1954), pp. 555-564.

[32] B. EPSTEIN, Simple estimators of the parameters of exponential

distributions when samples are censored, Ann. Inst. Statist. Math., 8 (1956), pp. 15-26.

[33] B. EPSTEIN, Statistical life test acceptance procedures, Technometrics, 2 (1960), pp. 435-446.

[34] B. EPSTEIN, Estimation from life test data, Technometrics, 2 (1960), pp. 447-454.

[35] B. EPSTEIN and M. SOBEL, Life testing, J. Amer. Statist. Assoc., 48 (1953), pp. 486-502.

[36] B. EPSTEIN and M. SOBEL, Some theorems relevant to life testing from an exponential distribution, Ann. Math. Statist., 25 (1954), pp. 373-381.

[37] B. EPSTEIN and M. SOBEL, Sequential life tests in the exponential case, Ann. Math. Statist., 26 (1955) pp. 82-93.

[38] B. EPSTEIN and C. K. TSAO, Some tests based on ordered observations from two exponential populations, Ann. Math. Statist., 24 (1953), pp. 458-466.

[39] J. D. ESARY, A. W. MARSHALL and F. PROSCHAN, Some reliability applications of the hazard transform, SIAM J. Appl. Math., 18 (1970), pp. 849-860.

[40] M. GNANADESIKAN and S. S. GUPTA, Selection procedures for multivariate normal distributions in terms of measures of dispersion. Technometrics 12 (1970), pp. 103-117.

[41] Z. GOVINDARAJULU, A supplement to Mendenhall's bibliography on life testing and related topics, J. Amer. Statist. Assoc.,

59 (1964), pp. 1231-1291.

[42] U. GRENANDER, On the theory of mortality measurement, Part II, Skand. Aktuarietidskr., 39 (1956), pp. 125-153.

[43] E. J. GUMBEL, Statistics of Extremes, Columbia University Press, New York, 1958.

[44] S. S. GUPTA, On a decision rule for a problem in ranking means, Mimeo. Series No. 150, Institute of Statistics, Univ. of North Carolina, Chapel Hill, North Carolina, 1956.

[45] S. S. GUPTA, Life test sampling plans for normal and lognormal distributions, Technometrics, 4 (1962), pp. 151-175.

[46] S. S. GUPTA, Probability integrals of the multivariate normal and multivariate t. Ann. Math. Statist. 34 (1963), pp. 792-828.

[47] S. S. GUPTA, On a selection and ranking procedure for gamma populations, Ann. Inst. Statist. Math., 14 (1963), pp. 199-216.

[48] S. S. GUPTA, On some multiple decision (selection and ranking) rules, Technometrics, 7 (1965), pp. 225-245.

[49] S. S. GUPTA, On some selection and ranking procedures for multivariate normal populations using distance functions, Multivariate Analysis, P. R. KRISHNAIAH, Academic Press, New York, 1966, pp. 457-475.

[50] S. S. GUPTA and P. A. GROLL, Gamma distribution in acceptance sampling based on life tests, J. Amer. Statist. Assoc., 56 (1961), pp. 942-970.

[51] S. S. GUPTA and WEN-TAO HUANG, A note on selecting a subset of normal populations with unequal sample sizes, Sankhyā Ser. A, to appear.

[52] S. S. GUPTA and G. C. McDONALD, On some classes of selection procedures based on ranks, Nonparametric Techniques in Statistical Inference, M. L. PURI, Cambridge University Press, London, 1970, pp. 491-514.

[53] S. S. GUPTA and K. NAGEL, On selection and ranking procedures and order statistics from the multinomial distribution, Sankhyā Ser. B, 29 (1967), pp. 1-34.

[54] S. S. GUPTA and K. NAGEL, On some contributions to multiple decision theory, Statistical Decision Theory and Related Topics, S. S. GUPTA and J. YACKEL, Academic Press, New York, 1971, pp. 79-102.

[55] S. S. GUPTA, K. NAGEL and S. PANCHAPAKESAN. On the order statistics from equally correlated normal random variables, Biometrika, 60 (1973), to appear.

[56] S. S. GUPTA and S. PANCHAPAKESAN, Some selection and ranking procedures for multivariate normal populations, Multivariate Analysis II, P. R. KRISHNAIAH, Academic Press, New York, 1969, pp. 475-505.

[57] S. S. GUPTA and S. PANCHAPAKESAN, On a class of subset selection procedures, Ann. Math. Statist., 43 (1972), pp. 814-822.

[58] S. S. GUPTA and S. PANCHAPAKESAN, On multiple decision procedures,

J. Mathematical and Physical Sci., 6 (1972), pp. 1-72.

[59] S. S. GUPTA and S. PANCHAPAKESAN, On order statistics and some applications of combinatorial methods in statistics, A Survey of Combinatorial Theory, J. N. SRIVASTAVA, American Elserier, New York, 1973, pp. 217-250.

[60] S. S. GUPTA and T. J. SANTNER, Selection of a restricted subset of normal populations containing the one with largest mean, Mimeo. Series No. 299, Department of Statistics, Purdue Univ., West Lafayette, Indiana, 1972.

[61] S. S. GUPTA and M. SOBEL, On a statistic which arises in selection and ranking problems, Ann. Math. Statist., 28 (1957), pp. 957-967.

[62] S. S. GUPTA and M. SOBEL, On selecting a subset which contains all populations better than a standard, Ann. Math. Statist., 29 (1958), pp. 235-244.

[63] S. S. GUPTA and M. SOBEL, Selecting a subset containing the best of several binomial populations, Contributions to Probability and Statistics, I. OLKIN, Stanford University Press, Stanford. California, 1960, pp. 224-248.

[64] S. S. GUPTA and M. SOBEL, On selecting the subset containing the population with the smallest variance, Biometrika, 49 (1962), pp. 495-507.

[65] S. S. GUPTA and M. SOBEL, On the smallest of several correlated F-statistics, Biometrika, 49 (1962), pp. 509-523.

[66] S. S. GUPTA and W. J. STUDDEN, On some selection and ranking procedures with applications to multivariate populations, Essays in Probability and Statistics, R. C. BOSE, University of North Carolina Press, Chapel Hill, North Carolina, 1970, pp. 327-338.

[67] G. H. HARDY, J. E. LITTLEWOOD and G. PÓLYA, Inequalities, Cambridge University Press, London, 1959.

[68] M. HOLLANDER and F. PROSCHAN, Testing whether new is better than used, Ann. Math. Statist., 43 (1972), pp. 1136-1146.

[69] J. H. K. KAO, A graphical estimation of mixed Weibull parameters in life testing electron tubes, Technometrics, 1 (1959), pp. 389-407.

[70] M. G. KENDALL, A new measure of rank correlation, Biometrika, 30 (1938), pp. 81-93.

[71] J. KIEFER and J. WOLFOWITZ, Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. Ann. Math. Statist., 27 (1956), pp. 887-906.

[72] M. J. LAWRENCE, Inequalities and tolerance limits for s-ordered distributions, Rep. ORC 66-37, Operations Research Center, Univ. of California, Berkeley, California, 1966.

[73] E. L. LEHMANN, Some Model I problems of selection, Ann. Math. Statist., 32 (1961), pp. 990-1012.

[74] H. B. MANN, Nonparametric tests against trend, Econometrica, 13 (1945), pp. 245-249.

[75] A. W. MARSHALL, I. OLKIN and F. PROSCHAN, Monotonicity of ratios of means and other applications of majorization, Inequalities-I, O. SHISHA, Academic Press, New York, 1967, pp. 177-203.

[76] A. W. MARSHALL and F. PROSCHAN, Maximum likelihood estimation for distributions with monotone failure rate, Ann. Math. Statist., 36 (1965), pp. 69-77.

[77] A. W. MARSHALL and F. PROSCHAN, Classes of distributions applicable in replacement with renewal theory implications, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, University of California Press, Berkeley and Los Angeles, California, 1972, pp. 395-415.

[78] W. MENDENHALL, A bibliography on life testing and related topics, Biometrika, 45 (1958), pp. 521-543.

[79] K. NAGEL, On subset selection rules with certain optimality properties, Mimeo. Series No. 222, Department of Statistics, Purdue University, West Lafayette, Indiana, 1970.

[80] S. PANCHAPAKESAN, On a subset selection procedure for the most probable event in a multinomial distribution, Statistical Decision Theory and Related Topics, S. S. GUPTA and J. YACKEL, Academic Press, New York, 1971, pp. 275-298.

[81] E. PARZEN, On estimation of a probability density and mode, Ann. Math. Statist., 33 (1962), pp. 1065-1076.

[82] R. L. PLACKETT, Analysis of life test data, Techrometrics,

1 (1959), pp. 9-19.

[83] B. L. S. PRAKASA RAO, Estimation for distributions with monotone failure rate, Ann. Math. Statist., 41 (1970), pp. 507-519.

[84] F. PROSCHAN and R. PYKE, Tests for monotone failure rate, Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III, University of California Press, Berkeley and Los Angeles, 1967, pp. 293-312.

[85] R. PYKE, Spacings, J. Roy. Statist. Soc. Ser. B, 7 (1965), pp. 395-449.

[86] R. PYKE, Spacings revisited, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, University of California Press, Berkeley and Los Angeles, California, 1972, pp. 417-427.

[87] M. H. RIZVI and M. SOBEL, Nonparametric procedures for selecting a subset containing the population with the largest α-quantile, Ann. Math. Statist., 38 (1967), pp. 1788-1803.

[88] M. H. RIZVI, M. SOBEL and G. C. WOODWORTH, Nonparametric ranking procedures for comparison with a control, Ann. Math. Statist., 39 (1968), 2075-2093.

[89] T. J. SANTNER, A restricted subset selection approach to ranking and selection problems, Mimeo. Series No. 318, Department of Statistics, Purdue University, West Lafayette, Indiana, 1973.

[90] M. SOBEL and J. A. TICHENDORF, Acceptance sampling with new life test objectives, Proc. of Fifth National Symposium on

Unclassified

Security Classification

**DOCUMENT CONTROL DATA - R&D**

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

| | |
|---|---|
| 1. ORIGINATING ACTIVITY (Corporate author) Purdue University | 2a. REPORT SECURITY CLASSIFICATION Unclassified |
| | 2b. GROUP |

3. REPORT TITLE Some Multiple Decision and Related Problems with Special Reference to Restricted Families of Distributions and Applications to Reliability Theory

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report, July, 1973

5. AUTHOR(S) (Last name, first name, initial)
Gupta, Shanti S. and Panchapakesan, S.

| 6. REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| July 1973 | 92 | 95 |

| | |
|---|---|
| 8a. CONTRACT OR GRANT NO. N00014-67-A-0226-00014 | 9a. ORIGINATOR'S REPORT NUMBER(S) Mimeo Series #324 |
| b. PROJECT NO. | |
| c. | 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) |
| d. | |

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited

| | |
|---|---|
| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY Office of Naval Research Washington, D.C. |

13. ABSTRACT

This paper deals with procedures for selecting a subset from k given populations so as to include the "best" with a specified guaranteed minimum probability. Some general results relating to subset selection and specific procedures for important classes of distributions are reviewed with special emphasis on restricted families of probability distributions. Such families are defined through partial order relations and are extensively considered in reliability theory. A selection problem for tail-ordered family of distributions is considered and tables are provided for constants needed to implement the procedure. A general partial order relation is defined through a class of real valued functions and a related selection problem is discussed. These results provide a unified view of earlier known results. The rest of the paper gives a brief survey of some important results pertaining to restricted families of distributions such as the star-ordered and convex-ordered distributions. These results relate to life test sampling plans, inequalities for linear combinations of order statistics, estimation of failure rate function and some tests of hypotheses.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification