[1] This paper is based in part on the author's 1974 Ph.D. dissertation at Purdue University. This research was supported in part by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University.

AMS 1970 Subject Classifications: Primary 62G31; Secondary 62G45, 62G70

Key Words and Phrases: Ranking and Selection; Sequential Selection; Tail Orderings; Stochastic Orderings; Quantiles; Nonparametric Selection.

ASYMPTOTICALLY NONPARAMETRIC SEQUENTIAL SELECTION PROCEDURES[1]

by

Raymond J. Carroll

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series #944
August 1974

ASYMPTOTICALLY NONPARAMETRIC SEQUENTIAL SELECTION PROCEDURES[1]

by

Raymond J. Carroll

University of North Carolina

SUMMARY

Let $\Pi_1,\ldots,\Pi_K$ be K independent populations with distributions $F(x;\theta_i)$ $(i = 1,\ldots,K)$ which are stochastically ordered or tail ordered. The ranking goal is to select the stochastically largest population or the population with the lightest tail. Sequential selection rules in the spirit of Robbins, Sobel, and Starr (1968) and Geertsema (1972) are proposed and studied. The above problems are corollaries of a general Theorem proposed, and the solutions are nonparametric in an asymptotic sense.

[1]This paper is based in part on the author's 1974 Ph.D. dissertation at Purdue University. This research was supported in part by the Office of Naval Research Contract N00014-67-A-0226-00014 at Purdue University.

1. INTRODUCTION

Let $\Pi_1,\ldots,\Pi_K$ be K independent populations with distributions $F(x;\theta_i)$ $(i = 1,\ldots,K)$, where θ_i is some unknown indexing parameter. If $\theta_i \lesssim \theta_j$ indicates that θ_j is preferred to θ_i , the most fundamental problem in ranking theory revolves around finding the unknown population $\Pi_{[K]}$ which is most preferred. For example, if $F(x;\theta_i) = F(x-\theta_i)$, then $\theta_i \lesssim \theta_j$ could mean $\theta_i \le \theta_j$, in which case the goal would be to select the largest location parameter.

Suppose the true (unknown) preference ordering is $\theta_{[1]} \lesssim \cdots \lesssim \theta_{[K]}$. Using a natural decision procedure, Bechhofer's (1954) indifference zone method consists of determining the smallest sample size $n(\delta)$ for which the probability of correctly selecting the preferred population whenever $d\big(F(x;\theta_{[i]}), F(x;\theta_{[K]})\big) \ge \delta > 0$ is at least as large as a specified constant P^* , where $\delta > 0$ is known and $d(\circ,\circ)$ is some distance function. For the normal means case, using sample means,

(1.1) $$d\big(F(x;\theta_i), F(x;\theta_j)\big) = \theta_j - \theta_i$$

(1.2) $$N(\delta) = \text{first integer } n \ge (b\sigma/\delta)^2 \text{ , where}$$

(1.3) $$P^* = \int_{-\infty}^{\infty} \Phi^{K-1}(x+b)\,d\Phi(x) \text{ .}$$

For future use, define

(1.4) $$\Omega(\delta) = \{(\theta_1,\ldots,\theta_K) \mid \theta_{[K]} - \theta_{[i]} \ge \delta > 0 \quad i = 1,\ldots,K-1\} \text{ .}$$

If σ^2 is unknown, Robbins, Sabel, and Starr (1968) and Geertsema (1972) investigate and generalize the natural rule which takes $N(\delta)$ observations from each population, where

$$N(\delta) = \text{first integer } n \geq (bS_{nK}/\delta)^2 \tag{1.5}$$

and S_{nK}^2 is the usual pooled sample variance based on $K(n - 1)$ degrees of freedom. Geertsema (1972), using translation invariant estimates, derives asymptotic (as $\delta \to 0$) results which are nonparametric in the sense that, under certain conditions, they hold independently of the underlying distribution F .

$N(\delta)$ is a stopping rule of Chow and Robbins type (1965); in order to investigate its asymptotic properties, the following definition is needed:

<u>DEFINITION 1.1</u> (Anscombe (1952)) A sequence of r.v.'s $\{\gamma_n\}$ is said to be *uniformly continuous in probability* if $\forall\ \varepsilon > 0$, $\sigma > 0$, $\exists\ J(\varepsilon,\sigma)$ and $c(\varepsilon,\sigma) \ni$ if $n \geq J(\varepsilon,\sigma)$

$$P\{|\gamma_m - \gamma_n| < \sigma\omega_n^{-1}\ \forall \text{ integers } m \ni |m - n| < n\, c(\varepsilon,\sigma)\} > 1 - \varepsilon , \tag{1.6}$$

where ω_n is a sequence of norming constants such that for some real θ and d.f. F , $P\{\gamma_n - \theta \leq x\omega_n^{-1}\} \to F(x)$.

The sample means satisfy (1.6) so that from Anscombe (1952), as $\delta \to 0$,

$$\delta^{-1}b(\overline{X}_{i,N(\delta)} - \theta_i) \xrightarrow{L} \Phi , \tag{1.7}$$

where "$\xrightarrow{L}$" indicates convergence in law. Hence

(1.8) $$\liminf_{\delta \to 0} P(CS) = P^* ,$$

where CS indicates a correct selection and the inf is taken over $\Omega(\delta)$.

This paper attempts to generalize the location parameter results to such preference patterns as stochastic ordering (Lehmann (1959)) and tail ordering (Doksum (1969)). In Section 2, the sequential rules to be used are presented. Section 3 motivates and presents the results and includes some applications. Section 4 gives the proofs.

2. STRUCTURE

The initial structure of Section 1 will be assumed. For $i = 1,\ldots,K$, independent observations $X_{i1},\ldots,X_{in}$ are taken according to a distribution $F(\cdot;\ \theta_i)$ and statistics $T_i(n)$, $g_i(n)$ are formed such that

(2.1) $$\frac{\sqrt{n}}{\sigma(\theta_i)} \left(T_i(n) - \mu(\theta_i)\right) \xrightarrow{L} \Phi \quad \text{as} \quad n \to \infty$$

(2.2) $$T_i(n) \to \mu(\theta_i) \quad \text{a.s.}$$

(2.3) $$g_i(n) \to \sigma(\theta_i) \quad \text{a.s., and} \quad g_i(n) > 0 \quad \text{a.s.} ,$$

where $\mu(\theta_i) = \mu_i$ and $\sigma(\theta_i) = \sigma_i$ are constants. The preference pattern and distance function will be

(2.4) $$\theta_i \precsim \theta_j \iff \mu_i \le \mu_j ,$$

(2.5) $$d\left(F(x;\ \theta_{[K]})\ ;\ F(x;\ \theta_{[i]})\right) = \mu(\theta_{[K]}) - \mu(\theta_{[i]}) ,$$

and $\Omega(\delta)$ will be as in (1.4), with the obvious replacement of $\mu(\theta_{[K]})$ for $\theta_{[K]}$. Since no simple location parameter structure is assumed, there will be K independent stopping rules $N_1(\delta),\ldots,N_K(\delta)$ defined by

(2.6) $N_i(\delta)$ = first integer $n \geq \left(bg_i(n)/\delta\right)^2$ $(i = 1,\ldots,K)$.

For $i = 1,\ldots,K$, we will take $N_i(\delta)$ observations from π_i, form $T_i\left(N_i(\delta)\right)$, and select the population with the largest $T_i\left(N_i(\delta)\right)$. Examples considered in this paper for stochastically ordered families will include cases where the $T_i(n)$ are sample means or sample medians; the interquartile range will be used for selecting the population with the lightest tail. It will be assumed throughout that the $\{T_i(n)\}$ satisfy (1.6).

REMARK 2.1. The stopping rules (2.6) being independent has certain drawbacks. However, improvement here must await results in the location case, for which there does not appear to be a nonparametric rule which eliminates obviously inferior populations and satisfies (1.8).

3. RESULTS AND APPLICATIONS

The structure of Section 2 (especially (2.4) and (2.5)) will be assumed and the goal will be to guarantee (1.8). At this point, no specific assumptions concerning the preference pattern (2.4) will be made; only later will such orders as stochastic orderings or tail orderings be used.

The major difficulty encountered in guaranteeing (1.8) is the lack of a recognizable least favorable configuration (see Bechhofer (1954)). Letting

(3.1) $$\underline{\theta} = (\theta_1,\ldots,\theta_K)$$

(3.2) $$\Omega^* = \{\underline{\theta} \mid \mu(\theta_i) \le \ldots \le \mu(\theta_K)\}$$

(3.3) $$P(\underline{\theta},\delta) = P_{\underline{\theta}}\left\{\delta^{-1}b\left[T_K\left(N_K(\delta)\right) - T_i\left(N_i(\delta)\right) - \mu(\theta_K) + \mu(\theta_i)\right] \ge -b\right\} ,$$

For $i = 1,\ldots,K - 1$

it is easy to see that

(3.4) $$\liminf P(CS) \ge \liminf P(\underline{\theta},\delta) = Q^* \quad \text{(say)},$$

where the latter inf is taken over Ω^* . Now, in the location parameter case, $N_i(\delta) \equiv N(\delta)$ $(i = 1,\ldots,K)$ and

(3.5) $$P(\underline{\theta},\delta) \equiv P(\underline{0},\delta) ,$$

so that

(3.6) $$\liminf P(CS) \ge \lim P(\underline{0},\delta) = P^* .$$

However, in the cases which are dealt with here, the parameter point at which $P(CS)$ attains its infinum is as yet unknown. However, there exists $(\underline{\theta}^n, \delta_n)$ such that

(3.7) $$P(\underline{\theta}^n, \delta_n) \to Q^* \text{ and } \delta_n \to 0 \text{ as } n \to \infty .$$

If there were a $\underline{\theta}$ such that $\underline{\theta}^n \to \underline{\theta}$, since

(3.8) $$P(\underline{\theta}, \delta_n) \to P^*$$

$$P(\underline{\theta}^n, \delta_n) \to Q^* ,$$

some sort of a continuity criterion for $P(\cdot,\delta)$ suggests itself. By the Helly-Bray Lemma, it is clear that a sufficient condition for $Q^* \geq P^*$ is

$$(3.9)\quad P\left\{\delta_n^{-1} b\left[T_i\left(N_i(\delta_n)\right) - \mu(\theta_i^n)\right) \leq z\right\} \to \Phi(z) \quad \text{as} \quad n \to \infty \quad (i = 1,\ldots,K) .$$

This fact, together with the continuity criterion, suggests Theorem 3.1. Since the sufficient condition (3.9) depends only on the one-dimensional distributions, the rest of the paper will use the generic terms $T(n)$, $g(n)$, $N(\delta)$, and $\mu(\theta)$.

This intuitive continuity criterion suggests that $F(x;\theta)$ should be close to $F(x;\theta_0)$ if $|\mu(\theta) - \mu(\theta_0)|$ is small, so that

(A1) $$\sigma(\theta) \to \sigma(\theta_0) \quad \text{as} \quad \mu(\theta) \to \mu(\theta_0) .$$

(A2) For each θ_0 and for all $\varepsilon > 0$, there is a d such that

$$P_\theta\{g_n > d \quad \text{for} \quad n = 1,2,\ldots\} > 1 - \varepsilon \quad \text{if} \quad |\mu(\theta) - \mu(\theta_0)| < \varepsilon .$$

The other continuity conditions needed are that (2.1), (2.2), and (2.3) hold uniformly for $\mu(\theta)$ close to $\mu(\theta_0)$; this is summed up in

<u>THEOREM 3.1</u> Suppose that for each θ_0, $\varepsilon > 0$, $\sigma > 0$, and $z \varepsilon R'$, $\exists J$, $\gamma > 0$, $c > 0$, such that if $n \geq J$ and $|\mu(\theta) - \mu(\theta_0)| \leq \gamma$, then

(A3) $P_\theta\{|g(n) - \sigma(\theta)| > \sigma(\theta)H$ for some $n \geq J\} \leq \varepsilon$, where

$$0 < H < \varepsilon \quad \text{is such that} \quad |(1 \pm H)^2 - 1| < c/4$$

(A4) $$\left|P_\theta\left\{\sigma(\theta)^{-1} n^{1/2}\left(T(n) - \mu(\theta)\right) \leq z \pm \sigma\right\} - \Phi(z \pm \sigma)\right| \leq \varepsilon$$

(A5) $P_\theta\{n^{1/2}|T(m) - T(n)| > \sigma$ for some $m \ni |m - n| < cn\} \le \varepsilon$

Then, if Ω^* is compact and (A1) and (A2) are true, (1.8) holds.

REMARK 3.1. Theorem 3.1 is nonparametric in the sense that it holds independently of the underlying class of distributions $\{F(x;\theta)\}$, as long as this class satisfies (A1) - (A5) and Ω^* is compact.

REMARK 3.2. Although the proof of Theorem 3.1 depends on the rules (2.6), the Theorem is a fixed-sample result and thus can be verified in individual cases. Ω^* compact is used to guarantee that a subsequence of (θ_n, δ_n) converges; this condition does not appear to be serious in practice, since most measuring devices are finite.

REMARK 3.3. It is quite easy to see that the conditions of Theorem 3.1 suffice for the goals (i) selecting the t largest $\mu(\theta_{[K-t+1]}),\ldots,\mu(\theta_{[K]})$; (ii) selecting the smallest $\mu(\theta_{[1]})$; (iii) as selecting a restricted subset (see Santner (1974)).

LEMMA 3.1. (Location Parameters) If $F(x;\theta) = F(x - \theta)$ and (2.1) - (2.3) hold, then (1.8) holds and Ω^* need not be compact.

LEMMA 3.2. (Location Parameters with Unequal Scales) Suppose $F(x; (\mu,\sigma)) = F\left(\frac{x-\mu}{\sigma}\right)$, where $\sigma\varepsilon K$ compact and $\mu\varepsilon R$. Then, if (2.1) - (2.3) hold, (1.8) holds.

The following Proposition 3.1 is adapted from Geertsema (1972) and Bahadur (1966) and follows in a similar manner.

PROPOSITION 3.1. Let $c > 0$ and $0 < \alpha < 1$, and suppose $F(x)$ has a bounded second derivative in a neighborhood of the α^{th} population quantile ξ, with $F'(\xi) = f(\xi) > 0$. Let $b_n = [n\alpha] - cn^{1/2}/2$ and $a_n = [n\alpha] + cn^{1/2}/2$, and define $X_{[an]}$, $X_{[bn]}$ to be the $[an]^{th}$ and $[bn]^{th}$ order statistics. Then

$$(3.10) \quad c^{-1}\big(n\alpha(1-\alpha)\big)^{1/2}|X_{[an]} - X_{[bn]}| \to \big(\alpha(1-\alpha)\big)^{1/2}/f(\xi) \quad \text{a.s.}\ .$$

It is clear that some sort of "closeness" condition must be placed on the underlying distributions in order that the continuity criteria (A1) - (A5) hold. One convenient Lipschitz type condition is

(3.11) For each θ_0, $\exists$ numbers $B_\theta \to 1$ and $C_\theta \to 1$ as $\mu(\theta) \to \mu(\theta_0)$ such that for all x,y,

$$C_\theta|x - y| \le \left|F^{-1}\big(F(x;\ \theta_0);\ \theta\big) - F^{-1}\big(F(y;\ \theta_0);\ \theta\big)\right| \le B_\theta|x - y|\ .$$

LEMMA 3.3. (Sample Quantiles in Stochastic Orderings) Suppose that the distributions $F(x;\theta)$ are stochastically ordered and let $\mu(\theta)$ be the α^{th} population quantile. Suppose the conditions of Proposition 3.1 hold. Define $T(n) = X_{[n\alpha]}$ and $g(n)$ as in (3.10). Then, if Ω^* is compact, and (3.11) holds for all x and y in some neighborhood of $\mu(\theta_0)$, (1.8) holds.

LEMMA 3.4. (Sample Means in Stochastic Orderings) Suppose the distributions $F(x;\theta)$ are stochastically ordered, that $T(n)$ is the sample mean, and $g^2(n)$ is the sample variance. Suppose that $\mu(\theta) = F^{-1}\left\{F\big(\mu(\theta_0)\ ;\ \theta_0\big)\ ;\ \theta\right\}$ (which is true if $F(\cdot;\theta)$ is symmetric), and that for all θ_0, (3.12) $$\int \left(x - \mu(\theta)\right)^2 dF(x;\theta)$$

is bounded in a neighborhood of θ_0. If (3.11) holds and Ω^* is compact, then (1.8) holds.

LEMMA 3.5. (Interquantile Ranges in Tail Orderings) Suppose that the $F(x;\theta)$ are tail ordered and it is desired to select the population with the lightest tail. If $0 < \alpha_1 < 1/2 < \alpha_2 < 1$, and $\mu_1(\theta)$, $\mu_2(\theta)$ are the α_1^{st} and α_2^{nd} quantiles of $F(x;\theta)$, set $T(n) = X_{[n\alpha_2]} - X_{[n\alpha_1]}$ and let $g(n)$ be the obvious estimate based on functions of the type (3.10). If Ω^* is compact and (3.11) holds for all x,y in some neighborhood of $\mu_1(\theta)$ (and $\mu_2(\theta)$), then (1.8) holds.

4. PROOFS

DEFINITION 4.1. Let

(4.1)
$$M(\delta,\theta) = [b\sigma(\theta)/\delta]^2$$
$$M_1(\delta,\theta) = [b\sigma(\theta)(1 - H)/\delta]^2$$
$$M_2(\delta,\theta) = [b\sigma(\theta)(1 + H)/\delta]^2 .$$

Recall that $M(\delta,\theta)/N(\delta) \to 1$ a.s. under $F(\cdot;\theta)$ as $\delta \to 0$.

PROOF OF THEOREM 3.1. It is sufficient to prove (3.9). Then, by using (A2) and (A3), one sees that $\exists$ δ_0, J, and γ such that if $\delta \le \delta_0$ and $|\mu(\theta) - \mu(\theta_0)| \le \gamma$

(4.2) $$P_\theta\{N(\delta) \le J\} \le \varepsilon .$$

Since $b\sigma(\theta)/\delta\big(M(\delta,\theta)\big)^{1/2} = \big(b\sigma(\theta)/\delta\big)/[b\sigma(\theta)/\delta]$, one can easily show that

$$(4.3) \qquad \left|P_\theta\left\{b\delta^{-1}\left[T\big(M(\delta,\theta)\big) - \mu(\theta)\right] \le z \pm \sigma\right\} - \Phi(z \pm \sigma)\right| \le \varepsilon ,$$

$$(4.4) \quad P_\theta\left\{b\delta^{-1}\left|T(m) - T\big(M(\delta,\theta)\big)\right| > \sigma \text{ for some } m \ni \left|\frac{m}{M(\delta,\theta)} - 1\right| < c\right\} \le \varepsilon .$$

Choosing δ_0 and γ small enough that

$$|M_i(\delta,\theta) - M(\delta,\theta)| < cM(\delta,\theta) \quad (i = 1,2) ,$$

from (A3) and (4.2),

$$(4.5) \qquad P_\theta\{|N(\delta) - M(\delta,\theta)| > cM(\delta,\theta)\} \le 2\varepsilon .$$

Now,

$$\begin{aligned}(4.6) \quad & P_\theta\left\{\delta^{-1}b\left[T\big(N(\delta)\big) - \mu(\theta)\right] \le z\right\} - P_\theta\left\{\delta^{-1}b\left[T\big(M(\delta,\theta)\big) - \mu(\theta)\right] \le z + \sigma\right\} \\ & \le 2\varepsilon + P_\theta\left\{b\delta^{-1}\left|T\big(N(\delta)\big) - T\big(M(\delta,\theta)\big)\right| \ge \sigma \text{ and } |N(\delta)/M(\delta,\theta) - 1| \le c\right\} \\ & \le 2\varepsilon + P_\theta\left\{b\delta^{-1}\left|T(m) - T\big(M(\delta,\theta)\big)\right| \ge \sigma \text{ for some } m \ni |m/M(\delta,\theta) - 1| \le c\right\} \\ & \le 3\varepsilon .\end{aligned}$$

Similarly,

$$(4.7)\quad P_\theta\left\{\delta^{-1}b\left[T\left(N(\delta)\right) - \mu(\theta)\right] \le z\right\} \ge P_\theta\left\{\delta^{-1}b\left[T\left(M(\delta,\theta)\right) - \mu(\theta)\right] \le z - \sigma\right\} - 3\varepsilon .$$

By choosing σ small, $\delta \le \delta_0$, and $|\mu(\theta) - \mu(\theta_0)| \le \gamma$, this means from (4.3) that

$$(4.8)\qquad \left|P_\theta\left\{\delta^{-1}b\left[T\left(N(\delta)\right) - \mu(\theta)\right] \le z\right\} - \Phi(z)\right| \le 4\varepsilon ,$$

and hence that (3.9) holds.

PROOF OF LEMMA 3.3. For convenience, assume that (3.11) holds; one can extend to Lemma 3.3 from Bahadur (1966). Since

$$(4.9)\qquad \sigma(\theta) = \sigma(\theta_0)\, \frac{d}{dx} F^{-1}\left(F(x;\ \theta_0)\ ;\ \theta\right)\ \Big|\ x = \mu(\theta_0) ,$$

(A1) and (A2) follow from (3.11). (A3) follows from the probability integral transformation, (3.11), and (A2). For (A4), assuming z positive, n large, and $|\mu(\theta) - \mu(\theta_0)|$ small, since $F^{-1}\left(F(x;\ \theta_0)\ ;\ \theta\right)$ is increasing,

$$(4.10)\qquad \Phi(C_\theta z) - \varepsilon \le P_\theta\left\{\sigma(\theta)^{-1}n^{1/2}\left(T(n) - \mu(\theta)\right) \le z\right\} \le \Phi(B_\theta z) + \varepsilon .$$

(A5) follows from the integral transformation, (3.11), and Proposition 3.1.

PROOF OF LEMMA 3.4. Assume that $|C_\theta - 1| \le |B_\theta - 1|$, and let $g^2(n;\theta)$ be the sample variance obtained from $F^{-1}\left(F(X_1;\ \theta_0)\ ;\ \theta\right),\ldots,F^{-1}\left(F(X_n;\ \theta_0)\ ;\ \theta\right)$. Then, from (3.11), one can show that

$$(4.11)\qquad |g^2(n;\theta) - g^2(n;\ \theta_0)| \le |B_\theta^2 - 1|G_n \quad \text{(say)},$$

where $G_n \to c > 0$ almost surely. With arbitrarily large probability under θ_0 , (4.12) $$|g(n;\theta) - g(n;\theta_0)| \le |B_\theta^2 - 1| G_n^* ,$$

where $G_n^* \to c^* > 0$ almost surely. (A1) follows from (4.11) and (A2) and (A3) follow from (4.12). The proof of (A4) is somewhat involved. It is possible to show that, if

(4.13) $$F^{-1}\left(F(X_i;\theta_0) ; \theta\right) - F^{-1}\left(F(\theta_0;\theta_0) ; \theta\right) = H_i(\theta)(X_i - \theta) ,$$

then (A4) will follow if it can be shown that $\forall\ \varepsilon > 0 , \beta > 0 , \exists\ N(\varepsilon) , \eta(\varepsilon)$ such that $n \ge N(\varepsilon)$ and $|\theta - \theta_0| < \eta(\varepsilon)$ imply

(4.14) $$P_{\theta_0}\left\{\sqrt{n}\left|\frac{1}{n}\sum_{i=1}^{n}\left(H_i(\theta) - 1\right)(X_i - \theta_0)\right| > \varepsilon\right\} < \beta .$$

It turns out that (4.14) is true if for any sequence $(n_1\theta_1)$, $(n_2, \theta_2), \ldots$ with $n_j \to \infty$ and $\theta_j \to \theta_0$, the distribution of

(4.15) $$n_j^{-1/2}\sum_{i=1}^{n_j}\left(H_i(\theta_j) - 1\right)(X_i - \theta_0)$$

converges to that of a random variable putting all mass at 0 . Since $|H_i(\theta) - 1| \le |B_\theta - 1|$, if $E|X_i - \theta_0|^2$ exists, one may use the degenerate convergence criterion (Loéve (1963), page 317) to achieve the result. The proof of (A5) requires an extension of Anscombe's (1952) proof that the sample mean satisfies (2.6), and uses (A4), (3.12), and Kolmogorov's Inequality.

PROOF OF LEMMA 3.5. Again, as in Lemma 3.3, assume (3.11). (A1) - (A3) follow in a manner similar to the proof of Lemma 3.4, although $g(n;\theta)$ is derived from Proposition 3.1 and the expression for $\sigma(\theta)$. The proof of (A4) involves (3.11), the monotonicity of $F^{-1}\left(F(x;\theta_0) ; \theta\right)$, and decomposing the probability space into the 4 sets where $X_{[n\alpha_i]} - \mu_i(\theta) > 0$ or < 0 for $i = 1,2$. (A5) follows from Lemma 3.3.

REMARK 4.1. Note that a stochastic ordering was not really used in Lemma 3.3 or 3.4 except to insure that (2.4) makes sense. One could also combine Lemma 3.1 with Lemmas 3.2 - 3.5 to extend the set over which the infinum in (1.8) is taken.

ACKNOWLEDGEMENT I would like to express my deepest appreciation to my advisor Shanti S. Gupta for his advice and support.

REFERENCES

[1] Anscombe, F.J. (1952). Large sample theory of sequential estimation. *Proc. Camb. Phil. Soc.* (48) 600-617.

[2] Bahadur, R.R. (1966). A note on Quantiles in large samples. *Ann. Math. Statist.* (37) 577-580.

[3] 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.

[4] Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* (36) 463-467.

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

[6] Geertsema, J.C. (1972). Nonparametric sequential procedures for selecting the best of k populations. *J. Am. Statist. Assoc.* (67) 614-616.

[7] Lehmann, E.L. (1959). Testing Statistical Hypotheses, John Wiley, New York.

[8] Loéve, M. (1963). Probability Theory, D. Van Nostrand, Princeton.

[9] Robbins, H., Sobel, M., and Starr, N. (1968). A sequential procedure for selecting the largest of k means. *Ann. Math. Statist.* (39) 88-92.

[10] Santner, T.J. (1974). A restricted subset selection approach to ranking and selection problems. To appear in *Ann. Math. Statist.*