*The Annals of Statistics*
1979, Vol. 7, No. 3, 615–643

# DESIGN OF EXPERIMENTS FOR SELECTION FROM ORDERED FAMILIES OF DISTRIBUTIONS[1]

BY JEFFREY H. HOOPER AND THOMAS J. SANTNER[2]
*Bell Laboratories, Inc., and*
*Cornell University*

Given $G \in \mathcal{S}$ a space of cumulative distribution functions and a weak (reflexive and transitive) order relation $\prec$ on $\mathcal{S}$, the subclass of $\mathcal{S}$ given by $\{F \in \mathcal{S} | F < G\}$ is called an ordered family of distributions. Suppose $\pi_1, \ldots, \pi_k$ represent $k$ populations with distributions only known to belong to some specified ordered family. The general problem is to design an experiment to select $\pi_i$'s having large (small) $\alpha$-quantities. A preferred population is defined to be any population with $\alpha$-quantile "near" the $t$th largest $\alpha$-quantile and a correct selection occurs if the subset of populations selected contains at least a prespecified number, $r$, of preferred populations. The design problem is solved for both fixed and random subset size selection procedures under star and tail ordering. Tables of the sample sizes required to guarantee prespecified minimum probabilities of correct selection are given for the case of selecting from continuous IFRA populations. Comparisons are made with the optimal procedure for selecting from exponential populations. Properties of the proposed rules are discussed.

**1. Introduction.** Given a space $\mathcal{S}$ of cumulative distribution functions (cdf's) and a weak (reflexive and transitive) order relation, $\prec$, on $\mathcal{S}$, a subclass of $\mathcal{S}$ is called an ordered family of distributions (with respect to $G \in \mathcal{S}$) if it has the form $\{F \in \mathcal{S} | F \prec G\}$. Two familiar examples of weak orderings are convex ordering and star ordering which yield the increasing failure rate (IFR) and increasing failure rate on the average (IFRA) families when $G(x) = 1 - e^{-x}$.

Weak orderings and the ordered family model assumption have proved to be of considerable importance in the literature: (a) for providing characterizations of skewness and kurtosis (Van Zwet (1964)) (b) in power studies of certain rank tests (Doksum (1969)), (c) as models for the lifetimes of coherent systems (Birnbaum, Esuary and Marshall (1966)) and (d) as models for the lifetimes of systems subject to random shocks (Esuary, Marshall and Proschan (1973) and Barlow and Proschan (1975)).

This paper studies the problem of designing a single stage experiment for the selection of the distribution(s) with the largest $\alpha$-quantile(s) where $\alpha \in (0, 1)$ is given and when the model assumptions are specified by an ordered family of distributions. More precisely, suppose observations are to be taken from $k$ populations labeled $\Pi_1, \cdots, \Pi_k$. For each $i$ in $\{1, \cdots, k\}$ (a) let $F_i$ denote the cdf of

Received September 1977.
[1]Research supported by National Science Foundation Grant ENG75-10487, Office of Naval Research Contract N00014-75-C-0586 and U.S. Army Research Office Contracts DAAG29-73-C-0008 and DAAG29-77-C-0003.
[2]Second author on leave at Biometry Br., National Cancer Institute.
*AMS* 1970 *subject classifications.* Primary 62G99; secondary 62K99.
*Key words and phrases.* Star ordering, tail ordering, fixed size selection procedure, restricted subset selection procedure.

population $\Pi_i$, (b) $\{X_{ij} : 1 \leqslant j \leqslant n\}$ be a random sample from $\Pi_i$, (c) $x_\alpha(F_i)$ be the $\alpha$-quantile of $F_i$, (d) $F_{[i]}$ be the cdf with the $i$th smallest $\alpha$-quantile, and (e) $\Pi_{(i)}$ denote the (unknown) population with cdf $F_{[i]}$. Then $x_\alpha(F_{[1]}) \leqslant \cdots \leqslant x(F_{[k]})$ are the ordered $\alpha$-quantiles. Assume samples from different $\Pi_i$'s are independent, there is no prior knowledge of the pairing of $\Pi_i$ and $\Pi_{(i)}$, the $F_i$'s come from a specified ordered family of distributions, and the experimenter is interested in selection of $\Pi_{(k-t+1)}, \cdots, \Pi_{(k)}$, the $t$ populations with largest $\alpha$-quantiles. We adopt the preferred population formulation in which $\Pi_i$ is called a preferred population if $x_\alpha(F_i)$ is "close" (in a sense made precise in Section 3) to $x_\alpha(F_{[k-t+1]})$ and consider two classes of selection problems:

1. selection of a subset of prespecified size $s(r \leqslant t, r \leqslant s < k - t + r)$ so as to include at least $r$ preferred populations; and
2. selection of a random size subset containing at most $s(1 \leqslant s \leqslant k)$ populations so as to include at least one ($r = 1$) preferred population.

The problem is to design the experiment so that (1) or (2) is guaranteed with at least prespecified probability $P^*$ no matter what the true vector of cdf's. This approach combines aspects of Gupta's subset selection formulation (1956, 1965) and Bechhofer's indifference zone formulation (1954) and will be shown to be a strengthening of the latter.

Previous work on selection procedures for nonparametric problems can be categorized according to (1) the parameter of interest for selection, (2) the formulation of the design requirement and (3) the model assumptions. One group of papers considers problems of selection for location parameters or more complicated linear models usually under the assumption that the error distributions are continuous and symmetric. They include: Lehmann (1963), Bartlett and Govindarajulu (1968), Randles (1970), Bhapkar and Gore (1971), Puri and Sen (1972), Ghosh (1973) and Gupta and Huang (1974). Related papers include Patel (1976) who considers the problem of selecting the distribution with the largest scale parameter when the common form of the distribution is assumed to be IFR and Gupta and McDonald (1969) who consider selection in terms of an arbitrary parameter under which the distributions form a stochastically increasing family of distributions. A second group of papers are those which consider selection in terms of $\alpha$-quantiles. Rizvi and Sobel (1967), Rizvi, Sobel and Woodworth (1968) and Barlow and Gupta (1969) study subset selection formulations of the problem while Sobel (1967) and Desu and Sobel (1971) use an indifference zone formulation. For a complete review of the literature see Lee and Dudewicz (1974).

In contrast, the present work does not assume the $F_i$'s are stochastically ordered by a location, scale or other parameter as do the papers in the first group nor does it assume they are arbitrary continuous cdf's as do many of the papers in the second group. Rather the ordered family assumption falls between these two extremes.

Section 2 studies weak orderings and the corresponding ordered families. Sections 3 and 4 formulate a fixed size subset selection problem and a random size subset selection problem respectively and give their solution for two classes of distributions. Finally Section 5 describes some properties of the two classes of procedures considered in Sections 3 and 4.

**2. Weak orderings and ordered families of distributions.** Given a finite or infinite interval, $I$, of the real line, $R$, let $\mathcal{F}_I$ denote the set of distributions $F$ which satisfy:

(1) $F$ is absolutely continuous with respect to Lebesgue measure;

(2) $S(F) \equiv \{x \in R \mid F(x+\varepsilon) - F(x-\varepsilon) > 0 \forall \varepsilon > 0\}$, the support of $F$, is a convex set;

(3) $S(F) \subset I$.

In particular when $I = R$ and $[0, \infty)$, $\mathcal{F}_I$ will be denoted as $\mathcal{F}$ and $\mathcal{F}_0$ respectively. For $F \in \mathcal{F}_I$ define $F^{-1}(y)$ on $[0, 1]$ as $F^{-1}(y) \equiv \inf\{x \in R \mid F(x) \geqslant y\}$ where the infimum of the empty set is taken to be $+\infty$.

REMARK 2.1. For $F, G \in \mathcal{F}_I$ the following hold.

(i) $S(F)$ is a closed interval of $R$;

(ii) $I(F) \equiv \{x \in R \mid 0 < F(x) < 1\}$ is an open interval of $R$;

(iii) $F$ is strictly increasing on the open interval $I(F)$;

(iv) $F^{-1}(\cdot)$ is continuous and strictly increasing on $(0, 1)$;

(v) $\phi(x) \equiv G^{-1}(F(x))$ is the unique strictly increasing function on $I(F)$ such that if $X$ has cdf $F$ then $\phi(X)$ has cdf $G$.

A weak ordering, $\prec$, on $\mathcal{F}_I$ is a relation on $\mathcal{F}_I$ which is reflexive and transitive. Every weak ordering $\prec$ on $\mathcal{F}_I$ can be "extended" to a partial order by first defining an equivalence relation $\simeq$ on $\mathcal{F}_I$ by $F \simeq G \Leftrightarrow F \prec G$ and $G \prec F$. The resulting set of equivalence classes $\{E(F) \mid F \in \mathcal{F}_I\}$ is partially ordered by the relation $\preceq$ which is well defined by $E(F) \preceq E(G) \Leftrightarrow F \prec G$. We now define a class of weak orderings which is a modification of a class proposed by Panchapakesan (1969).

Let $\mathcal{H}$ be a set of functions $h : I^2 \to I$ satisfying

$$h(x_1, x_2) \leqslant \max\{x_1, x_2\} \forall (x_1, x_2) \in I^2.$$

DEFINITION 2.1. For $F, G \in \mathcal{F}_I$, *$F$ is said to be $\mathcal{H}$-ordered w.r.t. $G$* (written $F \prec_{\mathcal{H}} G) \Leftrightarrow G^{-1}F(h(x_1, x_2)) \leqslant h(G^{-1}F(x_1), G^{-1}F(x_2)) \forall h \in \mathcal{H}$ and $x_1, x_2$ such that $x_1, x_2$ and $h(x_1, x_2) \in I(F)$.

LEMMA 2.1. *$\mathcal{H}$ ordering is a weak ordering on $\mathcal{F}_I$.*

PROOF. Reflexivity is straightforward. To show transitivity suppose $F \prec_{\mathcal{H}} G$, $G \prec_{\mathcal{H}} J$ and $h \in \mathcal{H}$ and $x_1, x_2$ satisfy $x_1, x_2$ and $h(x_1, x_2) \in I(F)$. Then $G^{-1}F(x_1)$, $G^{-1}F(x_2)$ and $G^{-1}F(h(x_1, x_2)) \in I(G)$ and hence

$$G^{-1}Fh(x_1, x_2) \leqslant h\big(G^{-1}F(x_1), G^{-1}F(x_2)\big) \leqslant \max\big\{G^{-1}F(x_1), G^{-1}F(x_2)\big\} \tag{2.1}$$

implies $h(G^{-1}F(x_1), G^{-1}F(x_2)) \in I(G)$ by convexity. So

$$\begin{aligned} J^{-1}F(h(x_1, x_2)) &= J^{-1}GG^{-1}F(h(x_1, x_2)) \\ &\leqslant J^{-1}G\big(h\big(G^{-1}F(x_1), G^{-1}F(x_2)\big)\big) \quad \text{by } (2.1) \\ &\leqslant h\big(J^{-1}F(x_1), J^{-1}F(x_2)\big) \quad \text{since } G \prec_{\mathcal{H}} J \end{aligned}$$

and the proof is completed.

The motivation for introducing $\mathcal{H}$-ordering is that it provides a convenient framework for embedding several well-known orderings.

EXAMPLE 2.1. Let $I = [0, \infty)$ and $\mathcal{H}_1$ consist of the single function

$$\begin{aligned} h(x_1, x_2) &= x_1 - x_2 \quad , x_1 > x_2; \\ &= 0 \quad , x_1 \leqslant x_2. \end{aligned}$$

Now $h(x_1, x_2) \leqslant \max\{x_1, x_2\}$ for any $x_1, x_2 \geqslant 0$ and so $\mathcal{H}_1$ is a weak ordering on $\mathcal{F}_0$. Furthermore

$$\begin{aligned} F \prec_{\mathcal{H}_1} G &\Leftrightarrow G^{-1}F(t - x) \leqslant G^{-1}F(t) - G^{-1}F(x) \quad \text{whenever} \quad x, t, t - x \in I(F); \\ &\Leftrightarrow G^{-1}F \quad \text{is superadditive on} \quad I(F). \end{aligned}$$

In particular $F \prec_{\mathcal{H}_1} G(x) = 1 - e^{-x} \Leftrightarrow \bar{F}(t - x)\bar{F}(x) \geqslant \bar{F}(t)$ whenever $x$, $t$ and $t - x \in I(F)$ where $\bar{F}(y) = 1 - F(y) \Leftrightarrow F$ is an absolutely continuous NBU (new better than used) distribution with convex support.

Before giving further examples recall the following definitions. For $F, G \in \mathcal{F}_0$(1) $F$ is *convex ordered* w.r.t. $G(F \prec_c G)$ means $G^{-1}F(x)$ is convex on $I(F)$ and (2) $F$ is *star ordered* w.r.t. $G(F <_* G)$ means $G^{-1}F(x)$ is starshaped on $I(F)$, i.e., $G^{-1}F(\lambda x) \leqslant \lambda G^{-1}F(x)$ whenever $x, \lambda x \in I(F)$ and $\lambda \in [0, 1]$. Convex ordering was introduced as an alternative to the standarized third moment inequality definition of skewness. Namely $F <_c G$ is interpreted to mean "$G$ is more skewed to the right than $F$" (see Van Zwet (1964) page 9). Since the star-shaped property can be thought of as a weakening of convexity (see Bruckner and Ostrow (1962)), $F <_* G$ can also be interpreted as an ordering according to skewness. For $F, G \in \mathcal{F}$, $F$ is said to be *tail ordered* w.r.t. $G(F \prec_t G)$ iff $G^{-1}F(x) - x$ is nondecreasing on $I(F)$. When $F \prec_t G$ then $G$ is interpreted as having heavier tails than $F$ (see Doksum (1969)).

EXAMPLE 2.2. Take $I = [0, \infty)$ and $\mathcal{H}_2 = \{h_\lambda | 0 \leqslant \lambda \leqslant 1\}$ where $h_\lambda(x_1, x_2) \equiv \lambda x_1 + (1 - \lambda)x_2 \leqslant \max\{x_1, x_2\} \forall \lambda \in [0, 1]$; hence $\mathcal{H}_2$ ordering is a weak ordering. It is straightforward to check that $F \prec_{\mathcal{H}_2} G \Leftrightarrow G^{-1}F(\cdot)$ is convex on $I(F)$. So $\mathcal{H}_2$ yields convex ordering.

EXAMPLE 2.3. Again let $I = [0, \infty)$ and define $\mathcal{H}_3 = \{h_\lambda | 0 \leqslant \lambda \leqslant 1\}$ where $h_\lambda(x_1, x_2) = \lambda x_1 \leqslant \max\{x_1, x_2\} \forall 0 \leqslant \lambda \leqslant 1$ and so $\mathcal{H}_3$ defines a weak ordering. It is straightforward to verify that $\mathcal{H}_3$ yields star ordering.

EXAMPLE 2.4. Choose $I = R$ and $\mathcal{H}_4 = \{h_\delta | 0 \leqslant \delta < \infty\}$ where $h_\delta(x_1, x_2) = x_1 - \delta \leqslant \max\{x_1, x_2\} \forall \delta \in [0, \infty)$; $\mathcal{H}_4$ ordering is a weak ordering which reduces to tail ordering.

The remainder of this section is devoted to deriving some properties of star and tail ordering which will be used in the later sections.

REMARK 2.2. It is straightforward to show that the set of distributions equivalent to $F \in \mathcal{F}_0$ under star ordering ($F \in \mathcal{F}$ under tail ordering) is $E_*(F) \equiv \{F(\delta x) | 0 < \delta < \infty\}$ ($E_t(F) \equiv \{F(x + \beta) | \beta \in R\}$), the set of all scale changes (location shifts) of $F$.

LEMMA 2.2. (a) *For* $F, G \in \mathcal{F}_0$, $F \prec_* G \Leftrightarrow G^{-1}F(x)/x$ *is nondecreasing on* $R^+(F) \equiv \{x > 0 | F(x) < 1\}$.

(b) *For* $F, G \in \mathcal{F}$, $F \prec_t G \Leftrightarrow G^{-1}F(x) - x$ *is nondecreasing on* $R(F) \equiv \{x \in R | F(x) < 1\}$.

PROOF. Write $R^+(F)$ as the disjoint union $\{x > 0 | F(x) = 0\} \cup I(F)$ and $\forall x \in \{x > 0 | F(x) = 0\}$, $G^{-1}F(x)/x = G^{-1}(0)/x = -\infty$ and the result follows since $G^{-1}F(x)/x$ is nondecreasing on $I(F)$. The reverse implication is obvious and the proof for $\prec_t$ is similar.

An important class of distributions preserving weak orderings are the distributions of the order statistics. If $X_{(q)} (1 \leqslant q \leqslant n)$ is the $q$th order statistic based on $X_1, \cdots, X_n$ i.i.d. $F$ then $P_F[X_{(q)} \leqslant x] = B(F(x); q, n)$ where

$$B(p; q, n) = \frac{n!}{(n-q)!\,(q-1)!} \int_0^p x^{q-1}(1-x)^{n-q} dx.$$

THEOREM 2.1. *If* $F, G \in \mathcal{F}_I$ *and* $F \prec_{\mathcal{H}} G$ *then* $B(F; q, n) \prec_{\mathcal{H}} B(G; q, n)$ *holds* $\forall 1 \leqslant q \leqslant n$.

The proof is immediate from the definition of $\mathcal{H}$ ordering since $G^{-1}F(x) = G^{-1}B^{-1}BF(x)$ where $B(\cdot)$ and $B^{-1}(\cdot)$ denote $B(x; q, n)$ and its inverse respectively.

The final part of this section will describe stochastic bounds for several classes of distributions. For $G$ in $\mathcal{F}_0$ with $I(G) = (0, \infty)$ and $0 < \xi < \infty$ define

$$\mathcal{F}_*(G) = \{F \in \mathcal{F}_0 | F <_* G\} \tag{2.2}$$

$$\mathcal{F}_*(G, \xi) = \{F \in \mathcal{F}_*(G) | x_\alpha(F) = \xi\}. \tag{2.3}$$

REMARK 2.3. $\mathcal{F}_*(G) = \mathcal{F}_*(G_\lambda) \forall \lambda > 0$ where $G_\lambda(x) \equiv G(\lambda x)$ since $\forall \lambda > 0$, $G_\lambda \in E_*(G)$ and $F \in \mathcal{F}_*(G) \Leftrightarrow F \prec_* G \Leftrightarrow F \prec_* G_\lambda \Leftrightarrow F \in \mathcal{F}_*(G_\lambda)$.

EXAMPLE 2.5. Let $G(x) = 1 - e^{-x}$ or 0 as $x > 0$ or $x \leqslant 0$ respectively. Then $G^{-1}(y) = -\ln(1-y)$ for $0 \leqslant y < 1$ and $F \in \mathcal{F}_*(G) \Leftrightarrow F$ is continuous on $[0, \infty)$ and $-\ln(1 - F(x))/x$ is nondecreasing on $R^+(F) \Leftrightarrow F$ is a continuous IFRA distribution. In particular $\mathcal{F}_*(G) = \mathcal{F}_*(G_\lambda) \forall \lambda > 0$ where $G_\lambda(x) = 1 - e^{-\lambda x}$, $x > 0$.

Associate with $\mathcal{F}_*(G)$ the families of distribution

$$F_{*,s}(x|\xi) = G(xG^{1}(\alpha)/\xi), \qquad x < \xi \tag{2.4}$$
$$= 1, \qquad x \geqslant \xi$$
$$F_{*,l}(x|\xi) = 0, \qquad x < \xi \tag{2.5}$$
$$= G(xG^{-1}(\alpha)/\xi), \qquad x \geqslant \xi$$

where $0 < \xi < \infty$.

REMARK 2.4. $G^{-1}F_{*,s}(\cdot|\xi)$ and $G^{-1}F_{*,l}(\cdot|\xi)$ are starshaped on $R^+(F_{*,s}(\cdot|\xi))$ and $R^+(F_{*,l}(\cdot|\xi))$ respectively and $x_\alpha(F_{*,s}(\cdot|\xi)) = x_\alpha(F_{*,l}(\cdot|\xi)) = \xi$. However, neither $F_{*,s}(\cdot|\xi)$ nor $F_{*,l}(\cdot|\xi)$ is in $\mathcal{F}_0$ and consequently neither is in $\mathcal{F}_*(G)$.

THEOREM 2.2. *For all* $0 < \xi < \infty$ *and* $F \in \mathcal{F}_*(G, \xi)$,

$$F_{*,s}(\cdot|\xi) \prec_{st} F \prec_{st} F_{*,l}(\cdot|\xi).$$

PROOF. If $R^+(F) = (0, M)$ then $x \leqslant 0 \Rightarrow F(x) = G(x) = F_{*,s}(x|\xi) = F_{*,l}(x|\xi) = 0$ and $\forall 0 < x < \xi \leqslant y < M$

$$G^{-1}F(x)/x \leqslant G^{-1}F(\xi)/\xi = G^{-1}(\alpha)/\xi \leqslant G^{-1}F(y)/y$$
$$\Leftrightarrow F(x) \leqslant G(xG^{-1}(\alpha)/\xi) \leqslant G(yG^{-1}(\alpha)/\xi) \leqslant F(y).$$

If $M = +\infty$ the proof is complete, while if $M < +\infty$ then $\forall y \geqslant M,\ F(y) = 1 \geqslant G(yG^{-1}(\alpha)/\xi)$ and the result follows. See Figure 2.1.

Since both $\{F_{*,s}(\cdot|\xi)|0 < \xi < \infty\}$ and $\{F_{*,l}(\cdot|\xi)|0 < \xi < \infty\}$ are (stochastically increasing) scale parameter families the result can be slightly strengthened as follows.

COROLLARY 2.1. *For* $0 < \xi' \leqslant \xi \leqslant \xi''$

(a) $F \in \mathcal{F}_*(G, \xi') \Rightarrow F \prec_{st} F_{*,l}(\cdot|\xi)$; *and*

(b) $F \in \mathcal{F}_*(G, \xi'') \Rightarrow F_{*,s}(\cdot|\xi) \prec_{st} F$.

REMARK 2.5. It can be shown that the bounds $F_{*,s}(\cdot|\xi)$ and $F_{*,l}(\cdot|\xi)$ are tight in the sense that they are the stochastically largest and smallest distributions respectively satisfying Corollary 2.1.

The analagous results for tail ordered families are as follows. For $G \in \mathcal{F}$ with $I(G) = R$ and $\xi \in R$ define

$$\mathcal{F}_t(G) = \{F \in \mathcal{F}|F \prec_t G\}$$
$$\mathcal{F}_t(G, \xi) = \{F \in \mathcal{F}_t(G)|x_\alpha(F) = \xi\}$$
$$F_{t,s}(x|\xi) = G(x + G^{-1}(\alpha) - \xi), \qquad x < \xi$$
$$= 1, \qquad x \geqslant \xi$$
$$F_{t,l}(x|\xi) = 0, \qquad x < \xi$$
$$= G(x + G^{-1}(\alpha) - \xi), \qquad x \geqslant \xi.$$

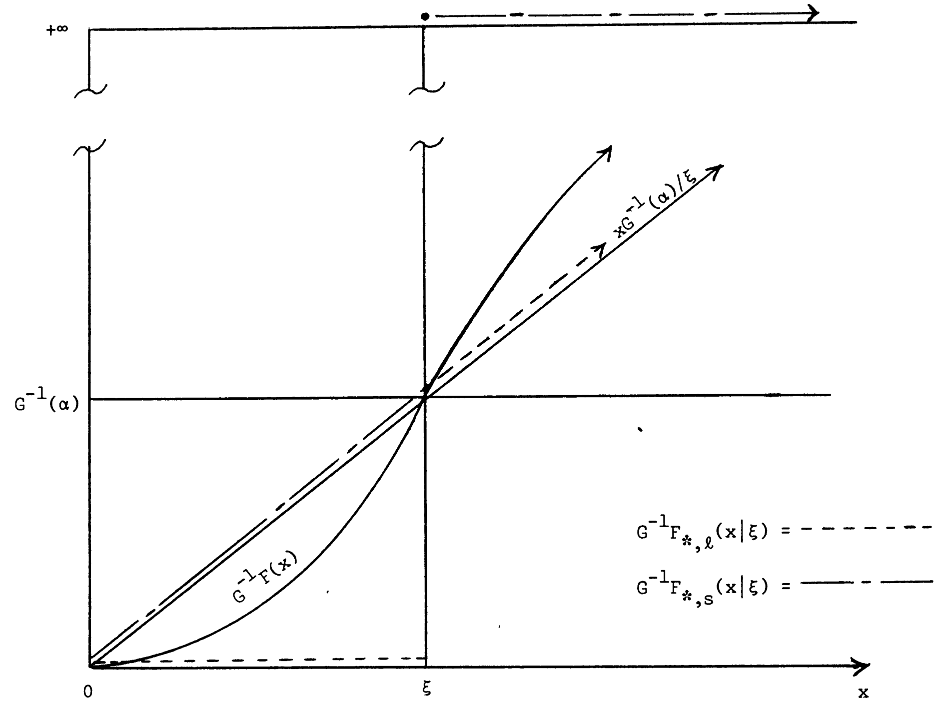


FIG. 2.1.

EXAMPLE 2.6. When $G(x) = 1/(1 + e^{-x})$, $x \in R$ then $G^{-1}(y) = -\ln((1-y)/y)$ on $(0, 1) \Rightarrow$

$$\mathcal{F}_t(G) = \left\{ F \in \mathcal{F} \,\middle|\, \frac{\bar{F}(t+x)/F(t+x)}{\bar{F}(t)/F(t)} \leqslant e^{-x} \forall x \geqslant 0 \quad \text{and} \quad t \in I(F) \atop \text{such that} \quad (t+x) \in I(F) \right\}$$

where $\bar{F}(y) = 1 - F(y)$ is the reliability function of $F$. In other words $\mathcal{F}_t(G)$ is the set of all continuous distributions such that the ratio of the odds of surviving beyond $t + x$ to the odds of surviving beyond $t$ is uniformly bounded above (in $t$) by $e^{-x}$.

REMARK 2.6. $G^{-1}F_{t,s}(x|\xi) - x$ and $G^{-1}F_{t,l}(x|\xi) - x$ are nondecreasing on $\{x \in R | F_{t,s}(x|\xi) < 1\}$ and $\{x \in R | F_{t,l}(x|\xi) < 1\}$ respectively and $x_\alpha(F_{t,s}(\cdot|\xi)) = x_\alpha(F_{t,l}(\cdot|\xi)) = \xi$ but neither is in $\mathcal{F}$ and hence not $\mathcal{F}_t(G)$. Both $\{F_{t,s}(\cdot|\xi) | \xi \in R\}$ and $\{F_{t,l}(\cdot)\xi) | \xi \in R\}$ are (stochastically increasing) locations parameter families.

THEOREM 2.3. *For all $\xi \in R$ and $F \in \mathcal{F}_t(G, \xi)$, $F_{t,s}(\cdot|\xi) \prec_{st} F \prec_{st} F_{t,l}(\cdot|\xi)$.*

PROOF. Similar to that of Theorem 2.3.

COROLLARY 2.2. *For $\xi' \leqslant \xi \leqslant \xi''$*

(a) $F \in \mathcal{F}_t(G, \xi') \Rightarrow F \prec_{st} F_{t,l}(\cdot|\xi)$;

(b) $F \in \mathcal{F}_t(G, \xi'') \Rightarrow F_{t,s}(\cdot|\xi) \prec_{st} F$.

Finally it should be noted that $F_{t,s}(\cdot|\xi)$ and $F_{t,l}(\cdot|\xi)$ are the stochastically largest and smallest distributions respectively satisfying Corollary 2.2.

**3. Fixed size subset selection procedures.**

3.1. *Preferred population formulation.* When faced with the problem of selecting at least $r$ of the $t$ best of $k$ populations ($1 \leqslant r \leqslant t < k$) an experimenter may well be indifferent between the selection of one of the $t$ best populations and one of the $(k - t)$ worst populations when they are sufficiently close together. This is the rationale of the indifference zone formulation (IZF) of selection problems in which the procedure is only required to guarantee correct selection (CS) with prespecified probability, $P^*$, when the $t$th and $(t + 1)$st best populations are sufficiently "far" apart (see Mahamunulu (1967)).

Alternatively, if the $t$th best and one of the $(k - t)$ worst populations are sufficiently "close" so that an experimenter would be satisfied with the selection of either one, then one can define a preferred population as any population satisfying this criterion and a correct selection as the selection of at least $r$ preferred populations. The preferred population formulation (PPF) requires the procedure make a correct selection with at least probability $P^*$ no matter what the true vector $\mathbf{F}$ of distributions.

Suppose $x_\alpha(F_i) \in \Xi$ where $\Xi$ is a known interval of the real line. Then, following Santner (1975), consider measures of closeness to $\pi_{(k-t+1)}$ defined in terms of a function $p : \Xi \to R$ satisfying:

$$p(\cdot) \text{ is continuous;} \tag{3.1}$$

$$p(\xi) \geqslant \gamma \equiv \inf\{\xi \in \Xi\} \ \forall \xi \in \Xi; \tag{3.2}$$

$$p(\cdot) \text{ is strictly increasing on } \ \Xi' \equiv \{\xi \in \Xi | p(\xi) > \gamma\}; \tag{3.3}$$

$$p(\xi) < \xi \ \forall \xi \in \Xi'. \tag{3.4}$$

Definition 3.1. $\pi_i$ is a *preferred population* (relative to $p(\cdot)$) iff $x_\alpha(F_i) > p(x_\alpha(F_{[k-t+1]}))$. Clearly at least $\pi_{(k-t+1)}, \cdots, \pi_{(k)}$ are preferred populations and the true population configuration will contain between $t$ and $k$ preferred $\pi_i$'s.

Remark 3.1. The PPF yields a strengthening of the probability guarantee over the IZF. Consider the problem of selecting the $t(r = t)$ *best* $\pi_i$'s when the true $\mathbf{F} \in \Omega \equiv \mathcal{F}_I^k$ where $\Omega' = \{\mathbf{F} \in \Omega | x_\alpha(\mathrm{F}_{[k-t]}) > p(x_\alpha(\mathrm{F}_{[k-t+1]}))\}$ and $\overline{\Omega}' = \Omega \setminus \Omega'$ are the indifference zone and preference zone respectively. The corresponding problem based on the PPF is to select at least $t$ preferred populations (relative to $p(\cdot)$). For $\mathbf{F} \in \overline{\Omega}'$ the $t$ best populations are the only preferred $\pi_i$'s and both the IZF and PPF guarantee their selection with probability $P^*$. For $\mathbf{F} \in \Omega'$ the PPF still guarantees the selection of at least $t$ preferred $\pi_i$'s while the IZF guarantees nothing.

Two specific choices for $p(\cdot)$ are now derived. Recall that for $F \in \mathcal{F}_*(G)$ any scale change of $F(x)$, say $F(\delta x)$, is equivalent to $F(x)$. Hence we require that $\pi_i$ be preferred relative to $\pi_{(k-t+1)}$ independent of the selected measurement units. More precisely this means $\pi_i$ be close to $\pi_{(k-t+1)}$ when $F_i \in \mathcal{F}_*(G, \xi'/\delta)$ and $F_{[k-t+1]} \in \mathcal{F}_*(G, \xi/\delta)$ independent of $\delta > 0$, i.e., $p(\cdot)$ must satisfy $\forall \xi', \xi \in (0, \infty)$

$$\xi' > p(\xi) \Leftrightarrow \xi'/\delta > p(\xi/\delta) \forall \delta > 0. \tag{3.5}$$

It can be shown that the unique $p(\cdot)$ satisfying (3.1) – (3.5) is $p(\xi) = c^*\xi$ where $0 < c^* < 1$. In the tail ordered case a similar argument based on the location invariant classes $E_t(G)$ leads to the choice $p(\xi) = \xi - d^*$ where $0 < d^* < \infty$. Hence $\pi_i$ will be called preferred in the star (tail) ordered case iff $x_\alpha(F_i) > c^* x_\alpha(F_{[k-t+1]})$ $(x_\alpha(F_{[k-t+1]}) - d^*)$.

In the remainder of the paper $G$ is a fixed distribution in $\mathcal{F}_0(\mathcal{F})$ with $I(G) = (0, \infty)$ $(R)$ in discussions of the star (tail) ordered model. Let

$$\Omega_* = \{\mathbf{F} = (F_1, \cdots, F_k) | F_i \in \mathcal{F}_*(G) \forall i\} \tag{3.6}$$

and

$$\Omega_t = \{\mathbf{F} = (F_1, \cdots, F_k) | F_i \in \mathcal{F}_t(G) \forall i\}. \tag{3.7}$$

Given $t(1 \leqslant t < k)$, $r(1 \leqslant r \leqslant t)$, $s(r \leqslant s < k - t + r)$ and $\alpha(0 < \alpha < 1)$ the goal is to select a subset of size $s$ containing at least $r$ preferred populations (a correct section). Note that for $k - t + r \leqslant s \leqslant k$ a correct selection must automatically occur. Let $N_\alpha \equiv \min\{n \geqslant 1 | 1 \leqslant (n+1)\alpha \leqslant n\}$ and for each $n \geqslant N_\alpha$ define the procedure.

$R_1(s, n)$: take independent random samples of size $n$ from each population and select $\pi_i \Leftrightarrow T_i \geqslant T_{[k-s+1]}$ where $T_{[1]} \leqslant \cdots \leqslant T_{[k]}$ are the ordered sample $\alpha$-quantiles $T_1, \cdots, T_k$.

*Design requirement*: given $P^* \in (\sum_{j=r}^{\min\{t,s\}} \binom{t}{j}\binom{k-t}{s-j}/\binom{k}{s}, 1)$ and $0 < c^* < 1$ $(0 < d^* < \infty)$ determine the smallest $n \geqslant N_\alpha$ so that

$$P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] \geqslant P^* \forall \mathbf{F} \in \Omega_*(\Omega_t) \tag{3.8}$$

where the event $[\mathrm{CS}|R_1(s, n)]$ denotes a correct selection using $R_1(s, n)$.

3.2. *Selection from star ordered families.* For fixed $\alpha \in (0, 1)$ it follows, without loss of generality, from Remark 2.3, that $G$ can be chosen so that $G^{-1}(\alpha) = 1$. Define the following subsets of $\Omega_*$. For $0 < \xi < \infty$ and $t \leqslant i \leqslant k$ let

$$\begin{aligned} \Omega_*(i) &= \{\mathbf{F} \in \Omega_* | x_\alpha(F_{[k-i]}) \leqslant c^* x_\alpha(F_{[k-t+1]}) \\ &\quad < x_\alpha(F_{[k-i+1]})\}, \qquad t \leqslant i < k \\ &= \{\mathbf{F} \in \Omega_* | c^* x_\alpha(F_{[k-t+1]}) < x_\alpha(F_{[1]})\}, \qquad i = k, \end{aligned} \tag{3.9}$$

$$\Omega_{*,\xi} = \{\mathbf{F} \in \Omega_* | x_\alpha(F_{[k-t+1]}) = \xi\}, \tag{3.10}$$

$$\Omega_*(i, \xi) = \Omega_*(i) \cap \Omega_{*,\xi}. \tag{3.11}$$

It follows that $\Omega_* = \cup_{\xi>0}\Omega_{*,\xi} = \cup_{\xi>0} \cup_{i=t}^{k} \Omega_*(i, \xi)$ and, for example, $\Omega_*(i, \xi)$ is the set of configurations for which there are exactly $i$ preferred populations and the $t$th largest $\alpha$-quantile is $\xi$.

It is easy to see that for $\mathbf{F} \in \Omega_*(i)$ where $r + k - s \leqslant i \leqslant k$ that $P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] = 1$. Hence $\inf_{\Omega_*} P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] = \min_{i \in \mathcal{I}} \inf_{\Omega_*(i)} P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)]$ where $\mathcal{I} = \{i | t \leqslant i < r + k - s\}$. For $\mathbf{F} \in \Omega_*$ and $i \in \mathcal{I}$ define

$$T_{(i)} \equiv \text{sample } \alpha\text{-quantile from } \pi_{(i)};$$

$$W_i \equiv r\text{th largest of } \{T_{(k-i+1)}, \cdots, T_{(k)}\}; \text{ and}$$

$$A_i \equiv \{W_i \geqslant \text{at least } (r + k - s - i) \text{ of } T_{(1)}, \cdots, T_{(k-i)}\}.$$

REMARK 3.2. It can be shown that the $A_j$'s form a nondecreasing sequence of events and that $\chi_{A_i}(\mathbf{T})$, the characteristic function of $A_i$, is nonincreasing in any of $T_{(1)}, \cdots, T_{(k-i)}$ and is nondecreasing in any of $T_{(k-i+1)}, \cdots, T_{(k)}$. This latter property and an application of Theorem 2.1 implies that $P_{\mathbf{F}}[A_i] = E_{\mathbf{F}}[\chi_{A_i}(\mathbf{T})]$ is nonincreasing in any of $F_{[1]}, \cdots, F_{[k-i]}$ and is nondecreasing in any of $F_{[k-i+1]}, \cdots, F_{[k]}$ under stochastic ordering $\prec_{st}$ (see Mahumunulu (1967) or Alam and Rizi (1966)).

LEMMA 3.1. $\inf_{\Omega_*(i)} P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] = \inf_{\Omega_*(i, 1)} P_{\mathbf{F}}[A_i]$ for $i \in \mathcal{I}$.

PROOF. Pick $\mathbf{F} \in \Omega_*(i, \xi)$ then $P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] = P_{\mathbf{F}}[A_i] = P_{\mathbf{F}}[W_i/\xi \geqslant$ at least $(r + k - s - i)$ of $T_{(1)}/\xi, \cdots, T_{(k-i)}/\xi]$. Now $\mathbf{T}' = \mathbf{T}/\xi x$ has the vector of cdf's $\mathbf{F}'(x) = (F_1(\xi x), \cdots, F_k(\xi x))$ which is in $\Omega_*(i, 1)$ by the equivalence of $F_i$ under scale changes and the definition of preferred $\pi_i$. So $P_{\mathbf{F}}[A_i] = P_{\mathbf{F}'}[A_i]$ and the proof is complete.

REMARK 3.3. An equivalent way of viewing the proof of Lemma 3.1 is as an application of invariance. The problem is invariant under the group of all possible common scale changes of the raw data. Hence the risk under a 0-1 loss structure becomes the probability of incorrect selection and is constant over orbits—the orbit of $\mathbf{F}$ consisting of all common scale changes of the components of $\mathbf{F}$.

Next a lower bound for $P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)]$ over $\mathbf{F} \in \Omega_{*,1}$ is obtained. Given $i \in \mathcal{I}$ and $\mathbf{F} \in \Omega_*(i, 1)$ it follows from Corollary 2.1 that

$$
\begin{aligned}
&F_{[j]} \prec_{st} F_{*,l}(\cdot|c^*), && 1 \leqslant j \leqslant k - i \\
&F_{*,s}(\cdot|1) \prec_{st} F_{[j]}, && k - t + 1 \leqslant j \leqslant k \\
&F_{*,s}(\cdot|c^*) \prec_{st} F_{[j]}, && \text{for any other } j
\end{aligned}
\tag{3.12}
$$

since $x_\alpha(F_{[j]}) \leqslant c^*$ for $1 \leqslant j \leqslant k - i$, $1 \leqslant x_\alpha(F_{[j]})$ for $k - t + 1 \leqslant j \leqslant k$ and $c^* < x_\alpha(F_{[j]})$ for all remaining $j$'s. Let $\mathbf{F}(i)$ be any configuration having $k - i$, $t$ and $i - t$ (if $i > t$) components of the types $F_{*,l}(\cdot|c^*)$, $F_{*,s}(\cdot|1)$ and $F_{*,s}(\cdot|c^*)$ respectively. Then

$$
\begin{aligned}
P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] &= P_{\mathbf{F}}[A_i], \\
&\geqslant P_{\mathbf{F}(i)}[A_i] && \text{from (3.12) and Remark 3.2,} \\
&\geqslant P_{\mathbf{F}(i)}[A_t] && \text{since } A_t \subset A_i, \\
&\geqslant P_{\mathbf{F}(t)}[A_t],
\end{aligned}
$$

where the final inequality holds since when $i > t$ the $j$th($k - i + 1 \leqslant j \leqslant k - t$) ordered component of $\mathbf{F}(i)$ is $F_{*,s}(\cdot|c^*) \prec_{st} F_{*,l}(\cdot|c^*)$, the $j$th ordered component of $\mathbf{F}(t)$.

To calculate the lower bound $P_{\mathbf{F}(t)}[A_t]$ let

$$\text{(3.13)} \qquad \begin{aligned} H(y) &= B(B(y); k - t - s + r, k - t) \\ J(y) &= B(B(y); t - r + 1, t) \end{aligned}$$

where $B(y) = B(y; q, n)$, $q = [(n + 1)\alpha]$ and $[\cdot]$ denotes the greatest integer function. Hence $H(F_{*,l}(y|c^*))$ is the cdf of the $(k - t - s + r)$th smallest of $\{T_{(1)}, \cdots, T_{(k-t)}\}$ and $J(F_{*,s}(y|1))$ is the cdf of the $(t - r + 1)$th smallest of $\{T_{(k-t+1)}, \cdots, T_{(k)}\}$ all under $\mathbf{F}(t)$. Therefore

$$\begin{aligned} P_{\mathbf{F}(t)}[A_t] &= \textstyle\int_0^\infty H(F_{*,l}(y|c^*))dJ(F_{*,s}(y|1)) \\ &= \textstyle\int_{c^*}^1 H(G(y/c^*))dJ(G(y)) \\ &\quad + H(G(1/c^*))(1 - J(\alpha)) \\ &= \textstyle\int_{G(c^*)}^{\alpha} H\big(G(G^{-1}(y)/c^*)\big)dJ(y) \\ &\quad + H(G(1/c^*))(1 - J(\alpha)) \\ &= F_n^*(r, t, s, k, \alpha, c^*) \end{aligned}$$

or simply $F_n^*$. Since $\mathbf{F}(t) \not\in \Omega_*$ (the component distributions are not in $\mathcal{F}_*(G)$) it is not immediate that $F_n^*$ is the infimum of the PCS over $\Omega_*$. The next result establishes that this is the case.

THEOREM 3.1. $\inf_{\Omega_*} P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] = F_n^*(r, t, s, k, \alpha, c^*)$.

PROOF. From Lemma 3.1 and the discussion above it suffices to exhibit a sequence of configurations $\{\mathbf{F}_j\}$ in $\Omega_*$ so that $\lim_{j\to\infty} P_{\mathbf{F}_j}[\mathrm{CS}|R_1(s, n)] = F_n^*$. A sequence in $\Omega_*$ which approximates $\mathbf{F}(t)$ is the obvious candidate. Let $L \equiv [1/c^*] + 1$ and for $j \geqslant L$ define

$$\begin{aligned} F_{l,j}(y) &= 0 & y &< c^* - j^{-1} \\ &= G[j(y - (c^* - j^{-1}))] & c - j^{-1} &\leqslant y < c^* \\ &= G(y/c^*) & c^* &\leqslant y \\ F_{s,j}(y) &= G(y) & y &< 1 \\ &= G(1 + j(y - 1)) & 1 &\leqslant y. \end{aligned}$$

Then $F_{l,j}(\cdot) \to_w F_{*,l}(\cdot|c^*) \Rightarrow H(F_{l,j}(\cdot)) \to_w H(F_{*,l}(\cdot|c^*))$ and $F_{s,j}(\cdot) \to_w F_{*,s}(\cdot|1) \Rightarrow J(F_{s,j}(\cdot)) \to_w J(F_{*,s}(\cdot|1)$ where $\to_w$ denotes weak convergence of distributions. Furthermore $F_{l,j} \in \mathcal{F}_*(G, c^*)$ and $F_{s,j} \in \mathcal{F}_*(G, 1)$ and so $\mathbf{F}_j$ with $t$ components $F_{l,j}$

and $(k - t)$ components $F_{s,j}$ is in $\Omega_*(t)$. Therefore

$$\begin{aligned} P_{\mathbf{F}_j}[A_t] &= \int_0^\infty H(F_{l,j}(y))\,dJ(F_{s,j}(y)) \\ &= \int_0^{c^*} H(F_{l,j}(y))\,dJ(F_{*,s}(y|1)) + \int_{c^*}^\infty H(F_{*,l}(y|c^*))\,dJ(F_{s,j}(y)) \\ &\qquad \text{since } F_{l,j}(y) = F_{*,l}(y|c^*) \quad \text{on } [c^*, \infty) \quad \text{and} \\ &\qquad F_{s,j}(y) = F_{*,s}(y|1) \quad \text{on } [0, 1) \\ &= \int_0^{c^*} H(F_{l,j}(y))\,dJ(F_{*,s}(y|1)) + \int_0^\infty H(F_{*,l}(y|c^*))\,dJ(F_{s,j}(y)) \\ &\qquad \text{since } F_{*,l}(y|c^*) = 0 \quad \text{on } [0, c^*) \\ &\to 0 + \int_0^\infty H(F_{*,l}(y|c^*))\,dJ(F_{*,s}(y|1)) = F_n^* \end{aligned}$$

where the first convergence follows from dominated convergence since $H(F_{l,j}(y)) \to 0$ a.e.$[J(F_{*,s}(\cdot|1))]$ on $[0, c^*]$ and the second follows from the weak convergence of $J(F_{s,j}(\cdot))$ to $J(F_{*,s}(\cdot|1))$ and the fact that $H(F_{*,l}(\cdot|c^*))$ is bounded and continuous a.e. $[J(F_{*,s}(\cdot|1))]$ on $[0, \infty)$. The proof is completed.

It can be shown that $F_n^* \to 1$ as $n \to \infty$ and hence (3.8) can be guaranteed by a finite $n$ for any $P^* < 1$.

EXAMPLE 3.1. Given $\alpha \in (0, 1)$ choose

$$\begin{aligned} G(x) &= 1 - e^{x\ln(1-\alpha)}, & x &\geq 0; \\ &= 0, & x &< 0. \end{aligned}$$

$\Omega_*$ is then all configurations of continuous IFRA distributions. Note that $G(\cdot)$ is chosen so that $G^{-1}(\alpha) = 1$. Suppose it is desired to select at least $r = t$ preferred populations. In this case

$$\begin{aligned} F_n^* = t\int_{1-(1-\alpha)^{c^*}}^{\alpha} & B\left(B\left(1 - [1 - x]^{1/c^*}\right); k - s, k - t\right) dB(x) \\ & + B\left(B\left(1 - [1 - \alpha]^{1/c^*}\right); k - s, k - t\right)(1 - B(\alpha))^t. \end{aligned}$$

For computational purposes it is desirable to remove the composition of incomplete beta functions from inside the integral. After an integration by parts and a change of variables $F_n^*$ becomes

$$\begin{aligned} \text{(3.14)} \quad \binom{k-t}{k-s} \int_\alpha^{1-(1-\alpha)^{1/c^*}} & \left(1 - B\left(1 - [1 - y]^{c^*}\right)\right)^t (1 - B(y))^{s-t} dB^{k-s}(y) \\ & + \left(1 - B\left(1 - (1 - \alpha)^{c^*}\right)\right)^t B(B(\alpha); k - s, k - t). \end{aligned}$$

Expression (3.14) was evaluated on Cornell University's IBM 370/169 computer using a procedure based on Simpson's rule (see Shampine and Allen (1973)). Tables 1, 2 and 3 give the smallest odd sample sizes required to meet (3.8) for $\alpha = .25(.25).75$, $c^* = .65, .70$, $P^* = .75, .90, .95$, $k = 2(1)9$, $t = 1(1)[k/2]$, $r = t$, $s = t(1)[k/2]$.

One measure of the cost of using the nonparametric procedure $R_1(s, n)$ is obtained by comparing the sample size required by $R_1(s, n)$ to achieve a specified

TABLE 3.1.
*Sample sizes required by $R_1(s, n)$ when selecting from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\alpha = .25$ and $c^* = .65$.*

| $k$ | $t$ | $s$ | $P^*$ .75 | .90 | .95 |
|---|---|---|---|---|---|
| 2 | 1 | 1 | 39 | 87 | 131 |
| 3 | 1 | 1 | 55 | 115 | 163 |
| 4 | 1 | 1 | 69 | 135 | 187 |
| 4 | 1 | 2 | 27 | 67 | 103 |
| 4 | 2 | 2 | 87 | 155 | 211 |
| 5 | 1 | 1 | 79 | 147 | 203 |
| 5 | 1 | 2 | 35 | 79 | 119 |
| 5 | 2 | 2 | 103 | 179 | 235 |
| 6 | 1 | 1 | 89 | 159 | 215 |
| 6 | 1 | 2 | 43 | 91 | 131 |
| 6 | 1 | 3 | 23 | 59 | 95 |
| 6 | 2 | 2 | 119 | 195 | 255 |
| 6 | 2 | 3 | 59 | 115 | 155 |
| 6 | 3 | 3 | 127 | 207 | 267 |
| 7 | 1 | 1 | 97 | 169 | 191 |
| 7 | 1 | 2 | 47 | 99 | 143 |
| 7 | 1 | 3 | 27 | 67 | 103 |
| 7 | 2 | 2 | 127 | 211 | 271 |
| 7 | 2 | 3 | 71 | 127 | 171 |
| 7 | 3 | 3 | 143 | 227 | 287 |
| 8 | 1 | 1 | 103 | 177 | 235 |
| 8 | 1 | 2 | 53 | 107 | 151 |
| 8 | 1 | 3 | 31 | 75 | 111 |
| 8 | 1 | 4 | 23 | 59 | 87 |
| 8 | 2 | 2 | 139 | 219 | 283 |
| 8 | 2 | 3 | 79 | 139 | 183 |
| 8 | 2 | 4 | 55 | 99 | 139 |
| 8 | 3 | 3 | 155 | 239 | 299 |
| 8 | 3 | 4 | 91 | 151 | 195 |
| 8 | 4 | 4 | 163 | 247 | 307 |
| 9 | 1 | 1 | 109 | 183 | 243 |
| 9 | 1 | 2 | 59 | 115 | 163 |
| 9 | 1 | 3 | 35 | 83 | 123 |
| 9 | 1 | 4 | 23 | 63 | 95 |
| 9 | 2 | 2 | 147 | 231 | 291 |
| 9 | 2 | 3 | 87 | 151 | 195 |
| 9 | 2 | 4 | 59 | 111 | 147 |
| 9 | 3 | 3 | 167 | 251 | 313 |
| 9 | 3 | 4 | 99 | 163 | 211 |
| 9 | 4 | 4 | 175 | 263 | 325 |

$P^*$ to that required by the best procedure for the parametric problem with exponential populations. Suppose $\pi_i$ has cdf $F(x|\theta_i) = 1 - e^{-x/\theta_i}$, $x > 0$ where $\theta_i > 0$ but unknown, then $x_\alpha(F_i) = -\theta_i \ln(1-\alpha) \Rightarrow x_\alpha(F(\cdot|\theta_i)) > c^* x_\alpha(F_{[k-t+1]})$ $\Leftrightarrow \theta_i > c^*\theta_{[k-t+1]}$ where $\theta_{[i]}$ is the $i$th ordered $\theta_j$. In this case the optimal selection procedure is the natural procedure based on sample means rather than sample $\alpha$--quantiles (Eaton (1967)). Denote this procedure as $R_M(s, n)$. It is easily seen that

TABLE 3.2.
*Sample sizes required by $R_1(s, n)$ when selecting from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\alpha = .50$ and $c^* = .65, .70$.*

| | | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $P^*$ | | | $P^*$ | |
| $k$ | $t$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 2 | 1 | 1 | 19 | 45 | 67 | 27 | 65 | 97 |
| 3 | 1 | 1 | 29 | 59 | 85 | 43 | 85 | 123 |
| 4 | 1 | 1 | 37 | 69 | 97 | 53 | 101 | 141 |
| 4 | 1 | 2 | 13 | 33 | 51 | 19 | 47 | 73 |
| 4 | 2 | 2 | 43 | 79 | 109 | 63 | 117 | 157 |
| 5 | 1 | 1 | 43 | 77 | 105 | 61 | 113 | 153 |
| 5 | 1 | 2 | 17 | 39 | 59 | 25 | 57 | 85 |
| 5 | 2 | 2 | 53 | 91 | 121 | 77 | 133 | 177 |
| 6 | 1 | 1 | 47 | 83 | 111 | 69 | 121 | 163 |
| 6 | 1 | 2 | 21 | 45 | 67 | 31 | 67 | 97 |
| 6 | 1 | 3 | 11 | 29 | 47 | 17 | 43 | 67 |
| 6 | 2 | 2 | 61 | 101 | 131 | 87 | 147 | 191 |
| 6 | 2 | 3 | 31 | 57 | 79 | 43 | 83 | 115 |
| 6 | 3 | 3 | 65 | 107 | 137 | 95 | 155 | 199 |
| 7 | 1 | 1 | 51 | 87 | 117 | 75 | 129 | 171 |
| 7 | 1 | 2 | 25 | 51 | 73 | 37 | 73 | 105 |
| 7 | 1 | 3 | 13 | 35 | 51 | 19 | 49 | 75 |
| 7 | 2 | 2 | 67 | 107 | 139 | 97 | 155 | 203 |
| 7 | 2 | 3 | 35 | 65 | 87 | 51 | 95 | 127 |
| 7 | 3 | 3 | 73 | 115 | 147 | 107 | 169 | 215 |
| 8 | 1 | 1 | 53 | 91 | 121 | 79 | 135 | 179 |
| 8 | 1 | 2 | 27 | 51 | 79 | 41 | 81 | 113 |
| 8 | 1 | 3 | 17 | 39 | 57 | 23 | 55 | 83 |
| 8 | 1 | 4 | 11 | 29 | 45 | 15 | 41 | 63 |
| 8 | 2 | 2 | 71 | 113 | 145 | 103 | 165 | 211 |
| 8 | 2 | 3 | 41 | 71 | 95 | 59 | 103 | 137 |
| 8 | 2 | 4 | 27 | 51 | 69 | 37 | 73 | 101 |
| 8 | 3 | 3 | 79 | 123 | 155 | 117 | 179 | 225 |
| 8 | 3 | 4 | 47 | 77 | 101 | 67 | 111 | 145 |
| 8 | 4 | 4 | 83 | 127 | 159 | 121 | 185 | 231 |
| 9 | 1 | 1 | 57 | 95 | 125 | 83 | 139 | 185 |
| 9 | 1 | 2 | 31 | 59 | 83 | 45 | 87 | 121 |
| 9 | 1 | 3 | 19 | 43 | 61 | 27 | 61 | 89 |
| 9 | 1 | 4 | 11 | 31 | 49 | 17 | 45 | 69 |
| 9 | 2 | 2 | 75 | 119 | 149 | 109 | 173 | 219 |
| 9 | 2 | 3 | 45 | 77 | 101 | 65 | 111 | 145 |
| 9 | 2 | 4 | 29 | 55 | 75 | 43 | 81 | 109 |
| 9 | 3 | 3 | 85 | 129 | 161 | 125 | 189 | 235 |
| 9 | 3 | 4 | 51 | 83 | 107 | 75 | 121 | 155 |
| 9 | 4 | 4 | 89 | 135 | 167 | 131 | 193 | 243 |

TABLE 3.3.
*Sample sizes required by $R_1(s, n)$ when selecting from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\alpha = .75$ and $c^* = .65, .70$.*

| | | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $P^*$ | | | $P^*$ | |
| $k$ | $t$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 2 | 1 | 1 | 15 | 33 | 51 | 21 | 47 | 73 |
| 3 | 1 | 1 | 21 | 45 | 63 | 33 | 65 | 93 |
| 4 | 1 | 1 | 29 | 53 | 73 | 41 | 77 | 105 |
| 4 | 1 | 2 | 11 | 27 | 39 | 15 | 35 | 55 |
| 4 | 2 | 2 | 33 | 59 | 81 | 47 | 87 | 119 |
| 5 | 1 | 1 | 33 | 57 | 79 | 45 | 85 | 115 |
| 5 | 1 | 2 | 13 | 31 | 43 | 19 | 43 | 63 |
| 5 | 2 | 2 | 41 | 69 | 91 | 59 | 101 | 133 |
| 6 | 1 | 1 | 37 | 63 | 85 | 53 | 91 | 123 |
| 6 | 1 | 2 | 17 | 35 | 51 | 23 | 51 | 71 |
| 6 | 1 | 3 | 11 | 23 | 35 | 11 | 31 | 51 |
| 6 | 2 | 2 | 45 | 75 | 99 | 67 | 111 | 143 |
| 6 | 2 | 3 | 23 | 43 | 59 | 35 | 63 | 85 |
| 6 | 3 | 3 | 49 | 79 | 103 | 71 | 115 | 149 |
| 7 | 1 | 1 | 37 | 67 | 89 | 57 | 97 | 129 |
| 7 | 1 | 2 | 19 | 39 | 55 | 27 | 55 | 79 |
| 7 | 1 | 3 | 11 | 27 | 39 | 15 | 35 | 55 |
| 7 | 2 | 2 | 49 | 81 | 105 | 73 | 119 | 151 |
| 7 | 2 | 3 | 27 | 49 | 65 | 39 | 71 | 95 |
| 7 | 3 | 3 | 55 | 87 | 111 | 81 | 127 | 161 |
| 8 | 1 | 1 | 41 | 69 | 93 | 61 | 101 | 135 |
| 8 | 1 | 2 | 21 | 41 | 59 | 29 | 61 | 85 |
| 8 | 1 | 3 | 13 | 29 | 43 | 19 | 41 | 61 |
| 8 | 1 | 4 | 7 | 23 | 31 | 11 | 31 | 47 |
| 8 | 2 | 2 | 53 | 85 | 109 | 79 | 125 | 159 |
| 8 | 2 | 3 | 31 | 53 | 71 | 43 | 77 | 103 |
| 8 | 2 | 4 | 19 | 39 | 51 | 27 | 55 | 75 |
| 8 | 3 | 3 | 61 | 93 | 117 | 87 | 135 | 169 |
| 8 | 3 | 4 | 35 | 57 | 75 | 51 | 83 | 107 |
| 8 | 4 | 4 | 63 | 95 | 119 | 91 | 139 | 173 |
| 9 | 1 | 1 | 45 | 73 | 95 | 65 | 105 | 139 |
| 9 | 1 | 2 | 23 | 45 | 63 | 33 | 65 | 91 |
| 9 | 1 | 3 | 15 | 31 | 47 | 21 | 45 | 67 |
| 9 | 1 | 4 | 11 | 23 | 35 | 15 | 35 | 51 |
| 9 | 2 | 2 | 57 | 89 | 113 | 83 | 129 | 165 |
| 9 | 2 | 3 | 33 | 57 | 75 | 49 | 83 | 109 |
| 9 | 2 | 4 | 23 | 43 | 57 | 31 | 59 | 81 |
| 9 | 3 | 3 | 65 | 97 | 121 | 93 | 141 | 177 |
| 9 | 3 | 4 | 39 | 63 | 79 | 55 | 91 | 115 |
| 9 | 4 | 4 | 67 | 101 | 125 | 99 | 147 | 181 |

for $\Omega = (0, \infty)^k$

$$\inf_\Omega P_\theta\{\mathrm{CS}|R_M(s, n)\} = \int_0^\infty H(\Gamma(y|c^*, n))dJ(\Gamma(y|1, n))$$

$$\text{where} \quad \Gamma(y|\theta, n) = \int_0^y \frac{e^{-w/\theta} w^{n-1}}{(n-1)!\theta^n}\, dw.$$

When $r = t$ this integral reduces, after a change of variables, to

$$(3.15) \quad (k-s)\binom{k-t}{s-t}\int_0^1\{1 - \Gamma(c^*\Gamma^{-1}[x|1, n]|1, n)\}^t x^{k-s-1}(1-x)^{s-t}\, dx.$$

Expression (3.15) was evaluated using an algorithm based on Simpson's rule. Table 4 contains the smallest odd sample signs required to satisfy

$$(3.16) \qquad \inf_\Omega P_\theta[\mathrm{CS}|R_M(s, n)] > P^*$$

for $c^* = .65$, $.70$, $P^* = .75$, $.90$, $.95$, $k = 2(1)9$, $r = t = 1(1)[k/2]$ and $s = t(1)[k/2]$. These calculations show that the exponential procedure requires roughly half as many observations as the nonparametric IFRA procedure.

3.3. *Selection from tail ordered families.* Fix $G' \in \mathfrak{F}$ satisfying $I(G') = R$ and $\alpha \in (0, 1)$. Choose $G \in E_t(G')$ so that $G^{-1}(\alpha) = 0$ and, for $d^* \in (0, \infty)$, define

$$(3.17) \qquad \Omega_t \equiv \{\mathbf{F} \in \mathfrak{F}^k | F_i \prec_t G \,\forall 1 \leq i \leq k\},$$

$$(3.18) \qquad \Omega_{t,\xi} \equiv \{\mathbf{F} \in \Omega_t | x_\alpha(F_{[k-t+1]}) = \xi\}, \qquad \xi \in R,$$

$$(3.19) \qquad \begin{aligned} \Omega_t(i) &= \{\mathbf{F} \in \Omega_t | x_\alpha(F_{[k-i]}) \leq x_\alpha(F_{[k-t+1]}) - d^* \\ &\qquad < x_\alpha(F_{[k-i+1]})\},\ t \leq i < k \\ &= \{\mathbf{F} \in \Omega_t | x_\alpha(F_{[k-t+1]}) - d^* < x_\alpha(F_{[1]})\},\ i = k, \end{aligned}$$

$$(3.20) \qquad \Omega_t(i, \xi) = \Omega_t(i) \cap \Omega_{t,\xi}.$$

It follows that $\Omega_t = \cup_{\xi \in R}\Omega_{t,\xi} = \cup_{\xi \in R} \cup_{i=t}^k \Omega_t(i, \xi)$. By using location invariance arguments similar to the scale invariance arguments of Lemma 3.1 it can be shown that

LEMMA 3.2. $\inf_{\Omega_t} P[\mathrm{CS}|R, (s, n)] = \min_{i \in \mathcal{G}} \inf_{\Omega_t(i, 0)} P_{\mathbf{F}}[A_i]$ *where* $A_i$ *and* $\mathcal{G}$ *are defined as in subsection* 3.2.

From the stochastic bounds developed in Corollary 2.2 for $F \in \mathfrak{F}_t(G)$ and arguments similar to those preceding Theorem 3.1 a tight lower bound, $F_n^t$, can be constructed for $\inf_{\Omega_t} P[\mathrm{CS}|R_1(s, n)]$.

THEOREM 3.2. $\inf_{\Omega_t} P[\mathrm{CS}|R_1(s, n)] = P_{\mathbf{F}(t)}[A_t] = F_n^t = F_n^t(r, t, s, k, \alpha, d^*)$ *where* $\mathbf{F}(t)$ *is any configuration with* $t$ *components* $F_{t,s}(\cdot|0)$ *and* $(k - t)$ *components* $F_{t,t}(\cdot| - d^*)$ *and*

$$(3.21) \qquad \begin{aligned} F_n^t \equiv &\int_{G(-d^*)}^\alpha H(G(G^{-1}(y) + d^*))dJ(y) \\ &+ H(G(d^*))[1 - J(\alpha)]. \end{aligned}$$

TABLE 3.4.
*Sample sizes required by $R_M(s, n)$ when selecting from exponential populations with $\alpha$ arbitrary and $c^* = .65, .70$.*

| | | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $P^*$ | | | $P^*$ | |
| $k$ | $t$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 2 | 1 | 1 | 7 | 19 | 31 | 9 | 27 | 43 |
| 3 | 1 | 1 | 11 | 27 | 41 | 17 | 39 | 59 |
| 4 | 1 | 1 | 15 | 33 | 47 | 23 | 47 | 67 |
| 4 | 1 | 2 | 5 | 15 | 23 | 7 | 21 | 33 |
| 4 | 2 | 2 | 21 | 39 | 53 | 29 | 57 | 77 |
| 5 | 1 | 1 | 19 | 37 | 51 | 27 | 53 | 73 |
| 5 | 1 | 2 | 7 | 19 | 29 | 11 | 27 | 41 |
| 5 | 2 | 2 | 25 | 45 | 59 | 37 | 65 | 85 |
| 6 | 1 | 1 | 21 | 39 | 55 | 31 | 57 | 79 |
| 6 | 1 | 2 | 9 | 23 | 33 | 13 | 33 | 47 |
| 6 | 1 | 3 | 5 | 13 | 21 | 7 | 19 | 31 |
| 6 | 2 | 2 | 29 | 49 | 63 | 43 | 71 | 93 |
| 6 | 2 | 3 | 15 | 29 | 39 | 21 | 39 | 55 |
| 6 | 3 | 3 | 33 | 53 | 67 | 47 | 75 | 97 |
| 7 | 1 | 1 | 23 | 43 | 57 | 33 | 61 | 83 |
| 7 | 1 | 2 | 11 | 25 | 35 | 17 | 37 | 51 |
| 7 | 1 | 3 | 7 | 17 | 25 | 9 | 23 | 35 |
| 7 | 2 | 2 | 33 | 53 | 67 | 47 | 77 | 97 |
| 7 | 2 | 3 | 17 | 33 | 43 | 25 | 45 | 61 |
| 7 | 3 | 3 | 37 | 57 | 71 | 53 | 81 | 103 |
| 8 | 1 | 1 | 25 | 45 | 59 | 35 | 65 | 85 |
| 8 | 1 | 2 | 13 | 27 | 39 | 19 | 39 | 55 |
| 8 | 1 | 3 | 7 | 19 | 29 | 11 | 27 | 41 |
| 8 | 1 | 4 | 5 | 13 | 21 | 5 | 19 | 29 |
| 8 | 2 | 2 | 35 | 55 | 71 | 51 | 81 | 103 |
| 8 | 2 | 3 | 21 | 35 | 47 | 29 | 51 | 67 |
| 8 | 2 | 4 | 13 | 25 | 33 | 17 | 35 | 49 |
| 8 | 3 | 3 | 39 | 61 | 75 | 57 | 87 | 109 |
| 8 | 3 | 4 | 23 | 37 | 49 | 33 | 55 | 71 |
| 8 | 4 | 4 | 41 | 61 | 77 | 59 | 89 | 111 |
| 9 | 1 | 1 | 25 | 47 | 61 | 37 | 67 | 89 |
| 9 | 1 | 2 | 15 | 29 | 41 | 21 | 43 | 59 |
| 9 | 1 | 3 | 9 | 21 | 31 | 13 | 29 | 43 |
| 9 | 1 | 4 | 5 | 15 | 23 | 7 | 21 | 33 |
| 9 | 2 | 2 | 37 | 57 | 73 | 53 | 83 | 105 |
| 9 | 2 | 3 | 23 | 37 | 49 | 31 | 55 | 71 |
| 9 | 2 | 4 | 15 | 27 | 37 | 21 | 39 | 53 |
| 9 | 3 | 3 | 41 | 63 | 79 | 61 | 91 | 113 |
| 9 | 3 | 4 | 25 | 41 | 53 | 37 | 59 | 75 |
| 9 | 4 | 4 | 45 | 65 | 81 | 63 | 95 | 117 |

REMARK 3.4. The proof of Theorem 3.2 consists in constructing a sequence of configurations in $\Omega_t$ such that the corresponding sequence of probabilities of correct selection converge to $F_n^t$. It can also be shown that $F_n^t \to 1$ as $n \to \infty$ and hence (3.8) can be guaranteed for any $P^* < 1$.

## 4. Random size subset selection procedures.

4.1. *Preferred population formulation.* When the goal of an experiment is to screen a large number of populations to obtain a more manageable subset containing at least $r$ preferred populations then the procedure $R_1(s, n)$ has the undesirable characteristic of always selecting a subset of size $s$. This section studies the random size restricted subset selection procedures introduced by Gupta and Santner (1973) and Santner (1975). These rules are characterized by the properties: (1) the experimenter specifies an upper bound, say $s$, on the size of the selected subset and (2) the procedure is able to capitalize on configurations favorable to the experimenter by selecting fewer than the maximum number of $s\,\pi_i$'s.

Following Santner (1975), restricted subset selection procedures will be defined in terms of a sequence of functions $\Psi = \{\psi_n\}_{n=1}^{\infty}$, $\Psi_n : \Xi \to \Xi$ satisfying:

(4.1) $$\forall \xi \in \Xi \quad \text{and} \quad \forall n, \psi_n(\xi) > \xi;$$

(4.2) $$\forall n, \Psi_n \quad \text{is continuous and strictly increasing; and}$$

(4.3) $$\forall \xi \in \Xi, \psi_n(\xi) \to \xi \quad \text{as} \quad n \to \infty.$$

EXAMPLE 4.1. For $\Xi = (0, \infty)$ take $\psi_n(\xi) = \xi\rho^{e(n)}$ where $\rho \in (1, \infty)$ and $\{e(n)\}$ is any sequence of positive numbers converging to zero as $n \to \infty$.

EXAMPLE 4.2. For $\Xi = R$ take $\psi_n(\xi) = \xi + e(n)$ where $e(n)$ is as in the previous example.

For specified $\alpha \in (0, 1)$, and each $s$ in $\{1, \cdots, k\}$, $\Psi$ satisfying (4.1)—(4.3) and $n \geqslant N_\alpha$ define the procedure

$R_2(s, \Psi, n)$: take independent random samples of size $n$ from each population and select $\pi_i \Leftrightarrow T_i \geqslant \max\{T_{[k-s+1]}, \psi_n^{-1}(T_{[k]})\}$ where $T_{[1]} \leqslant \cdots \leqslant T_{[k]}$ are the ordered values of the sample $\alpha$-quantiles $T_1, \cdots, T_k$.

*Design requirement:* given $P^* \in (t/k, 1)$, $1 \leqslant t < k$, $1 \leqslant s \leqslant k - t + 1$, $\Psi$ and $0 < c^* < 1 (0 < d^* < \infty)$ determine the smallest $n \geqslant N_\alpha$ so that

(4.4) $$P_{\mathbf{F}}[\text{CS}|R_2(s, \Psi, n)] \geqslant p^* \forall \mathbf{F} \in \Omega_*(\Omega_t)$$

where the event $[\text{CS}|\text{R}_2(s, \Psi, n)]$ occurs iff at least one preferred $\pi_i$ is selected.

REMARK 4.1. In this formulation $\pi_i$ is preferred $\Leftrightarrow$ $x_\alpha(F_i) > c^* x_\alpha(F_{[k-t+1]})(x_\alpha(F_{[k-t+1]}) - d^*)$ in the star (tail) ordered model. However, only the goal corresponding to $r = 1$ of Section 3 is discussed. The monotonicity arguments used below in establishing the infimum of the PCS are invalid when $r > 1$. When $R_2(s, \Psi, n)$ selects $k - t + 1$ populations a correct selection automatically occurs and when $s = 1$ the rule reduces to $R_1(1, n)$. The results in subsections 4.2 and 4.3 are valid for $1 \leqslant t < k$ and $1 < s < k - t + 1$. The case $s = k - t + 1$

can be easily analysed by straightforward arguments similar to those developed below.

4.2 *Selection from star ordered families.* Given $\alpha \in (0, 1)$ choose $G$ as in Section 3.2 and for $1 \leqslant t < k$ define $\Omega_*$, $\Omega_*(i)$, $\Omega_{*,\xi}$ and $\Omega_*(i, \xi)$ by (3.6), (3.9), (3.10) and (3.11) respectively. Clearly $P_{\mathbf{F}}[\mathrm{CS}|R_2(s, \Psi, n)] = 1 \forall \mathbf{F} \in \Omega_k$. Let $\mathcal{G} \equiv \{t, \cdots, k-1\}$ and for $i \in \mathcal{G}$ and $\mathbf{F} \in \Omega_* \Omega_*(i)$ define $W_i = \max\{T_{(k-i+1)}, \cdots, T_{(k)}\}$ and

$$(4.5) \qquad B_i = \{W_i \geqslant \text{at least } (k-s+1-i) \text{ of } T_{(1)}, \cdots, T_{(k-i)}; \; W_i \rho^{e(n)} \geqslant \max\{T_{(j)}|1 \leqslant j \leqslant k-i\}\} \qquad ,t \leqslant i \leqslant k-s$$
$$= \{W_i \rho^{e(n)} \geqslant \max\{T_{(j)}|1 \leqslant j \leqslant k-i\}\} \qquad ,k-s < i \leqslant k-1$$

where as before, $T_{(i)}$ is the sample $\alpha$-quantile with cdf $F_{[i]}$. Then $\forall i \in \mathcal{G}$ and $\mathbf{F} \in \Omega_*(i)$, $P_F[\mathrm{CS}|R_2(s, \Psi, n)] = P_{\mathbf{F}}[B_i]$. Furthermore, the events $\{B_j\}$ form a nondecreasing sequence and $P_{\mathbf{F}}[B_i] = E_{\mathbf{F}}[\chi_{B_i}(\mathbf{T})]$ is nonincreasing in any of $F_{[1]}, \cdots, F_{[k-i]}$ and nondecreasing in any of $F_{[k-i+1]}, \cdots, F_{[k]}$ under $\prec_{st}$.

LEMMA 4.1. $\inf_{\Omega_*} P_{\mathbf{F}}[\mathrm{CS}|R_2(s, \Psi, n)] = \min_{i \in \mathcal{G}} \inf_{\Omega_*(i,1)} P[B_i]$.

PROOF. It suffices to show $\inf_{\Omega_*(i)} P_{\mathbf{F}}[B_i] = \inf_{\Omega_*(i,1)} P_{\mathbf{F}}[B_i]$. When $k-s+1 \leqslant i \leqslant k-1$ and $\mathbf{F} \in \Omega_*(i, \xi)$ then

$$P_{\mathbf{F}}[B_i] = P_{\mathbf{F}}[\max\{T_{(j)}\rho^{e(n)}/\xi | k-i+1 \leqslant j \leqslant k\} \geqslant \max\{T_{(j)}/\xi | 1 \leqslant j \leqslant k-i\}]$$
$$= P_{\mathbf{F}'}[B_i] \qquad \text{where } \mathbf{F}'(x) = (F_1(x\xi), \cdots, F_k(x\xi)) \qquad \text{is in } \Omega_*(i, 1)$$

as noted in the proof of Lemma 3.1. The proof for the case $t \leqslant i \leqslant k-s$ is similar.

Now the events $\{A_i\}$ and $\{B_i\}$ possess similar monotonicity properties and so the argument previously used in the construction of the lower bound $F_n^*$ remains valid and yields

$$(4.6) \quad \inf_{\Omega_*} P_{\mathbf{F}}[\mathrm{CS}|R_2(s, \Psi, n)] \geqslant P_{\mathbf{F}(t)}[B_t] \equiv R_n^*(t, s, k, \alpha, c^*) = R_n^*, \qquad \text{say,}$$

where $\mathbf{F}(t)$ is defined in Section 3.2.

The lower bound will now be evaluated. Let $T_{[q]}$ be the $q$th ordered sample $\alpha$-quantile, $\mathfrak{D} \equiv \{k-t+1, \cdots, k\}$, $\overline{\mathfrak{D}} \equiv \{1, \cdots, k-t\}$, for $q > t$ let $\left\{\mathfrak{S}_\nu^q | 1 \leqslant \nu \leqslant \binom{k-t}{q-t}\right\}$ be the collection of all subsets of size $(q-t)$ from $\overline{\mathfrak{D}}$ and let $\overline{\mathfrak{S}}_\nu^q = \overline{\mathfrak{D}} - \mathfrak{S}_\nu^q$. Then

$$R_n^* = \sum_{q=k-s+1}^{k} P_{\mathbf{F}(t)}[W_t = T_{[q]}; B_t]$$
$$= \sum_{q=k-s+1}^{k} \sum_{j=k-t+1}^{k} P_{\mathbf{F}(t)}[W_t = T_{(j)} = T_{[q]}; T_{(j)}\rho^{e(n)} \geqslant \max\{T_{(m)}|1 \leqslant m \leqslant k-t\}]$$

$$= \Sigma_{q=k-s+1}^{k} \Sigma_{j=k-t+1}^{k} \Sigma_{\nu=1}^{\binom{k-t}{q-t}} P_{\mathbf{F}(t)} \begin{bmatrix} T_{(l)} < T_{(j)} \forall l \in \mathfrak{S}_{\nu}^{q} \cup \mathfrak{D} \setminus \{j\} \\ T_{(j)} < T_{(l)} < T_{(j)} \rho^{e(n)} \forall l \in \bar{\mathfrak{S}}_{\nu}^{q} \end{bmatrix}$$

$$= \Sigma_{q=k-s+1}^{k} \binom{k-t}{q-t} \int_0^\infty \left\{ B\left(F_{*,l}(y|c^*)\right)\right\}^{q-t}$$

$$\left\{ D_n\left(F_{*,l}(y|c^*)\right)\right\}^{k-q} d\left\{ B\left(F_{*,s}(y|1)\right)\right\}^{t}$$

$$= \Sigma_{q=k-s+1}^{k} \binom{k-t}{q-t} \begin{bmatrix} \int_{G(c^*)}^{\alpha} \{B(G(G^{-1}(y)/c^*))\}^{q-t} \left\{ B\left(G\left(\frac{G^{-1}(y)\rho^{e(n)}}{c^*}\right)\right) \right. \\ \left. - B\left(G\left(\frac{G^{-1}(y)}{c^*}\right)\right)\right\}^{k-q} d\{B(y)\}^{t} \\ + \left\{B\left(G\left(\frac{1}{c^*}\right)\right)\right\}^{q-t} \left\{D_n\left(G\left(\frac{1}{c^*}\right)\right)\right\}^{k-q} \{1 - B^t(\alpha)\} \end{bmatrix}$$

where $D_n(F(y)) = B(F(y\rho^{e(n)})) - B(F(y))$. As in the case of $R_1(s, n)$ this lower bound is tight.

THEOREM 4.1. $\inf_{\Omega_*} P_{\mathbf{F}}[\text{CS}|R_2(s, \psi, n)] = R_n^*$.

The computation of $\lim_{j\to\infty} P_{\mathbf{F}_j}[\text{CS}|R_2(s, \Psi, n)]$ for the sequence $\{\mathbf{F}_j\}$ displayed in the proof of Theorem 3.1 yields the result. Furthermore, $R_n^* \to 1$ as $n \to \infty$ and (4.4) can be guaranteed by finite $n$ for any $P^* < 1$.

EXAMPLE 4.3. When $\Omega^*$ is the class of all continuous IFRA distributions $(G(x) = 1 - e^{x\ln(1-\alpha)})$ then $R_n^*(1, s, k, \alpha, c^*)$ becomes

(4.7)
$$\Sigma_{q=k-s+1}^{k} \binom{k-1}{q-1} \begin{bmatrix} \int_{1-(1-\alpha)^{c^*}}^{\alpha} \{B(1-(1-x)^{1/c^*})\}^{q-1} \{B(1-(1-x)^{\rho^{e(n)/c^*}}) \\ - B(1-(1-x)^{1/c^*})\}^{k-q} \\ dB(x) + \{B(1-(1-\alpha)^{1/c^*})\}^{q-1}\{1 - B(\alpha)\} \\ \times \{B(1-(1-\alpha)^{\rho^{e(n)/c^*}}) - B(1-(1-\alpha)^{1/c^*})\}^{k-q} \end{bmatrix}.$$

Tables of the smallest odd sample sizes required to satisfy (4.4) were computed from (4.7) based on $\rho = 4$ and $e(n) = n^{-\frac{1}{2}}$ for $\alpha = .25, .50, .75$, $c^* = .65, .70$, $P^* = .75, .90, .95$, $k = 3(1)9$ and $s = 2(1)\min\{4, k - 1\}$.

4.3 *Selection from tail ordered families.* This section studies the design problem for $R_2(s, \Psi, n)$ when $\mathbf{F} \in \Omega_t$ and $\Psi_n(\xi) = \xi - e(n)$ where $\{e(n)\}$ is a sequence of positive numbers decreasing to zero. The developments follow from modifications

TABLE 4.1.
*Sample sizes required by $R_2(s, \Psi, n)$ when selecting from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\rho = 4$, $e(n) = n^{-\frac{1}{2}}$ and $\alpha = .25$.*

| | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|
| | | | $P^*$ | | | $P^*$ | |
| $k$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 3 | 2 | 27 | 71 | 107 | 43 | 103 | 155 |
| 4 | 3 | 31 | 75 | 115 | 47 | 111 | 167 |
| 4 | 2 | 35 | 83 | 127 | 55 | 123 | 183 |
| 5 | 4 | 35 | 83 | 127 | 51 | 123 | 183 |
| 5 | 3 | 39 | 87 | 127 | 55 | 123 | 187 |
| 5 | 2 | 43 | 95 | 139 | 65 | 139 | 203 |
| 6 | 4 | 39 | 91 | 135 | 59 | 131 | 195 |
| 6 | 3 | 43 | 95 | 139 | 63 | 135 | 199 |
| 6 | 2 | 51 | 107 | 151 | 75 | 155 | 223 |
| 7 | 4 | 43 | 99 | 143 | 67 | 143 | 207 |
| 7 | 3 | 47 | 103 | 147 | 69 | 147 | 211 |
| 7 | 2 | 57 | 115 | 163 | 85 | 167 | 235 |
| 8 | 4 | 47 | 103 | 151 | 71 | 151 | 219 |
| 8 | 3 | 51 | 107 | 155 | 75 | 159 | 223 |
| 8 | 2 | 63 | 123 | 171 | 93 | 179 | 251 |
| 9 | 4 | 51 | 107 | 155 | 77 | 159 | 227 |
| 9 | 3 | 55 | 115 | 163 | 83 | 167 | 235 |
| 9 | 2 | 67 | 131 | 179 | 99 | 191 | 263 |

TABLE 4.2.
*Sample sizes required by $R_2(s, \Psi, n)$ when selecting from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\rho = 4$, $e(n) = n^{-\frac{1}{2}}$ and $\alpha = .50$.*

| | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|
| | | | $P^*$ | | | $P^*$ | |
| $k$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 3 | 2 | 11 | 31 | 49 | 17 | 45 | 71 |
| 4 | 3 | 11 | 31 | 49 | 17 | 47 | 73 |
| 4 | 2 | 15 | 39 | 59 | 23 | 55 | 83 |
| 5 | 4 | 13 | 33 | 53 | 19 | 49 | 77 |
| 5 | 3 | 13 | 35 | 55 | 21 | 53 | 79 |
| 5 | 2 | 19 | 45 | 65 | 29 | 65 | 95 |
| 6 | 4 | 15 | 37 | 57 | 21 | 55 | 83 |
| 6 | 3 | 17 | 39 | 61 | 25 | 59 | 87 |
| 6 | 2 | 23 | 51 | 73 | 35 | 73 | 105 |
| 7 | 4 | 17 | 39 | 61 | 25 | 59 | 89 |
| 7 | 3 | 19 | 43 | 65 | 29 | 63 | 95 |
| 7 | 2 | 27 | 55 | 79 | 39 | 81 | 113 |
| 8 | 4 | 19 | 43 | 63 | 27 | 63 | 93 |
| 8 | 3 | 21 | 47 | 69 | 31 | 69 | 101 |
| 8 | 2 | 31 | 59 | 83 | 45 | 87 | 121 |
| 9 | 4 | 19 | 45 | 67 | 29 | 67 | 97 |
| 9 | 3 | 23 | 51 | 73 | 35 | 75 | 107 |
| 9 | 2 | 33 | 63 | 87 | 49 | 91 | 127 |

TABLE 4.3.
*Sample sizes required by $R_2(s, \Psi, n)$ from $\Omega_*$ when $G(x) = 1 - e^{-x}$ with $\rho = 4$, $e(n) = n^{-\frac{1}{2}}$ and $\alpha = .75$.*

| | | $c^* = .65$ | | | $c^* = .70$ | | |
|---|---|---|---|---|---|---|---|
| | | | $P^*$ | | | $P^*$ | |
| $k$ | $s$ | .75 | .90 | .95 | .75 | .90 | .95 |
| 3 | 2 | 7 | 23 | 35 | 11 | 31 | 51 |
| 4 | 3 | 7 | 23 | 35 | 11 | 31 | 51 |
| 4 | 2 | 11 | 27 | 43 | 17 | 39 | 59 |
| 5 | 4 | 7 | 23 | 35 | 11 | 35 | 51 |
| 5 | 3 | 11 | 27 | 39 | 15 | 35 | 55 |
| 5 | 2 | 15 | 33 | 47 | 21 | 47 | 69 |
| 6 | 4 | 11 | 27 | 39 | 15 | 35 | 55 |
| 6 | 3 | 11 | 27 | 43 | 17 | 41 | 63 |
| 6 | 2 | 17 | 37 | 53 | 25 | 53 | 77 |
| 7 | 4 | 11 | 27 | 43 | 15 | 39 | 59 |
| 7 | 3 | 13 | 31 | 47 | 19 | 45 | 67 |
| 7 | 2 | 21 | 41 | 57 | 29 | 59 | 83 |
| 8 | 4 | 13 | 31 | 43 | 19 | 43 | 63 |
| 8 | 3 | 15 | 33 | 49 | 21 | 49 | 71 |
| 8 | 2 | 21 | 43 | 61 | 33 | 63 | 89 |
| 9 | 4 | 13 | 31 | 47 | 21 | 47 | 67 |
| 9 | 3 | 17 | 37 | 53 | 25 | 53 | 75 |
| 9 | 2 | 25 | 47 | 65 | 37 | 67 | 93 |

of the arguments of Section 4.2 along the lines of Section 3.3 and hence only the final results will be stated.

THEOREM 4.2. $\inf_{\Omega_t} P[CS|R_2(s, \Psi, n)] = P_{\mathbf{F}(t)}[B_t] = R_n^t$ *where* $\mathbf{F}(t)$ *is as in Theorem* 3.2 *and* $R_n^t = R_n^t(t, s, k, \alpha, d^*)$ *is given by*

$$\Sigma_{q=k-s+1}^{k}\binom{k-t}{q-t}$$

$$\left[\begin{array}{c} \int_{G(-d^*)}^{\alpha}\{B(G[G^{-1}[x]+d^*])\}^{q-t}\{B(G[G^{-1}[x]+e(n)+d^*]) \\ -B(G[G^{-1}[x]+d^*])\}^{k-q}d\{B(x)\}^t \\ +\{B(G[d^*])\}^{q-t}\{B(G[e(n)+d^*])-B(G[d^*])\}^{k-q}\{1-B^t(\alpha)\} \end{array}\right].$$

*Furthermore* $R_n^t \to 1$ *as* $n \to \infty$ *and so* (4.4) *can be guaranteed for any* $P^* < 1$.

EXAMPLE 4.4. If $\Omega_t$ consists of all configurations with continuous components having "lighter" tails than the logistic distribution then $G(x) = [1 + \exp(-(x - \ln[(1-\alpha)/\alpha]))]^{-1}$. The lower bound $R_n^t$ on the probability of correctly selecting at least one $\pi_i$ having $\alpha$-quantile within $d^*$ of the largest $\alpha$ – quantile ($t = 1$) is

$$\sum_{q=k-s+1}^{k}\binom{k-1}{q-1}\Bigg[\int_{\alpha/(\alpha+(1-\alpha)e^{d^*})}^{\alpha}\left\{B\left(\frac{x}{x+(1-x)e^{-d^*}}\right)\right\}^{q-1}\left\{B\left(\frac{x}{x+(1-x)e^{-[d^*+e(n)]}}\right)-B\left(\frac{x}{x+(1-x)e^{-d^*}}\right)\right\}^{k-q}dB(x)$$
$$+\left\{B\left(\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}}\right)\right\}^{q-1}\{1-B(\alpha)\}\times\left\{B\left(\frac{\alpha}{\alpha+(1-\alpha)e^{-[d^*+e(n)]}}\right)-B\left(\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}}\right)\right\}^{k-q}\Bigg].$$

**5. Properties of $R_1(s, n)$ and $R_2(s, \Psi, n)$.** The purpose of this section is to describe some small sample performance characteristics of $R_1(s, n)$ and $R_2(s, \Psi, n)$ including (1) the effect of skewness and kurtosis of $R_1(s, n)$, (2) their monotonicity properties, (3) their performance as $c^* \to 1$ and some large sample properties including (a) a study of the number of populations selected by $R_2(s, \Psi, n)$ and (b) a proposal for choosing the sequence $\Psi$.

The first set of results show that the more skewed or the heavier the tails of the component $F_i$'s the smaller the probability of correct selection using $R_1(s, n)$. The analysis follows Doksum (1969) who used weak orderings to study the effects of skewness and kurtosis on the power of monotone rank tests. It has previously been established in Remark 3.2 that the "closer together" the preferred and nonpreferred populations are stochastically, the smaller the PCS. In studying the effects of skewness and kurtosis it is necessary to eliminate this stochastic effect by restricting attention to configurations of distributions differing only in scale or location. In this set up the resulting scale or location parameters are measures of "stochastic distance." For $1 \leqslant t < k$ and $i \in \mathcal{I} = \{i | t \leqslant i < r + k - s\}$ let $\mathfrak{D}_i \equiv \{\mathbf{\Delta}_i = (\delta_1, \cdots, \delta_k) | \infty > \delta_1 \geqslant \delta_2 \geqslant \cdots \geqslant \delta_{k-i} = 1 > \delta_{k-i+1} \geqslant \cdots \geqslant \delta_k > 0$ and $c^* \geqslant \delta_{k-t+1} > c^*\delta_{k-i+1}\}$ and for $F \in \mathcal{F}_0$ and $\mathbf{\Delta}_i \in \mathfrak{D}_i$ let $\mathbf{F}(\mathbf{\Delta}_i) = (F_{\delta_1}, F_{\delta_2}, \cdots, F_{\delta_k})$ where $F_{\delta_j}(x) = F(\delta_j x)$.

REMARK 5.1. If $F \in \mathcal{F}_0$ has $F(\xi) = \alpha$ then $\chi_\alpha(F_{\delta_j}) = \xi/\delta_j$. Hence $\mathbf{F}(\mathbf{\Delta}_i)$ contains exactly $i$ preferred populations and $F_{[j]} = F_{\delta_j}$.

THEOREM 5.1. *If $F, H \in \mathcal{F}_0$ satisfy $F \prec_* H$ and either $H(x) < 1 \forall x \in R$ or $F(x) < 1 \forall x \in R$ then $\forall i \in \mathcal{I}$ and $\Delta_i \in \mathfrak{D}_i$*

$$P_{\mathbf{F}(\mathbf{\Delta}_i)}[\mathrm{CS}|R_1(s, n)] \geqslant P_{\mathbf{H}(\mathbf{\Delta}_i)}[\mathrm{CS}|R_1(s, n)].$$

PROOF. Suppose $T_{(j)}$ has cdf $B(F_{\delta_j})$, $1 \leqslant j \leqslant k$, and define

(5.1) $$T'_{(j)} = [B(H)]^{-1}B(F(T_{(j)})) = H^{-1}F(T_{(j)})$$

and

(5.2) $$T''_{(j)} = \left[B(H_{\delta_j})\right]^{-1}B\left(F_{\delta_j}(T_{(j)})\right) = H^{-1}F(\delta_j T_{(j)})/\delta_j.$$

It follows that $T''_{(j)}$ has cdf $B(H_{\delta_j})$ for $1 \leqslant j \leqslant k$ and $T'_{(k-i)} = T''_{(k-i)}$ has cdf $B(H)$ since $\delta_{k-i} = 1$.

Since $\delta_j \in (0, 1)$ for $k - i + 1 \leqslant j \leqslant k$ it follows from Lemma 2.2 that

(5.3) $$T'_{(j)} = H^{-1}F(T_{(j)}) \geqslant H^{-1}F(\delta_j T_{(j)})/\delta_j = T''_{(j)}$$

provided $T_{(j)}$ and $\delta_j T_{(j)} \in R^+(F)$. Now $\delta_j T_{(j)}$ has dist. $B(F)$ hence $\delta_j T_{(j)} \in R^+(F)$ a.s. $(F_{\delta_j})$. If $F(x) < 1 \forall x < \infty$ then $T_{(j)} \in R^+(F) = (0, \infty)$ while if $H(x) < 1 \forall x < \infty$ then $T_{(j)} \notin R^+(F) \Rightarrow F(T_{(j)}) = 1 \Rightarrow T'_{(j)} = H^{-1}(1) = \infty$. In either case (5.3) holds a.s. $(F_{\delta_j})$. Since $1 \leqslant \delta_j$ for $1 \leqslant j \leqslant k - i$ it also follows from $0 \leqslant T_{(j)} \leqslant \delta_j T_{(j)}$ that

(5.4) $$T'_{(j)} = H^{-1}F(T_{(j)}) \leqslant H^{-1}F(\delta_j T_{(j)})/\delta_j = T''_{(j)}$$

holds a.s. $(F_{\delta_j})$.

Because $\chi_{A_i}(\cdot)$ depends only on the ranks of the $T_{(i)}$'s and $H^{-1}F(\cdot)$ is strictly increasing on $I(F)$ and nondecreasing on $(0, \infty)$ it can be seen that

(5.5) $$\chi_{A_i}(T_{(1)}, \cdots, T_{(k)}) = \chi_{A_i}(T'_{(1)}, \cdots, T'_{(k)})$$

hence

(5.6) $$\chi_{A_i}(T_{(1)}, \cdots, T_{(k)}) \geqslant \chi_{A_i}(T''_{(1)}, \cdots, T''_{(k)})$$

holds a.s. $(\mathbf{F}(\Delta_i))$ from (5.3), (5.4) and (5.5). Taking expectations of both sides of (5.6) w.r.t. $\mathbf{F}(\Delta_i)$ and recalling that $T''_{(j)}$ has cdf $B(H_{\delta_j})$ completes the proof.

REMARK 5.2. If $R^+(F) = (0, b_F)$ and $R^+(H) = (0, b_H)$ where $\max(b_F, b_H) < \infty$ then (5.3) need not hold for $\delta_j \in (0, 1)$ and hence the proof is not valid in this case.

The results describing the effects of kurtosis on $P_{\mathbf{F}}[\text{CS}|R_1(s, n)]$ will now be stated. For $F, H \in \mathcal{F}$ recall that $F \prec_t H$ means that $H$ has heavier tails than $F$. For $\theta \in R$ and $F \in \mathcal{F}$ let $F_\theta(x) = F(x + \theta)$ and for $i \in \mathcal{I}$ define $\mathcal{F}_i = \{\Theta_i = (\theta_1, \cdots, \theta_k) | \infty > \theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_{k-i} = 0 > \theta_{k-i+1} \geqslant \cdots \geqslant \theta_k > -\infty; -d^* \geqslant \theta_{k-t+1} > \theta_{k-i+1} - d^*\}$ and for $F \in \mathcal{F}$ and $\Theta_i \in \mathcal{F}_i$ let $\mathbf{F}(\Theta_i) = (F_{\theta_1}, \cdots, F_{\theta_k})$. The configuration $\mathbf{F}(\boldsymbol{\theta}_i)$ contains exactly $i$ preferred populations. The following theorem shows the heavier the tails of the population distributions the smaller the probability of a correct selection.

THEOREM 5.2. *If $F, H \in \mathcal{F}$ satisfy $F \prec_t H$ and either $F(x) < 1 \forall x \in R$ or $H(x) < 1 \forall x \in R$ then $\forall i \in \mathcal{I}$ and $\Theta_i \in \mathcal{F}_i$*

$$P_{\mathbf{F}(\Theta_i)}[\text{CS}|R_1(s, n)] \geqslant P_{\mathbf{H}(\Theta_i)}[\text{CS}|R_1(s, n)].$$

REMARK 5.3. The proofs of Theorems 5.1 and 5.2 are not valid for the procedure $R_2(s, \Psi, n)$ because the characteristic functions of the events $B_i$ do not depend only on the ranks of the $T_{(j)}$'s.

The monotonicity properties of $R_1(s, n)$ and $R_2(s, \Psi, n)$ will next be investigated. Fix any $r, 1 \leqslant r \leqslant \min\{t, s\}$ when referring to $R_1(s, n)$ and $r = 1$ when referring to $R_2(s, \Psi, n)$. Given an arbitrary selection procedure $\mathcal{P}$ and any integer $i$ satisfying $r \leqslant i \leqslant k$ let

$$P_{\mathbf{F}}[S_i|\mathcal{P}] = P_{\mathbf{F}}[\mathcal{P} \text{ selects at least } r \text{ of } \pi_{(j_1)}, \cdots, \pi_{(j_i)}] \tag{5.7}$$

for any $S_i \equiv \{j_1, \cdots, j_i\} \subset \{1, \cdots, k\}$ of size $i$. For a given $S_i$ let $j_{[1]} < \cdots < j_{[i]}$ denote the ordered components of $S_i$.

DEFINITION 5.1. $\mathcal{P}$ is *monotone* w.r.t. $\Omega$ iff $\forall i, r \leqslant i \leqslant k$ and $\mathbf{F} \in \Omega$ $P_{\mathbf{F}}[S_i|\mathcal{P}] \leqslant P_{\mathbf{F}}[S_i'|\mathcal{P}]$ whenever $j_{[l]} \leqslant j_{[l]}' \forall l, 1 \leqslant l \leqslant i$.

Recall that for a given set of configurations $\Omega$, $\rho(\cdot)$ satisfying (3.1) – (3.4) and $t < k$ that $\Omega(i), t \leqslant i \leqslant k$, is the subset of $\Omega$ containing exactly $i$ preferred populations.

DEFINITION 5.2. $\mathcal{P}$ is *weakly monotone* w.r.t. $\Omega$ iff $\forall i, t \leqslant i \leqslant k$ and $\mathbf{F} \in \Omega(i)$, $P_{\mathbf{F}}[S_i|\mathcal{P}] \leqslant P_{\mathbf{F}}[S_i'|\mathcal{P}]$ whenever $S_i - (S_i \cap S_i') \subset \{1, \cdots, k - i\}$ and $S_i' - (S_i \cap S_i') \subset \{k - i + 1, \cdots, k\}$ where $-$ denotes set subtraction.

REMARK 5.4. The standard definitions of monotonicity used in parametric problems differ from those given here. In particular, Definition 4.1 of monotonicity in Santner (1975) is a special case ($i = r = 1$) of the present definition. Also note that monotonicity implies weak monotonicity.

LEMMA 5.1. *Any procedure $\mathcal{P}$ in $\{R_1(s, n)|n \geqslant N_\alpha\}$ or in $\{R_2(s, \Psi, n)|n \geqslant N_\alpha\}$ must satisfy $P_{\mathbf{F}}[S_i|\mathcal{P}] \leqslant P_{\mathbf{F}'}[S_i|\mathcal{P}] \forall i, r \leqslant i \leqslant k$ and $\mathbf{F}, \mathbf{F}' \in \Omega(\mathcal{F}) = \mathcal{F}^k$ such that $F_{[j]} \prec_{st} F_{[j]}' \forall j \in S_i$ and $F_{[j]}' \prec_{st} F_{[j]} \forall j \in \bar{S}_i \equiv \{1, \cdots, k\} - S_i$. The result is immediate since $[S_i|R_1(s, n)] = [R_1(s, n)$ selects at least $r$ of $\pi_{(j)}, j \in S_i]$ and $[S_i|R_2(s, \Psi, n)] = [R_2(s, \Psi, n)$ selects at least one of $\pi_{(j)}, j \in S_i]$ are nondecreasing in any of $T_{(j)}, j \in S_i$ and nonincreasing in any of $T_{(j)}, j \in \bar{S}_i$.*

Some additional notation will now be introduced. Let

$$\Omega_{st} = \{\mathbf{F} \in \Omega(\mathcal{F})|F_{[1]} \prec_{st} F_{[2]} \prec_{st} \cdots \prec_{st} F_{[k]}\},$$

$$\Omega_R(i) = \{\mathbf{F} \in \Omega(\mathcal{F})|x_\alpha(F_{[k-i]}) \leqslant p(x_\alpha(F_{[k-t+1]})) < x_\alpha(F_{[k-i+1]});$$
$$\min\{F_{[l]}(y)|1 \leqslant l \leqslant k - i\} \geqslant \max\{F_{[l]}(y)|k - i < l \leqslant k\} \forall y \in R\} \quad \text{for } t \leqslant i \leqslant k - 1,$$
$$= \{\mathbf{F} \in \Omega(\mathcal{F})|p(x_\alpha(F_{[k-t+1]})) < x_\alpha(F_{[1]})\} \text{ for } i = k,$$

and $\Omega_R = \cup_{i=t}^k \Omega_R(i)$. Note that $\Omega_{st} \subset \Omega_R$.

THEOREM 5.3. *The classes of procedures $\{R_1(s, n)|n \geqslant N_\alpha\}$ and $\{R_2(s, \Psi, n)|n \geqslant N_\alpha\}$ are all* monotone w.r.t. *any $\Omega_1 \subset \Omega_{st}$ and* weakly monotone w.r.t. *any $\Omega_2 \subset \Omega_R$.*

PROOF. To show monotonicity it suffices to consider the following case. Fix $\mathbf{F} \in \Omega_1$, $r \leqslant i \leqslant k$, $S_i \subset \{1, \cdots, k\}$, $j \in S_i$ and $m > j$ with $m \in \bar{S}_i$ and let $S_i' = S_i \cup \{m\} - \{j\}$. Lemma 5.1 implies

$$(5.8) \qquad \begin{aligned} P_{\mathbf{F}}[S_i | R_1(s, n)] &= P_{(F_{[1]}, \cdots, F_{[j]}, \cdots, F_{[m]}, \cdots, F_{[k]})}[S_i | R_1(s, n)] \\ &\leqslant P_{(F_{[1]}, \cdots, F_{[m]}, \cdots, F_{[m]}, \cdots, F_{[k]})}[S_i | R_1(s, n)] \\ &= P_{(F_{[1]}, \cdots, F_{[m]}, \cdots, F_{[m]}, \cdots, F_{[k]})}[S_i' | R_1(s, n)] \\ &\leqslant P_{\mathbf{F}}[S_i' | R_1(s, n)]. \end{aligned}$$

Since $R_2(s, \Psi, n)$ satisfies Lemma 5.1 the same proof also shows that it is monotone w.r.t. $\Omega_1$. The proofs that $R_1(s, n)$ and $R_2(s, \Psi, n)$ are weakly monotone w.r.t. $\Omega_2$ are similar.

REMARK 5.5. In most parametric selection problems where the scalar of interest, say $\lambda_i$, is estimated by $T_i$ having cdf $F(\cdot|\lambda_i)$ the following holds: $\lambda_i < \lambda_j \Rightarrow F(\cdot|\lambda_i) \prec_{st} F(\cdot|\lambda_j)$ and hence the stronger property of monotonicity holds. However $x_\alpha(F_i) < x_\alpha(F_j) \not\Rightarrow F_i \prec_{st} F_j$ and hence the property of weak monotonicity is introduced here. An example will be given later to show that neither $R_1(s, n)$ nor $R_2(s, \psi, n)$ is even weakly monotone w.r.t. $\Omega_*$ or $\Omega_t$.

REMARK 5.6. Since $\mathcal{F}_I \subset \mathcal{F}$ for any internal $I \subset R$, Theorem 5.3 holds if $\Omega_*$ and $\Omega_k$ are defined as the appropriate subsets of $\mathcal{F}_I^k$ rather than $\mathcal{F}^k$.

The next result, obtainable from straightforward computations, describes the behavior of the infimum of $P[\mathrm{CS}|R_1(s, n)]$ over $\Omega_*$ as $c^*$ increases to 1.

LEMMA 5.2. $\lim_{c^* \to 1^-} \inf_{\Omega_*} P[\mathrm{CS}|R_1(s, n)] = B[B[\alpha];\, k - t - s + r,\, k - t] \times (1 - B[B(\alpha);\, t - r + 1,\, t]) < (k - t - s + r - 1) \binom{k-t}{s-j} \Sigma_{j=r}^{t} \binom{t}{j-r} / (k - t - s + j - 1) \binom{k}{s+t-j}$.

When $r = t$ the right hand side reduces to $\binom{k-t}{s-t} / \binom{k}{s}$ which is the probability of making a correct selection by randomly selecting $s$ populations when exactly $t$ are preferred. So for any fixed sample size $n \geqslant N_\alpha$ there exists $c^*$ sufficiently close to 1 and an $\mathbf{F} \in \Omega_*(t)$ for which $P_{\mathbf{F}}[\mathrm{CS}|R_1(s, n)] < P_{\mathbf{F}}[\mathrm{CS}|$ choose $s$ populations at random]. To gain an intuitive feel for this result choose an arbitrary $\mathbf{F} \in \Omega_*(t, 1)$. Then $\forall x \in [c^*, 1]$

$$(5.9) \qquad F_{[j]}(x) > F_{[i]}(x) \forall 1 \leqslant j \leqslant k - t \quad \text{and} \quad k - t + 1 \leqslant i \leqslant k.$$

However (5.9) need not hold for $x \notin [c^*, 1]$. In fact given $\varepsilon > 0$ there exists $\mathbf{F} = \mathbf{F}(\varepsilon) \in \Omega_*(t, 1)$ such that for all $x \notin [c^* - \varepsilon, 1 + \varepsilon]$

$$(5.10) \qquad F_{[j]}(x) < F_{[i]}(x) \forall 1 \leqslant j \leqslant k - t \quad \text{and} \quad k - t + 1 \leqslant i \leqslant k.$$

In such a case unless $n$ is sufficiently large to detect the situation in $[c^*, 1]$ the procedure $R_1(s, n)$ will tend to reflect (5.10) and choose populations in $\{\pi_{(1)}, \cdots, \pi_{(k-t)}\}$.

The preceding argument can be used to show that $R_1(t, n)$ is not weakly monotone w.r.t. $\Omega_*$. For fixed $n$ and $c^*$ sufficiently close to 1 choose $\mathbf{F} \in \Omega_*(t)$ such that

$$P_{\mathbf{F}}[\mathrm{CS}|R_1(t, n)] = P_{\mathbf{F}}[\mathbf{k}_t|R_1(t, n)] < 1/\binom{k}{t} \tag{5.11}$$

where $\mathbf{k}_t = \{k - t + 1, \cdots, k\}$. Let $\{\mathbf{j}(l)|1 \leqslant l \leqslant \binom{k}{t}\}$ be the collection of all sets of size $t$ from $\{1, \cdots, k\}$. Then

$$\Sigma_{l=1}^{\binom{k}{t}} P_{\mathbf{F}}[\mathbf{j}(l)|R_1(t, n)] = 1$$

and (5.11) $\Rightarrow$ there exists a $j^*(l)$ such that

$$P_{\mathbf{F}}[\mathbf{j}^*(l)|R_1(t, n)] > 1/\binom{k}{t} > P_{\mathbf{F}}[\mathbf{k}_t|R_1(t, n)]$$

and hence $R_1(t, n)$ is not weakly monotone. Similar examples exist when $\Omega_*$ and $R_1(s, n)$ are replaced by $\Omega_t$ and $R_2(s, \Psi, n)$ respectively.

Some results will now be developed for the number of populations selected by $R_2(s, \Psi, n)$. Let $E(i, n) = [T_{(i)} \geqslant \max\{T_{[k-s+1]}, \psi_n^{-1}(T_{[k]})\}]$ be the event that $\pi(i)$ is selected, $P(i, n) = E_{\mathbf{F}}[E(i, n)]$ and $S(n) = \Sigma_{i=1}^{k}\chi_{E(i,n)}(\mathbf{T})$ the number of populations selected by $R_2(s, \Psi, n)$. It is straightforward to obtain an expression for $E_{\mathbf{F}}[S(n)]$ for arbitrary $\mathbf{F}$ based on this representation for $S(n)$.

For all $F \in \mathcal{F}$ and $\alpha \in (0, 1)$ it is well known that $T_{(i),n} \to_p \chi_\alpha(F)$ where $T_{(i),n}$ has cdf $B[F(y); q, n](n \geqslant N_\alpha)$. The following results are a direct consequence of this fact and the proof of Theorem 5.2 and Corollaries 5.1 and 5.2 of Santner (1975).

THEOREM 5.4. *For any* $\mathbf{F} \in \Omega(\mathcal{F})$ *such that* $\chi_\alpha(F_{[k]} > \chi_\alpha(F_{[k-1]})$

(1) $P_F(i, n) \to \begin{matrix} 1 & , i = k \\ 0 & , 1 \leqslant i < k \end{matrix}$ *as* $n \to \infty$,

(2) $Z_i(n) \to_{L^2} \begin{matrix} 1 & , i = k \\ 0 & , 1 \leqslant i < k \end{matrix}$ *as* $n \to \infty$ *and*

(3) $S(n) \to_{L^2} 1$ *and* $E_{\mathbf{F}}[S(n)] \to 1$ *as* $n \to \infty$ *where* $\to_{L^2}$ *denotes convergence in the* $L^2$ *norm.*

The final topic of this section is a proposal for choosing the constant $\rho \in (1, \infty)$ when the sequence $\Psi_n(\xi) = \xi\rho^{e(n)}$ is used to determine the rule $R_2(s, \Psi, n)$. The discussion is limited to the star ordered case but similar results hold for the tail ordered case.

Let $[\mathrm{NPS}|R_2(s, \Psi, n)] = [R_2(s, \Psi, n)$ selects at least one nonpreferred population$]$ and for $\rho \in (1, \infty)$ let $n(\rho)$ be the smallest $n \geqslant N_\alpha$ satisfying both

$$\inf_{\Omega_*} P_{\mathbf{F}}[\mathrm{CS}|R_2(s, \Psi, n)] \geqslant P^* \tag{5.12}$$

and

$$\sup_{\Omega_*} P_{\mathbf{F}}[\text{NPS}|R_2(s, \Psi, n)] \leqslant \varepsilon^* \tag{5.13}$$

for specified $\varepsilon^* \in (0, 1)$ and $P^* \in (1/k, 1)$. The optimal choice of $\rho$ minimizes $n(\rho)$. Arguments similar to those of Section 4.2 show the supremum of (5.13) occurs at $\mathbf{F}(1, 1)$. Calculation gives

$$P_{\mathbf{F}(1,1)}[\text{NPS}|R_2(s, \Psi, n)] = 1 - \int_0^\infty \{B(F_{*,l}(y/\rho^{e(n)}|c^*))\}^{k-1} dB(F_{*,s}(y|1))$$

$$= \begin{cases} 1 - \int_{c^*\rho^{n^{-\frac{1}{2}}}}^{1} \left[B\left(G\left(y/c^*\rho^{n^{-\frac{1}{2}}}\right)\right)\right]^{k-1} dB(G(y)) \\ -\left[B\left(G\left(1/c^*\rho^{n^{-\frac{1}{2}}}\right)\right)\right]^{k-1}[1 - B(\alpha)], & c^*\rho^{n^{-\frac{1}{2}}} \geqslant 1 \\ 1 & , c^*\rho^{n^{-\frac{1}{2}}} < 1. \end{cases} \tag{5.14}$$

REMARK 5.7. For fixed $\rho \in (1, \infty)$ the righthand side of (5.14) converges to zero as $n \to \infty$ and hence $\forall \varepsilon^* \in (0, 1)$ there exists $n \geqslant N_\alpha$ satisfying (5.13).

For each $\rho \in (1, \infty)$ let $n_1(\rho)$ be the smallest $n \geqslant N_\alpha$ satisfying (5.12) and let $n_2(\rho)$ be the smallest $n \geqslant N_\alpha$ satisfying (5.13). Since both $\sup_{\Omega_*} P_{\mathbf{F}}[\text{NPS}|R_2(s, \Psi, n)]$ and $\inf_{\Omega_*} P_{\mathbf{F}}[\text{CS}|R_2(s, \Psi, n)]$ are nondecreasing in $\rho$ it follows that $n_1(\rho)$ is nonincreasing in $\rho$ and $n_2(\rho)$ is nondecreasing in $\rho$. An optimal choice of $\rho$ is $\rho^*$ such that

$$\max\{n_1(\rho^*), n_2(\rho^*)\} = \min_{1<\rho<\infty} \max\{n_1(\rho), n_2(\rho)\}. \tag{5.15}$$

The solution of (5.15) can be obtained via the same techniques employed in the construction of Tables 4.1–4.3 (see Hooper (1977)).

## REFERENCES


[1] ALAM, K. and RIZVI, M. H. (1966). Selection from multivariate normal populations. *Ann. Inst. Statist. Math.* **18** 307–318.

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

[3] BARLOW, R. E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, N.Y.

[4] BARTLETT, N. S. and GOVINDARAJULU, Z. (1968). Some distribution-free statistics and their application to the selection problem. *Ann. Inst. Statist. Math.* **20** 79–97.

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

[6] BHAPKAR, V. P. and GORE, A. P. (1971). Some selection procedures based on $U$-statistics for the location and scale problems. *Ann. Inst. Statist. Math.* **23** 375–386.

[7] CARROLL, R. J. (1974). Some contributions to the theory of parametric and nonparametric sequential ranking and selection. Ph.D. Thesis, Dept. of Statist., Purdue University.

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

[9] BIRNBAUM, Z. W., ESARY, J. D. and MARSHALL, A. W. (1966). A stochastic characterization of wear-out for components and systems. *Ann. Math. Statist.* **37** 816–825.

[10] DESU, M. and SOBEL, M. (1971). Nonparametric procedures for selecting fixed size subsets. In *Statistical Decision Theory and Related Topics*. (ed. S. S. Gupta and J. Yackel). 255–274. Academic Press, New York.

[11] DOKSUM, K. (1969). Starshaped transformations and the power of rank tests. *Ann. Math. Statist.* **40** 1167–1176.

[12] EATON, MORRIS L. (1967). Some optimum properties of ranking procedures. *Ann. Math. Statist.* **36** 124–137.

[13] ESARY, J. D., MARSHALL, A. W. and PROSCHAN, F. (1973). Shock models and wear processes. *Ann. Probability* **1** 627–649.

[14] GHOSH, M. (1973). Nonparametric selection procedures for symmetric location parameter problems. *Ann. Statist.* **1** 773–779.

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

[16] GUPTA, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics* **7** 225–245.

[17] GUPTA, S. S. and HUANG, D. Y. (1974). Nonparametric subset selection procedures for the $t$ best populations. *Bull. Inst. Math. Acad. Sinica* **2** 377–386.

[18] GUPTA, S. S. and MCDONALD, G. C. (1970). On some classes of selection procedures based on ranks. In *Nonparametric Techniques in Statistical Inference*, (ed. M. L. Puri) Univ. Cambridge Press, Cambridge.

[19] GUPTA, S. S. and SANTNER, T. J. (1973). On selection and ranking procedures – a restricted subset selection rule. *Proc. 39th Session Internat. Statist. Inst.* Vienna, Austria **1** 409–417.

[20] HOOPER, J. H. (1977). Selection procedures for ordered families of distributions. Ph.D. Thesis, School of Operations Research and Industrial Engineering, Cornell University.

[21] LEE, Y. J. and DUDEWICZ, E. J. (1974). Nonparametric ranking and selection procedures. Technical Report No. 105, Department of Statistics, Ohio State Univ.

[22] LEHMANN, E. L. (1963). A class of selection procedures based on ranks. *Math. Ann.* **150** 268–275.

[23] MAHAMUNULU, D. M. (1967). Some fixed-sample ranking and selection problems. *Ann. Math. Statist.* **38** 1079–1091.

[24] PANCHAPAKESAN, S. (1969). A class of selection procedures including procedures for restricted families of distributions. Mimeograph series No. 189, Depart. Statist., Purdue Univ.

[25] PATEL, J. K. (1976). Ranking and selection of IFR populations based on means. *J. Amer. Statist. Assoc.* **71** 143–146.

[26] PURI, M. L. and SEN, P. K. (1972). On some selection procedures in two-way layouts. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **22** 242–250.

[27] RANDLES, R. H. (1970). Some robust selection procedures. *Ann. Math. Statist.* **41** 1640–1645.

[28] RIZVI, M. H. and SOBEL, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest $\alpha$-quantile. *Ann. Math. Statist.* **38** 1788–1803.

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

[30] SANTNER, T. J. (1975). A restricted subset selection approach to ranking and selection problems. *Ann. Statist.* **3** 334–349.

[31] SHAMPINE, L. and ALLEN, R. (1973). *Numerical Computing: an Introduction*. W. B. Saunders Company, Philadelphia.

[32] SOBEL, M. (1967). Nonparametric procedures for selecting the $t$ populations with the largest $\alpha$-quantiles. *Ann. Math. Statist.* **38** 1804–1816.

[33] VAN ZWET, W. R. (1964). *Convex Transformations of Random Variables. Mathematical Centre Tract* 7, Math. Centrum, Amsterdam.

QUALITY THEORY AND RELIABILITY DEPT.
BELL LABORATORIES
BUILDING WB, RM IJ303
CRAWFORDS CORNER RD.
HOLMDEL, N. J. 07733

BIOMETRY BRANCH
NATIONAL CANCER INSTITUTE
LANDOW BUILDING, 5C-09
BETHESDA, MARYLAND 20014