The Annals of Statistics
1978, Vol. 6, No. 3. 658-672

AN ADAPTIVE SOLUTION TO RANKING AND SELECTION PROBLEMS[1]

BY Y. L. TONG

University of Nebraska, Lincoln

An adaptive approach is considered as an alternative to the classical indifference-zone formulation of the problems of ranking and selection. With a fixed γ^*, the proposed procedure calls for the termination of sampling when the estimated probability of correct selection exceeds γ^* for the first time. Asymptotic properties of this procedure are proved as $\gamma^* \to 1$, and Monte Carlo results show that the procedure is well behaved even for moderate γ^*. Since the stopping variables depend on the estimators of the ordered parameters, distributions of the estimators as functions of the parameters are carefully studied via majorization.

1. Introduction. Let $\mathscr{F} = \{F(x, \theta) : \theta \in \Theta\}$ be a family of distributions where the parameter space Θ is a subset of the real line. For $i = 1, \cdots, k$ let $\{X_{ij}\}_{j=1}^{\infty}$ be a sequence of i.i.d. random variables with distribution function $F(x, \theta_i) \in \mathscr{F}$. Denoting $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$ (a point in the product parameter space) and defining the ordered parameters $\theta_{[1]} \leqq \cdots \leqq \theta_{[k]}$, the problem is to select the population associated with the largest parameter $\theta_{[k]}$.

The indifference-zone formulation of the problem, as originally considered in a pioneering paper of Bechhofer [1], may be briefly described as follows: After taking the observations X_{ij}'s for $i = 1, \cdots, k$ and $j = 1, \cdots, n$, one chooses an appropriate statistic T, observes

$$t_i = t_i^{(n)} = T(X_{i1}, \cdots, X_{in}), \qquad i = 1, \cdots, k; \tag{1.1}$$

then applies the decision rule "always select the population corresponding to the maximum of $(t_1, \cdots, t_k)$." Now for a fixed $\delta^* > 0$, denote

$$\Omega = \Omega(\delta^*) = \{\boldsymbol{\theta} \mid \psi(\theta_{[k]}, \theta_{[k-1]}) \geqq \delta^*\}, \tag{1.2}$$

where ψ is the distance function considered by Bechhofer–Kiefer–Sobel ([2], page 37). Then the set of all $\boldsymbol{\theta}$ not in Ω is regarded as the indifference-zone and the best population (with parameter $\theta_{[k]}$) is considered to be sufficiently distinct from the rest if and only if $\boldsymbol{\theta} \in \Omega$. With a preassigned $\gamma^* \in (1/k, 1)$, the main problem under the indifference-zone formulation is to determine the sample size $n^* = n(k, \gamma^*, \delta^*)$ such that the probability of a correct selection (CS) satisfies

$$\inf_{\boldsymbol{\theta} \in \Omega} P_{\boldsymbol{\theta}}(\text{CS}) \geqq \gamma^*. \tag{1.3}$$

Received December 1976; revised June 1977.

[1] Research supported in part by the National Science Foundation under grant No. MCS76-05783.

AMS 1970 *subject classifications.* Primary 62F07; Secondary 62L12.

Key words and phrases. Sequential multiple decision rules, ranking and selection, adaptive methods, sequential estimation, majorization and spacings.

This classical approach certainly has its mathematical elegance and has played a predominant role in the area of ranking and selection problems. However, it appears that the solution of the problem under this approach is rather "artificial" because of the lack of connection between the value of δ^* and the true parameter $\boldsymbol{\theta}$. Since δ^* is to be specified by the experimenter and the probability of correct selection is guaranteed to be at least γ^* only when $\psi(\theta_{[k]}, \theta_{[k-1]}) \geqq \delta^*$, in reality the true distance is unknown and there is no knowledge regarding the *true* probability of correct selection in a given real-life situation. In view of this difficulty, it is natural to consider more realistic approaches to this problem as alternatives.

Suppose that, for a fixed n, the distribution of t_i defined in (1.1) is $G_n(y, \theta_i)$. If G_n belongs to a location parameter family of continuous distributions, then we can write $G_n(y, \theta_i) = G_n(y - \theta_i)$. Defining $p = k - 1$ and

$$\psi(\theta_{[k]}, \theta_{[k-1]}) = \theta_{[k]} - \theta_{[k-1]} = \Delta \,, \tag{1.4}$$

$$\delta_i = \theta_{[k]} - \theta_{[i]} \,, \qquad i = 1, \cdots, p \tag{1.5}$$

(with $\delta_p = \Delta$), the probability of correct selection is, for every fixed n,

$$\begin{aligned} P_\theta(\mathrm{CS}) &= P_\theta[\textstyle\bigcap_{i=1}^p \{t_{(i)} \leqq t_{(k)}\}] \\ &= \textstyle\int \prod_{i=1}^p G_n(y + \delta_i)\, dG_n(y) = \gamma_1^{(n)} \,, \quad \text{say}, \end{aligned} \tag{1.6}$$

where $t_{(i)}$ has a cdf $G_n(y - \theta_{[i]})$ $(i = 1, \cdots, k)$. Note that $P_\theta(\mathrm{CS})$ depends on $\boldsymbol{\theta}$ only through $(\delta_1, \cdots, \delta_p)$. (Unless specified otherwise, all integrals in this paper are from $-\infty$ to ∞.)

A natural approach to this problem is, for a fixed n, the estimation of the probability of correct selection by estimating the parameters $\delta_1, \cdots, \delta_p$. This was considered recently by Olkin–Sobel–Tong [8]. Another approach seems to be the development of new adaptive procedures which will yield a probability of approximately γ^* under the true configuration, hence (1.3) is "approximately" satisfied when sampling terminates, at least for large γ^*. Along this line a class of adaptive sequential procedures is considered in this paper. Those procedures are stated in Section 2. Since they involve the estimators of the ordered parameters, the distributions of the estimators as functions of the true parameters play an important role. This is studied in detail in Section 3, and new probability inequalities are obtained through majorization. The limiting behavior of the class of procedures, as $\gamma^* \to 1$ (which is the case of interest in most applications), is studied in Section 4. Special results are given in Section 5 for the normal family when the common variance is either known or unknown. Section 6 contains some Monte Carlo results concerning the average sample number and the observed probability of correct selection under different configurations. It can be seen that the procedures are well behaved even for moderate γ^* values.

Although this paper concerns only locations parameter families, similar results can be obtained through obvious modifications for scale parameter families.

2. A class of adaptive sequential procedures. Our basic approach here is to estimate the probability of correct selection sequentially by estimating the true parameters involved, and stop for the first time when the estimated probability is at least γ^*. Suppose that, after observing $\mathbf{X}_1, \cdots, \mathbf{X}_n$ ($\mathbf{X}_j = (X_{1j}, \cdots, X_{kj})$) one computes $t_i^{(n)}$ ($i = 1, \cdots, k$) and estimates the δ_i's by

$$\hat{\delta}_i = \hat{\delta}_i^{(n)} = t_{[k]}^{(n)} - t_{[i]}^{(n)}, \qquad i = 1, \cdots, p; \tag{2.1}$$

where $t_{[1]}^{(n)} \leqq \cdots \leqq t_{[k]}^{(n)}$ are the ordered t values. Following the discussions given in [8] for every n the probability of a correct selection may be estimated by

$$\hat{\gamma}_1^{(n)} = \int \prod_{i=1}^{p} G_n(y + \hat{\delta}_i)\, dG_n(y). \tag{2.2}$$

Then, for a predetermined γ^*, one may keep taking observations until n is "large enough." This is defined in the following

PROCEDURE R_1. (a) Observe the sequence $\{\mathbf{X}_j\}_1^\infty$, one vector at a time. Stop with $N_1 = n$ where

$$n = \text{the smallest integer satisfying} \quad \hat{\gamma}_1^{(n)} \geqq \gamma^*. \tag{2.3}$$

(b) When sampling terminates, select the population corresponding to $t_{[k]}^{(n)}$.

Under this procedure, the stopping rule is now directly affected by the true configuration. In addition to R_1, two auxiliary procedures are also defined below. Our main reasons for introducing them are to give bounds on N_1 and to provide comparisons. (For reasons to be seen in Section 6, they do not perform well comparing to R_1; hence, they are not actually recommended for adoption in practical applications.) Suppose that, after observing the estimators $\hat{\delta}_i$'s, one obtains

$$\hat{\Delta} = \hat{\Delta}_n = \hat{\delta}_p^{(n)}, \qquad \hat{\bar{\delta}} = \hat{\bar{\delta}}_n = \frac{1}{p} \sum_{i=1}^{p} \hat{\delta}_i^{(n)}; \tag{2.4}$$

and evaluate

$$\hat{\gamma}_2^{(n)} = \int G_n^p(y + \hat{\Delta})\, dG_n(y), \tag{2.5}$$

$$\hat{\gamma}_3^{(n)} = \int G_n^p(y + \hat{\bar{\delta}})\, dG_n(y). \tag{2.6}$$

In $\hat{\gamma}_2^{(n)}$ and $\hat{\gamma}_3^{(n)}$ the values of $\hat{\delta}_i$'s in (2.2) are replaced by $\hat{\Delta}$ and $\hat{\bar{\delta}}$, respectively. They are, of course, appropriate estimators of

$$\gamma_2^{(n)} = \int G_n^p(y + \Delta)\, dG_n(y), \tag{2.7}$$

$$\gamma_3^{(n)} = \int G_n^p(y + \bar{\delta})\, dG_n(y), \tag{2.8}$$

respectively with $\bar{\delta} = (1/p) \sum_1^p \delta_i$. Here $\gamma_2^{(n)}$ is the lower bound of the true probability of correct selection (for fixed n) under the "least favorable" configuration, and $\gamma_3^{(n)}$ is an upper bound of the true probability for log concave distributions (for proof, see Theorem 2.1). In the special case when the δ_i ($\hat{\delta}_i$) values are close, then certainly the $\gamma_l^{(n)}$ ($\hat{\gamma}_l^{(n)}$) values are also close ($l = 1, 2, 3$). The evaluations of $\hat{\gamma}_2^{(n)}$ and $\hat{\gamma}_3^{(n)}$ involve joint probabilities of p exchangeable events. Inequalities on this type of probabilities are available [12] and, in special cases, their exact values may be found from existing statistical tables. We now define the Procedures R_2 and R_3 accordingly:

PROCEDURE R_l $(l = 2, 3)$. Same as Procedure R_1 with $N_l = n$, except equation (2.3) is replaced by

$$n = \text{the smallest integer satisfying } \quad \hat{\gamma}_l^{(n)} \geqq \gamma^* . \tag{2.9}$$

Note that R_1, R_2 and R_3 are identical for $k = 2$. For $k > 2$, it is obvious that Procedure R_2 is conservative. However, it is not so intuitively clear that R_3 is least conservative. This is shown in the following

THEOREM 2.1. *If* $\log G_n(y)$ *is a concave function of* y *for every* n, *then*

$$\hat{\gamma}_2^{(n)} \leqq \hat{\gamma}_1^{(n)} \leqq \hat{\gamma}_3^{(n)} \quad \text{a.s.} \tag{2.10}$$

holds for every $\boldsymbol{\theta}$ *and every* n. *Hence*

$$N_3 \leqq N_1 \leqq N_2 \quad \text{a.s.} \tag{2.11}$$

holds for every $\boldsymbol{\theta}$.

PROOF. The first inequality in (2.10) is obvious (actually for this one alone it does not need the log concavity of $G_n(y)$). The second inequality in (2.10) follows from a theorem in [8]. (2.11) follows from (2.10).

REMARK. It should be pointed out that, among many other distributions, a normal distribution function is log concave. To see this, let Φ and ϕ denote the $N(0, 1)$ cdf and pdf, respectively, then

$$\frac{d^2}{dy^2} \ln \Phi(y) = \frac{-\phi(y)}{\Phi^2(y)} [y\Phi(y) + \phi(y)] < 0$$

holds for all y.

3. Distribution functions of $\hat{\Delta}_n$ and $\hat{\hat{\delta}}_n$ and their bounds. Because of the important roles $\hat{\Delta}$ and $\hat{\hat{\delta}}$ play in the procedures stated above and in many other problems in multiple decision theory concerning ordered parameters, it becomes desirable to study the properties of the distributions of $\hat{\Delta}$ and $\hat{\hat{\delta}}$. In this section we consider their distribution functions for fixed n, obtain some bounds on the distribution functions through majorization (for definition, see [5], page 45), and find their limiting distributions as $n \to \infty$. Some of the results obtained will be used in Section 5.

For fixed n and every $z > 0$ consider the event

$$[\hat{\Delta}_n > z] = \bigcup_{j=1}^{k} \bigcap_{i \neq j} \{t_i \leqq t_j - z\} ;$$

and it follows that

$$\begin{aligned} P_{\boldsymbol{\theta}}[\hat{\Delta}_n \leqq z] &= 1 - \textstyle\sum_{j=1}^{k} \int \prod_{i \neq j} G_n(y + \theta_j - \theta_i - z)\, dG_n(y) \\ &= 1 - \textstyle\int \prod_{i=1}^{p} G_n(y + \delta_i - z)\, dG_n(y) \\ &\quad - \textstyle\sum_{j=1}^{p} \int \prod_{j \neq i} G_n(y + \theta_{[j]} - \theta_{[i]} - z)\, dG_n(y) . \end{aligned} \tag{3.1}$$

Denoting $\bar{\theta} = (1/k) \sum_1^k \theta_i$ and rewriting $\sum_{i \neq j} (\theta_j - \theta_i) = k(\theta_j - \bar{\theta})$, $\sum_{i \neq j} (\theta_{[j]} - \theta_{[i]}) = k(\theta_{[j]} - \bar{\theta})$, we observe

THEOREM 3.1. *If* $\ln G_n(y)$ *is a concave function of* y, *then*

$$\begin{aligned} P_\theta[\hat{\Delta}_n \leqq z] &\geqq 1 - \textstyle\sum_{j=1}^k \int G_n^{\,p}(y + (k(\theta_j - \bar{\theta})/p) - z)\, dG_n(y) \\ &= 1 - \textstyle\int G_n^{\,p}(y + \bar{\delta} - z)\, dG_n(y) \\ &\quad - \textstyle\sum_{j=1}^{\prime p} \int G_n^{\,p}(y + (k(\theta_{[j]} - \bar{\theta})/p) - z)\, dG_n(y)\,, \end{aligned} \tag{3.2}$$

$$P_\theta[\hat{\Delta}_n \leqq z] \leqq 1 - \textstyle\int G_n^{\,p}(y + \Delta - z)\, dG_n(y) \tag{3.3}$$

hold for all $\boldsymbol{\theta}$ *and all* $z \geqq 0$.

PROOF. For fixed j if $\ln G_n(y)$ is a concave function of y, then $\prod_{i \neq j} G_n(y + (\theta_j - \theta_i) - z)$ is a Schur-concave function of $(\theta_j - \theta_1, \cdots, \theta_j - \theta_k)$. Therefore (from [6]) the first integral in (3.1) is a Schur-concave function for every j. (3.2) follows from the fact that the vector $(\theta_j - \theta_1, \cdots, \theta_j - \theta_k)$ (after the components are properly reordered) majorizes the vector $(k(\theta_j - \bar{\theta})/p, \cdots, k(\theta_j - \bar{\theta})/p)$ for all j. The proof of (3.3) is obvious.

REMARKS. (a) The distribution of $\hat{\Delta}_n$ has the location translation invariance property and it depends on the θ_i's only through $(\theta_i - \bar{\theta})$, or equivalently, the δ_i's.

(b) The lower and upper bounds on the cdf of $\hat{\Delta}_n$ given above again involve a joint probability of exchangeable events. In particular, it follows from (3.3) and [12] that

$$\begin{aligned} P_\theta[\hat{\Delta}_n \leqq z] &\leqq 1 - \textstyle\prod_{l=1}^r \int G_n^{\,a_l}(y + \Delta - z)\, dG_n(y) \\ &\leqq 1 - \{\textstyle\int G_n(y + \Delta - z)\, dG_n(y)\}^p \end{aligned} \tag{3.4}$$

holds for all nonnegative integers $a_1, \cdots, a_r$ satisfying $\sum_1^r a_l = p$. If G_n is a normal cdf, then the bounds given in (3.2), (3.3) and (3.4) can be computed from [4], and the last integral in (3.4) depends only on a univariate normal probability.

(c) It should be noted that the bound in (3.2) is attainable when $\theta_i = \theta_j$ $(\delta_i = 0)$ for i, j. Also, when the δ_i values are not too small and they are approximately the same, then from (3.1) we have

$$P_\theta[\hat{\Delta}_n \leqq z] \cong 1 - \textstyle\int G_n^{\,p}(y + \bar{\delta} - z)\, dG_n(y)\,. \tag{3.5}$$

The distribution of $\hat{\bar{\delta}}_n$ as a function of $\boldsymbol{\theta}$ is more complicated. We shall not attempt to give its explicit functional form. In the following we give an inequality instead and show how the distribution of $\hat{\bar{\delta}}_n$ depends on $\boldsymbol{\theta}$. For fixed n let $g_n(y - \theta)$ denote the corresponding pdf with a cdf $G_n(y - \theta)$, and define $S = \{\mathbf{t} = (t_1, \cdots, t_k) \mid \hat{\bar{\delta}} \leqq z\}$ (a subset in R_k). Then clearly for every $z > 0$ we have

$$P_\theta[\hat{\bar{\delta}}_n \leqq z] = \int \underset{S}{\cdots} \int \textstyle\prod_{i=1}^k g_n(y_i - \theta_i)\, dy_1 \cdots dy_k\,. \tag{3.6}$$

THEOREM 3.2. *If* $\ln g_n(y)$ *is a concave function of* y, *then* $P_\theta[\hat{\hat{\delta}}_n \leqq z]$ *is a Schur-concave function of* $\boldsymbol{\theta}$.

PROOF. Using the identity $p\hat{\hat{\delta}}_n = k(t_{[k]} - \bar{t})$ with $\bar{t} = (1/k) \sum_1^k t_i$, it can be shown that S is a convex set and the proof follows immediately from a theorem in [7].

It might be tempting to claim that the distribution of $\hat{\Delta}_n$ given in (3.1) is also a Schur function of $\boldsymbol{\theta}$ for $k > 2$ ($p > 1$) under suitable conditions. The following example shows that this is not the case.

EXAMPLE. For $k = 3$ let t_1, t_2, t_3 be independent normal variables with means θ_i and a common variance τ^2. Consider three configurations $\boldsymbol{\theta}(1) = (c, c - \varepsilon, -2c + \varepsilon)$, $\boldsymbol{\theta}(2) = (c/2, 0, -c/2)$, $\boldsymbol{\theta}(3) = (\varepsilon, 0, -\varepsilon)$ where $c > 4\varepsilon > 0$. Clearly $\boldsymbol{\theta}(1) \succ \boldsymbol{\theta}(2) \succ \boldsymbol{\theta}(3)$ holds and for all j the probability $P_{\theta(j)}[\bigcap_{i=1}^3 \{|t_i - \theta_i(j)| < \varepsilon/2\}]$ can be made sufficiently large for small τ. Hence when τ is small enough the inequalities

$$P_{\theta(1)}[\hat{\Delta} \leqq z] \geqq P_{\theta(2)}[\hat{\Delta} \leqq z], \qquad P_{\theta(3)}[\hat{\Delta} \leqq z] \geqq P_{\theta(2)}[\hat{\Delta} \leqq z]$$

hold for small z.

The limiting distributions of $\hat{\Delta}_n$ and $\hat{\hat{\delta}}_n$ are given in the following theorem. The proof of the theorem follows from Theorem 3 of [8] and the fact that linear combinations of asymptotically normal variables are asymptotically normal variables.

THEOREM 3.3. *Assume that, as* $n \to \infty$, *there exists a* $\sigma > 0$ *such that* $n^{\frac{1}{2}}(t^{(n)} - \theta)/\sigma$ *converges to an* $N(0, 1)$ *variable in distribution. Then, for every* $\boldsymbol{\theta}$ *such that* $\theta_i \neq \theta_j$ $(i \neq j)$ *both* $n^{\frac{1}{2}}(\hat{\Delta}_n - \Delta)/2^{\frac{1}{2}}\sigma$ *and* $(n(k-1))^{\frac{1}{2}}(\hat{\hat{\delta}}_n - \bar{\delta})/k^{\frac{1}{2}}\sigma$ *converge to a* $N(0, 1)$ *variable in distribution as* $n \to \infty$.

4. Asymptotic behavior of the procedures. In this section we obtain results on consistency properties of the procedures for both the sample number and the probability of correct selection as $\gamma^* \to 1$.

For $n = 1, 2, \cdots$ consider the function

$$h_n(u) = \int G_n^{\,p}(y + u)\, dG_n(y) \tag{4.1}$$

for $u > 0$, and assume that G_n (hence h_n) possesses the following properties:

A1. For given γ, there exists a unique $u_n = u_n(\gamma)$ satisfying $h_n(u_n) = \gamma$ such that $u_n(\gamma) \uparrow \infty$ as $\gamma \uparrow 1$ for fixed n and $u_n(\gamma) \downarrow 0$ as $n \uparrow \infty$ for fixed γ.

A2. For given $\varepsilon > 0$, there exists an $\varepsilon' > 0$ and γ' such that

$$|u_m(\gamma)/u_n(\gamma) - 1| < \varepsilon' \quad \text{implies} \quad |m/n - 1| < \varepsilon \qquad \text{for all} \quad \gamma > \gamma'.$$

Note that A1 is easily satisfied when $h_n(u)$ is <1 for all u, continuous and strictly increasing in u for every n and $h_n(u) \uparrow 1$ as $n \uparrow \infty$ for every fixed $u > 0$. Now for given γ^* and $\boldsymbol{\theta}$ let u_n satisfy $h_n(u_n) = \gamma^*$, and let n_2, n_3 satisfy

$$n_2 = \inf\{n \geqq 1 \mid u_n \leqq \Delta\}, \qquad n_3 = \inf\{n \geqq 1 \mid u_n \leqq \bar{\delta}\}. \tag{4.2}$$

Clearly n_l is the smallest integer satisfying $\gamma_l^{(n)} \geqq \gamma^*$ ($l = 2, 3$). Under the procedures R_1, R_2 and R_3 we now study the behavior of N_1, N_2 and N_3.

THEOREM 4.1. *Assume that the conditions* A1, A2 *are satisfied and that $t^{(n)}$ is a strongly consistent estimator of θ. Then, for all $\boldsymbol{\theta}$ with* $\Delta > 0$,

(a) $P_{\boldsymbol{\theta}}[N_l < \infty] = 1$ *holds for* $l = 1, 2, 3$ *and for all* $\gamma^* < 1$;
(b) $N_l \to \infty$ a.s. *holds for* $l = 1, 2, 3$ *as* $\gamma^* \to 1$;
(c) $N_l/n_l \to 1$ a.s. *for* $l = 2, 3$ *as* $\gamma^* \to 1$.

PROOF. For fixed $\gamma^* < 1$, let u_n satisfy $h_n(u_n) = \gamma^*$. Clearly we have $\hat{\Delta}_n \to \Delta$ a.s. Hence for ω not in a null set there exists an $M_1(\omega)$ such that $\hat{\Delta}_n > \Delta/2$ holds for all $n > M_1$. Also, there exists an $M_2(\gamma^*)$ such that $u_n(\gamma^*) < \Delta/2$ holds for all $n > M_2$. Let $M = \max(M_1, M_2)$. Then, by the monotonicity of h_n, $N_2 = N_2(\omega, \gamma^*) < M$ holds and the proof of (a) follows from (2.11). Now for fixed $M > 0$ let $\gamma(n) < 1$ be large enough so that $\hat{\hat{\delta}}_n(\omega) < u_n$ holds ($n = 1, \cdots, M$), where u_n satisfies $h_n(u_n) = \gamma(n)$. Let $\gamma_0 = \max_{1 \leqq n \leqq M} \gamma(n)$. Then $N_3 = N_3(M, \omega) > M$ holds for every $\gamma^* > \gamma_0$ and the proof of (b) again follows from (2.11).

To prove (c), we first note that, from (4.2), (2.9) and the condition (A1), for every $\gamma^* < 1$

$$(u_{n_2}/\Delta) \leqq 1 < (u_{n_2-1}/\Delta), \qquad (4.3)$$
$$(u_{N_2}/\hat{\Delta}_{N_2}) \leqq 1 < (u_{N_2-1}/\hat{\Delta}_{N_2-1}) \quad \text{a.s.}$$

hold simultaneously. For arbitrary but fixed $\varepsilon > 0$ let ε' and γ' satisfy condition (A2). For ω not in a null set let $M_1(\omega, \varepsilon)$ be such that $|\Delta/\hat{\Delta}_{n_2} - 1| < \varepsilon'$ holds for all $n_2 > M_1$, and let γ'' be given by

$$\gamma'' = \inf\{\gamma \mid N_2(\omega, \gamma) > M_1\}.$$

Then for every $\gamma^* > \gamma_0 = \max(\gamma', \gamma'')$, we have, from (4.3),

$$1 - \varepsilon' < (\Delta/\hat{\Delta}_{N_2}) \leqq (u_{n_2-1}(\gamma^*)/u_{N_2}(\gamma^*)),$$
$$(u_{n_2}(\gamma^*)/u_{N_2-1}(\gamma^*)) < (\Delta/\hat{\Delta}_{N_2-1}) < 1 + \varepsilon';$$

which implies, from condition (A2),

$$1 - \varepsilon < (n_2 - 1)/N_2, \qquad n_2/(N_2 - 1) < 1 + \varepsilon.$$

The proof of the a.s. convergence of N_3/n_3 is similar.

For given $\boldsymbol{\theta}$ and γ^* let n_1 be the smallest integer satisfying $\gamma_1^{(n)} \geqq \gamma^*$ where $\gamma_1^{(n)}$ is defined in (1.6). To establish the a.s. convergence of N_1/n_1 under the procedure R_1, a stronger condition on G_n is needed; this is stated below.

A3. For every $n = 1, 2, \cdots$ and for every $\boldsymbol{\theta}$, G_n satisfies

$$\int \prod_{i=1}^{p} G_n(y + \delta_i)\, dG(y) = \int \prod_{i=1}^{p} H(y + u_n\delta_i)\, dH(y) = h^*(u_n), \quad \text{say}, \qquad (4.4)$$

for some cdf H and some $u_n > 0$; where (a) H is continuous, strictly increasing and $H(y) < 1$ for all y, (b) u_n tends to ∞ as $n \to \infty$, and (c) for fixed $\varepsilon > 0$, there exists an $\varepsilon' > 0$ and an $M(\varepsilon, \varepsilon') > 0$ such that $|u_m/u_n - 1| < \varepsilon'$ implies $|m/n - 1| < \varepsilon$ for all $m, n > M$.

THEOREM 4.2. *Assume that condition* A3 *is satisfied and that* $t^{(n)}$ *is a strongly consistent estimator of* θ. *Then, for all* $\boldsymbol{\theta}$ *with* $\Delta > 0$, $N_1/n_1 \to 1$ a.s. *as* $\gamma^* \to 1$.

PROOF. First note that, for fixed γ^* let u^* satisfy $h^*(u^*) = \gamma^*$, then n_1 satisfies

$$(u_{n_1-1}/u^*) < 1 \leqq (u_{n_1}/u^*) . \tag{4.5}$$

Since $t^{(n)} \to \theta$ a.s., for ω not in a null set there exists an $M_1(\omega)$ such that $|\hat{\delta}_i(\omega)/\delta_i - 1| < \varepsilon'$ holds for all $i = 1, \cdots, p$ and all $n > M_1$. Let γ_0 be large enough so that $N_1(\omega, \gamma_0) > M_0 = \max(M_1, M)$, where M satisfies (c) in condition A3. Then for every $\gamma^* > \gamma_0$ we have

$$\begin{aligned}\int \textstyle\prod_{i=1}^{p} H(y + (1 - \varepsilon')u_{N_1}\delta_i)\, dH(y) &\leqq \int \textstyle\prod_{i=1}^{p} H(y + u_{N_1}\hat{\delta}_i^{(N_1)}(\omega))\, dH(y) \\ &\leqq \int \textstyle\prod_{i=1}^{p} H(y + (1 + \varepsilon')u_{N_1}\delta_i)\, dH(y) .\end{aligned}$$

This along with

$$\begin{aligned}&\int \textstyle\prod_{i=1}^{p} H(y + u_{N_1-1}\hat{\delta}_i^{(N_1-1)}(\omega))\, dH(y) \\ &\qquad < \gamma^* \leqq \int \textstyle\prod_{i=1}^{p} H(y + u_{N_1}\hat{\delta}_i^{(N_1)}(\omega))\, dH(y) ,\end{aligned}$$

(4.5) and condition A3 implies

$$(1 - \varepsilon')u_{N_1-1} < u^* \leqq u_{n_1} , \qquad u_{n_1-1} < u^* \leqq (1 + \varepsilon')u_{N_1} .$$

The proof now follows from (c) in condition A3.

REMARK. We note that n_1 would be the sample size required to guarantee a probability of correct selection γ^* under a fixed-sample procedure if the values of $\delta_1, \delta_2, \cdots, \delta_p$ were known. Theorem 4.2 asserts that, under the Procedure R_1, one can do approximately equally well in terms of the random sample number without information on the δ_i's when γ^* is fairly large.

In the following theorem we establish the relationship between γ^* and the true probability of correct selection under the Procedures R_1, R_2 and R_3.

THEOREM 4.3. *Assume that the conditions imposed in Theorem* 4.1 *are satisfied and that* $\boldsymbol{\theta}$ *satisfies* $\Delta > 0$. *Then, for arbitrary but fixed* $\varepsilon > 0$, *there exists an* $\gamma_0^{(l)} = \gamma_0^{(l)}(\varepsilon, k, \boldsymbol{\theta})$ *such that*

$$|P_{\boldsymbol{\theta}}(\mathrm{CS}) - \gamma^*| < \varepsilon \tag{4.6}$$

holds for all $\gamma^* > \gamma_0^{(l)}$ *under the Procedure* R_l ($l = 1, 2, 3$). *In particular, we have*

$$P_{\boldsymbol{\theta}}(\mathrm{CS}) \to 1 \qquad as \quad \gamma^* \to 1 \tag{4.7}$$

under R_l ($l = 1, 2, 3$).

PROOF. Let $\varepsilon > 0$ be arbitrary but fixed. Since $t^{(n)} \to \theta$ a.s., there exists an $M(\varepsilon)$ such that $P[|t^{(n)} - \theta| > \Delta/2] \leqq \varepsilon/3p$ holds for all $n > M$. For $l = 1, 2, 3$ let $\gamma_0^{(l)}$ be large enough so that $P_{\boldsymbol{\theta}}[N_l > M] \geqq 1 - \varepsilon/3$. Without loss of generality we may assume $\theta_k = \theta_{[k]}$. Then for every $\gamma^* > \gamma_0^{(l)}$ we have

$$\begin{aligned}P_{\boldsymbol{\theta}}(\mathrm{CS}) &\geqq P_{\boldsymbol{\theta}}[\textstyle\bigcap_{i=1}^{p} \{t_i^{(N_l)} \leqq t_k^{(N_l)}\} \,|\, N_l > M] \cdot P_{\boldsymbol{\theta}}[N_l > M] - \varepsilon/3 \\ &\geqq (1 - \varepsilon/3)^2 - \varepsilon/3 > 1 - \varepsilon .\end{aligned}$$

This establishes (4.7). Now for $l = 1, 2, 3$ if $\gamma_0^{(l)} \geqq 1 - \varepsilon$ is large enough so that $P_\theta(\text{CS}) \geqq 1 - \varepsilon$ holds for all $\gamma^* \geqq \gamma_0^{(l)}$ under the Procedure R_l, then (4.6) holds. This completes the proof of the theorem.

REMARK. Theorem 4.3 implies that, under the Procedure R_l,

$$P_\theta(\text{CS}) \geqq \gamma^* - \varepsilon \tag{4.8}$$

holds for all $\gamma^* > \gamma_0^{(l)}$ $(l = 1, 2, 3)$. The value of $\gamma_0^{(l)}$ certainly depends on the probability distributions through $\boldsymbol{\theta}$ among other things. For fixed $\varepsilon > 0$ if (4.8) holds for a $\gamma_0^{(l)}$ which is not very close to 1, then even for moderate γ^* the Procedure R_l is on the safe side and it becomes practically useful provided the average sample number is not too large. The Monte Carlo results for the normal family given in Section 6 indicate that the Procedure R_1, the main procedure considered in this paper, performs remarkably well when γ^* is a small as 0.75. Details will be given in Section 6.

5. Normal family. In this section we consider the special case when the sequences of random variables $\{X_{ij}\}_{j=1}^{\infty}$ specified in Section 1 are i.i.d. normal variables with means θ_i $(i = 1, \cdots, k)$ and a common known variance σ^2. For every n let t_i (as defined in (1.1)) be $\bar{X}_i^{(n)} = (1/n) \sum_{j=1}^{n} X_{ij}$. Then $G_n(y) = \Phi(n^{\frac{1}{2}}y/\sigma)$ holds for all n, where Φ is the $N(0, 1)$ cdf. We now have

$$\gamma_1^{(n)} = \int \prod_{i=1}^{p} \Phi(y + n^{\frac{1}{2}}\delta_i/\sigma)\, d\Phi(y)\,, \tag{5.1}$$

$$\gamma_2^{(n)} = \int \Phi^p(y + n^{\frac{1}{2}}\Delta/\sigma)\, d\Phi(y)\,, \qquad \gamma_3^{(n)} = \int \Phi^p(y + n^{\frac{1}{2}}\bar{\delta}/\sigma)\, d\Phi(y)\,. \tag{5.2}$$

For every $\gamma^* \in (1/k, 1)$ let b satisfy

$$\int \Phi^p(y + b)\, d\Phi(y) = \gamma^* \tag{5.3}$$

(the values of b can be obtained from [4]). Then n_1 is the smallest integer satisfying

$$\int \prod_{i=1}^{p} \Phi(y + n^{\frac{1}{2}}\delta_i/\sigma)\, d\Phi(y) \geqq \gamma^*\,, \tag{5.4}$$

and n_2, n_3 are the smallest integers satisfying

$$n_2 \geqq b^2\sigma^2/\Delta^2\,, \qquad n_3 \geqq b^2\sigma^2/\bar{\delta}^2\,, \tag{5.5}$$

respectively. It is clear that the conditions (A1), (A2) and (A3) are satisfied. Hence Theorems 4.1, 4.2 and 4.3 apply.

For b satisfying (5.3), Equation (2.9) reduces to

$$N_2 = \text{the smallest integer } n \text{ such that } n \geqq b^2\sigma^2/\hat{\Delta}_n^2\,, \tag{5.6}$$

$$N_3 = \text{the smallest integer } n \text{ such that } n \geqq b^2\sigma^2/\hat{\bar{\delta}}_n^2\,. \tag{5.7}$$

Theorem 4.1 implies that $N_2/(b^2\sigma^2/\Delta^2) \to 1$ a.s. and $N_3/(b^2\sigma^2/\bar{\delta}^2)$ a.s. as $\gamma^* \to 1$. To obtain a result regarding the average sample number we modify the procedures slightly. Suppose that, with a sequence of positive real numbers $\{d_n\}$ satisfying $d_n \to 0$ as $n \to \infty$, we consider the stopping variables

(5.8) N_1' = the smallest integer n satisfying

$$\int \prod_{i=1}^{p} \Phi(y + n^{\frac{1}{2}}(\hat{\delta}_i^{(n)} + d_n)/\sigma)\, d\Phi(y) \geqq \gamma^* ,$$

(5.9) N_2' = the smallest integer n satisfying $n \geqq b^2\sigma^2/(\hat{\Delta}_n + d_n)^2$,

(5.10) N_3' = the smallest integer n satisfying $n \geqq b^2\sigma^2/(\hat{\hat{\delta}}_n + d_n)^2$.

Clearly $N_l' \leqq N_l$ a.s. holds and Theorems 4.1—4.3 apply when N_l is replaced by N_l' ($l = 1, 2, 3$). We now investigate the behavior of N_1', N_2' and N_3'. By adding d_n, $(\hat{\delta}_i^{(n)} + d_n)$, $(\hat{\Delta}_n + d_n)$ and $(\hat{\hat{\delta}}_n + d_n)$ are "bounded away" from zero, thus we are able to prove a result concerning the expected values of N_l' ($l = 1, 2, 3$). Note that in the following theorem the limiting behavior of EN_l' does not depend on a given choice of $\{d_n\}$. In applications one may choose $d_n = c/n^m$ for some $c > 0$ and $m > 0$, and d_n may be made to tend to zero as rapidly as one wishes by choosing m large enough.

THEOREM 5.1. *If there exists an* $m > 0$ *so that* $\liminf_{n\to\infty} n^m d_n > 0$, *then for every* $\boldsymbol{\theta}$ *with* $\Delta > 0$ *and for every* $\sigma^2 < \infty$

(5.11) $$N_l'/n_l \to 1 \quad \text{a.s.}, \qquad EN_l'/n_l \to 1$$

hold as $\gamma^* \to 1$ *for* $l = 1, 2, 3$.

PROOF. The a.s. convergence follows immediately from Lemma 1 of [3] and the fact that $d_n \to 0$. To prove $EN_2'/n_2 \to 1$, by Lemma 2 of [3] it suffices to show that $E_{\boldsymbol{\theta}} \sup_n \Delta^2/(\hat{\Delta}_n + d_n)^2 < \infty$. Applying (3.3), it follows that, for every n and every $v \in (1, \Delta^2/d_n^2)$,

$$\begin{aligned} P_{\boldsymbol{\theta}}&[\Delta^2/(\hat{\Delta}_n + d_n)^2 \geqq v] \\ &< \int [1 - \{1 - \Phi(-y - n^{\frac{1}{2}}((1 - v^{-\frac{1}{2}})\Delta + d_n)/\sigma)\}^p]\, d\Phi(y) \\ &< 2^p\Phi(-n^{\frac{1}{2}}((1 - v^{-\frac{1}{2}})\Delta + d_n)/2^{\frac{1}{2}}\sigma) \\ &< C_n \cdot \exp[-n((1 - v^{-\frac{1}{2}})\Delta + d_n)^2/4\sigma^2] = C_n \cdot r(n, v) , \quad \text{say,} \end{aligned}$$

where $C_n < C'$ holds for some finite C' for all n (the probability is zero for $v \geqq \Delta^2/d_n^2$). Therefore for every $v > 1$ we have

$$P_{\boldsymbol{\theta}}[\sup_n \Delta^2/(\hat{\Delta}_n + d_n)^2 \geqq v] = P_{\boldsymbol{\theta}}[\textstyle\bigcup_{n=1}^{\infty} \Delta^2/(\hat{\Delta}_n + d_n)^2 \geqq v] \leqq \sum_{n=1}^{\infty} s(n, v) ,$$

where

(5.12) $$\begin{aligned} s(n, v) &= C'r(n, v) \quad &&\text{if} \quad v < \Delta^2/d_n^2 , \\ &= 0 &&\text{otherwise.} \end{aligned}$$

Now consider $\sum_{j=2}^{\infty} \sum_{n=1}^{\infty} s(n, j)$. Since by the condition on $\{d_n\}$ there exists a $Q > 0$ such that $d_n \geqq Q \cdot n^{-m}$ holds for all n, it follows that, from (5.12), (for every fixed j) $s(n, j) = 0$ holds for all $n < (Q^2j/\Delta^2)^{1/2m} = D(j)$. Therefore

$$\begin{aligned} \sum_{j=2}^{\infty} \sum_{n=1}^{\infty} s(n, j) &\leqq C' \sum_{j=2}^{\infty} \sum_{n=D(j)+1}^{\infty} r(n, j) \\ &\leqq C' \sum_{j=2}^{\infty} \int_{D(j)}^{\infty} r(w, j)\, dw \\ &\leqq C \int_0^{\infty} \exp\{-\Delta(1 - 2^{-\frac{1}{2}})D(u)\}\, du < \infty \end{aligned}$$

holds for some finite C. This shows

$$E \sup_n \Delta^2/(\hat{\Delta}_n + d_n)^2 \leqq 5 + \sum_{j=2}^{\infty} P_\theta[\Delta^2/(\hat{\Delta}_n + d_n)^2 > j] < \infty$$

and $EN_2'/(b^2\sigma^2/\Delta^2) \to 1$ as $\gamma^* \to 1$. The rest of the proof follows immediately from the fact that $n_1 \leqq n_2$, $n_3 \leqq n_2$, $N_1' \leqq N_2'$, $N_3' \leqq N_2'$ a.s. and the dominated convergence theorem.

It should be noted that the stopping variables N_2 and N_2' are quite similar to that considered by Robbins, Sobel and Starr [9], except that in their solution σ^2 is estimated sequentially under the indifference-zone formulation, while the present solution is adaptive in which Δ is estimated sequentially and σ^2 is assumed to be known. If σ^2 is also unknown, then under the present approach the stopping variables may depend on both estimators of σ and Δ. Hence after observing $\{X_{ij}\}_{j=1}^n$ for $i = 1, \cdots, k$ $n \geqq 2$, one computes

$$(5.13) \qquad \bar{X}_i^{(n)} = \frac{1}{n} \sum_{j=1}^n X_{ij}\,, \qquad S_n^{\,2} = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i^{(n)})^2\,.$$

Define, for $l = 1, 2, 3$, N_l'' to be the stopping variable such that σ^2 is replaced by $S_n^{\,2}$ in (5.8), (5.9) and (5.10) respectively. Note that Theorems 4.1—4.3 also apply when N_l is replaced by N_l'' $(l = 1, 2, 3)$. In the following theorem we show that N_l'' and N_l' $(l = 1, 2, 3)$ possess a similar asymptotic behavior.

THEOREM 5.2. *Under the same conditions imposed in Theorem* 5.1

$$(5.14) \qquad N_l''/n_l \to 1 \quad \text{a.s.}, \qquad EN_l''/n_l \to 1$$

hold for all $\sigma^2 < \infty$ *and for* $l = 1, 2, 3$ *as* $\gamma^* \to 1$.

PROOF. The a.s. convergence follows from Lemma 1 of [3], $\hat{\delta}_i \to \delta_i$ $(i = 1, \cdots, p)$ a.s. and $S_n^{\,2} \to \sigma^2$ a.s. To show $EN_2''/n_2'' \to 1$, since

$$\begin{aligned} W &= \sup_n \Delta^2 S_n^{\,2}/\sigma^2(\hat{\Delta}_n + d_n)^2 \\ &\leqq \sup_n \Delta^2/(\hat{\Delta}_n + d_n)^2 \cdot \sup_n S_n^{\,2}/\sigma^2 = W_1 \cdot W_2\,, \quad \text{say,} \end{aligned}$$

a.s. and since the sequence of sample mean vectors and the sequence $\{S_n^{\,2}\}$ are independent, it follows that

$$E_{(\theta,\sigma^2)} W \leqq E_{(\theta,\sigma^2)} W_1 \cdot E_{(\theta,\sigma^2)} W_2$$

holds. Then from the proof of Theorem 5.1 it suffices to show that $E_{(\theta,\sigma^2)} \sup_n S_n^{\,2}/\sigma^2$ is finite, and this was already done in the proof of Theorem 3.3 of [11]. The rest of the proof follows from Lemma 2 of [3] and the dominated convergence theorem.

6. Monte Carlo results and concluding remarks. Two Monte Carlo studies were carried out on an IBM 360/65 computer at the University of Nebraska Computing Center. The studies concern two mean vectors $\boldsymbol{\theta}(1)$, $\boldsymbol{\theta}(2)$ of k normal populations for $k = 2, 4, 6, 8$ and 10 and for $\gamma^* = 0.75, 0.90, 0.95$ and 0.99. The two mean vectors given by

$$(6.1) \qquad \boldsymbol{\theta}(1) = (0.0, \cdots, 0.0, 0.8)\,, \qquad \boldsymbol{\theta}(2) = (0.0, 0.4, 0.8, 1.2, \cdots)$$

reflect two extreme cases. $\boldsymbol{\theta}(1)$ represents a slippage configuration under which all except one of the means are equal, while $\boldsymbol{\theta}(2)$ represents a configuration when the means are widely spread and equally spaced with an increment of 0.4. The common variance σ^2 was chosen to be 1. The formulas given in Table 5 of [10] were used to evaluate the probability integrals needed under the stopping variable N_1, and the table values in [4] were used for the stopping variables N_2 and N_3. The inital sample size was taken to be 2 in order to give protection against the undesirable situation of unusually early stopping with one observation. For each combination 400 cases were observed under each of the stopping variables, and the average sample numbers, their variances and the probabilities of correct selection were observed. In most cases the standard error of the average sample number was less than 5% of the average sample number. The values of n_1, n_2, n_3 under all stopping variables when $\boldsymbol{\theta} = \boldsymbol{\theta}(1)$ or under stopping variable N_2 or N_3 when $\boldsymbol{\theta} = \boldsymbol{\theta}(2)$ were obtained from (5.5) and the table in [4]. To find n_1 under N_1 the table in [8] was used through linear interpolations because $\boldsymbol{\theta}(2)$ is equally spaced.

The first study concerns the comparisons of N_l, N_l' and N_l'' ($l = 1, 2, 3$)

TABLE 1

Monte Carlo result—average sample number and observed probability of correct selection for $k = 4$

Stopping variable	$\boldsymbol{\theta}(1)$				$\boldsymbol{\theta}(2)$			
	$\gamma^* = 0.75$	0.90	0.95	0.99	$\gamma^* = 0.75$	0.90	0.95	0.99
	Average sample number							
n_1	5.00	10.00	14.00	23.00	8.00	21.00	34.00	68.00
N_1	6.26	12.87	18.47	30.09	7.34	21.26	34.99	62.65
N_1'	5.83	12.34	17.84	29.57	6.85	19.87	32.42	60.24
N_1''	7.96	15.15	20.50	31.45	9.38	23.60	36.58	63.32
n_2	5.00	10.00	14.00	23.00	18.00	38.00	54.00	91.00
N_2	9.86	17.22	23.22	35.56	14.45	35.36	47.91	80.55
N_2'	9.20	16.65	22.58	34.62	13.23	34.15	47.13	78.74
N_2''	11.64	19.41	25.03	36.69	17.22	37.54	50.49	81.10
n_3	5.00	10.00	14.00	23.00	5.00	10.00	14.00	23.00
N_3	4.10	8.98	13.57	24.12	4.60	9.97	14.25	23.97
N_3'	3.86	8.62	12.95	23.68	4.33	9.60	13.79	23.51
N_3''	5.99	11.64	16.28	25.62	6.27	11.90	16.13	25.45
	Observed probability of correct selection							
N_1	.7075	.8950	.9525	1.0000	.7850	.9450	.9875	1.0000
N_1'	.6750	.8850	.9500	1.0000	.7700	.9375	.9875	.9975
N_1''	.7800	.9375	.9750	.9900	.8550	.9500	.9550	.9975
N_2	.8350	.9400	.9875	1.0000	.8750	.9800	.9850	1.0000
N_2'	.8150	.9325	.9850	1.0000	.8650	.9775	.9975	1.0000
N_2''	.8750	.9675	.9875	.9925	.9225	.9775	.9875	1.0000
N_3	.5250	.7850	.8850	.9850	.6200	.8575	.9125	.9625
N_3'	.5025	.7750	.8725	.9825	.6050	.8500	.9125	.9625
N_3''	.6650	.8525	.9275	.9800	.7150	.8700	.9050	.9650

defined in Section 5 for $k = 4$ (the value of d_n was chosen to be $0.05\sigma/n^{\frac{1}{2}} = 0.05/n^{\frac{1}{2}}$ for convenience). Note that for this particular k value we have

$$\boldsymbol{\theta}(1) = (0.0, 0.0, 0.0, 0.8)\,, \qquad \boldsymbol{\theta}(2) = (0.0, 0.4, 0.8, 1.2)\,;$$

and the two mean vectors give a common $\bar{\delta}$ value. Hence this study also serves the purpose of illustrating how the solution of the problem depends on a particular configuration through spacings.

The numerical result given in Table 1 shows that, for fixed $l = 1, 2$ or 3, there is very little difference among N_l, N_l' and N_l''. Because of this fact attention was then restricted to the comparisons of N_1, N_2 and N_3 only in the second study. The results of the second study are summarized in the following tables, which cover a wide range of γ^* values and k values under both $\boldsymbol{\theta}(1)$ and $\boldsymbol{\theta}(2)$ and under

TABLE 2
Monte Carlo result—average sample number for $k = 2(2)10$

Stopping variable	$\boldsymbol{\theta}(1)$				$\boldsymbol{\theta}(2)$			
	$\gamma^* = 0.75$	0.90	0.95	0.99	$\gamma^* = 0.75$	0.90	0.95	0.99
$k = 2$								
n_1	2.00	6.00	9.00	17.00	6.00	21.00	34.00	68.00
N_1	2.98	7.81	12.56	21.40	3.85	13.69	26.97	71.02
$k = 4$								
n_1	5.00	10.00	14.00	23.00	8.00	21.00	34.00	68.00
N_1	6.26	12.87	18.47	30.09	7.34	21.26	34.99	62.65
n_2	5.00	10.00	14.00	23.00	18.00	38.00	54.00	91.00
N_2	9.86	17.22	23.22	35.56	14.45	35.36	47.91	80.55
n_3	5.00	10.00	14.00	23.00	5.00	10.00	14.00	23.00
N_3	4.10	8.98	13.57	24.12	4.60	9.97	14.25	23.97
$k = 6$								
n_1	7.00	12.00	16.00	26.00	8.00	21.00	34.00	68.00
N_1	9.04	17.97	23.56	34.97	8.22	22.95	33.41	61.83
n_2	7.00	12.00	16.00	26.00	25.00	46.00	63.00	101.00
N_2	15.68	24.65	30.10	42.31	22.25	41.40	56.10	90.03
n_3	7.00	12.00	16.00	26.00	4.00	8.00	10.00	17.00
N_3	4.48	11.00	16.02	26.37	3.51	5.96	7.83	12.36
$k = 8$								
n_1	8.00	13.00	18.00	27.00	8.00	21.00	34.00	68.00
N_1	10.68	20.34	26.63	39.31	8.22	23.85	35.26	65.32
n_2	8.00	13.00	18.00	27.00	29.00	52.00	69.00	108.00
N_2	19.39	29.40	35.53	48.52	28.81	49.83	64.65	101.71
n_3	8.00	13.00	18.00	27.00	3.00	5.00	6.00	9.00
N_3	5.16	12.36	18.05	28.88	2.69	3.84	4.91	7.41
$k = 10$								
n_1	9.00	14.00	19.00	29.00	8.00	21.00	34.00	68.00
N_1	12.61	22.44	28.65	40.83	7.63	20.60	32.98	64.70
n_2	9.00	14.00	19.00	29.00	33.00	56.00	74.00	113.00
N_2	22.37	32.00	38.23	52.39	27.47	51.14	67.61	102.34
n_3	9.00	14.00	19.00	29.00	2.00	3.00	4.00	6.00
N_3	5.41	12.98	18.67	30.05	2.20	2.77	3.44	4.98

TABLE 3
Monte Carlo result—observed probability of correct selection for $k = 2(2)10$

Stopping variable	$\boldsymbol{\theta}(1)$				$\boldsymbol{\theta}(2)$			
	$\gamma^* = 0.75$	0.90	0.95	0.99	$\gamma^* = 0.75$	0.90	0.95	0.99
$k = 2$								
N_1	.6000	.8175	.9575	.9925	.4225	.6125	.7725	.9725
$k = 4$								
N_1	.7075	.8950	.9525	1.0000	.7850	.9450	.9875	1.0000
N_2	.8350	.9400	.9875	1.0000	.8750	.9800	.9850	1.0000
N_3	.5250	.7850	.8850	.9850	.6200	.8575	.9125	.9625
$k = 6$								
N_1	.7075	.9275	.9700	.9800	.7875	.9575	.9950	1.0000
N_2	.8900	.9750	.9850	.9950	.9350	.9750	1.0000	1.0000
N_3	.4050	.7450	.8900	.9750	.6125	.7425	.8125	.8825
$k = 8$								
N_1	.8000	.9575	.9925	1.0000	.8300	.9550	.9800	.9975
N_2	.9425	.9950	1.0000	1.0000	.9675	.9925	.9975	1.0000
N_3	.5150	.8150	.9300	.9975	.5925	.6800	.7625	.8025
$k = 10$								
N_1	.8400	.9725	.9975	1.0000	.8225	.9475	.9825	1.0000
N_2	.9725	1.0000	1.0000	1.0000	.9650	.9950	1.0000	1.0000
N_3	.4700	.8300	.9350	1.0000	.6075	.6325	.6725	.7725

each of those three stopping variables. Note that for $k = 2$ we have

$$\boldsymbol{\theta}(1) = (0.0, 0.8), \qquad \boldsymbol{\theta}(2) = (0.0, 0.4).$$

In this special case those stopping variables are identical and the results given simply illustrate how the solution of the problem depends on the distance between the two normal populations. Also, the results for $k = 4$ are reproduced here for the purposes of comparison and completeness.

From Tables 2 and 3 it seems safe to draw the following conclusions:

(a) Although the Procedure R_2 does yield an observed probability of correct selection in excess of γ^* in all cases considered, it is too conservative and the average sample number tends to be unnecessarily large, especially under the slippage configuration ($\boldsymbol{\theta}(1)$).

(b) The average sample number (as a function of γ^*) behaves well under R_3, i.e., EN_3/n_3 is almost 1 for all γ^*. However, the procedure performs very poorly in the probability of correct selection under both configurations unless γ^* is extremely large.

(c) It appears that the Procedure R_1 performs remarkably well under all circumstances. The average sample numbers behave reasonably well; the probabilities of correct selection do exceed γ^* in almost all cases; and it shows improvements when k becomes large. This is not surprising because R_1 is most adaptive among the three procedures considered. Therefore, one should expect that it has the flexibility to make adjustments according to various configurations.

For implementing Procedure R_1 the computing facilities needed for calculating the exact values of those probability integrals are not really essential. Their values can be approximated by using existing statistical tables, and methods of approximations were already discussed in [8].

Acknowledgments. I wish to thank Professors I. Olkin and M. Sobel for several helpful discussions.

REFERENCES

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

[2] BECHHOFER, R. E., KIEFER, J. and SOBEL, M. (1968). *Sequential Identification and Ranking Procedures.* Inst. Math. Statist. Research Monograph Series **3**. Chicago Univ. Press.

[3] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[4] GUPTA, S. S. (1963). Probability integrals of multivariate normal and multivariate *t*. *Ann. Math. Statist.* **34** 792–828.

[5] HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1959). *Inequalities*, 2nd ed. Cambridge Univ. Press.

[6] MARSHALL, A. W. and OLKIN, I. (1974). Majorization in multivariate distributions. *Ann. Statist.* **2** 1189–1200.

[7] MUDHOLKAR, G. S. (1969). A generalized monotone character of d.f.'s and moments of statistics from some well-known populations. *Ann. Inst. Statist. Math.* **21** 277–285.

[8] OLKIN, I., SOBEL, M. and TONG, Y. L. (1976). Estimating the true probability of correct selection for location and scale parameter families. Technical Report No. 110, Dept. of Statistics, Stanford Univ.

[9] ROBBINS, H., SOBEL, M. and STARR, N. (1968). A sequential procedure for selecting the best of k populations. *Ann. Math. Statist.* **39** 88–92.

[10] STROUD, A. H. and SECREST, D. (1966). *Gaussian Quadrature Formulas.* Prentice-Hall, N.J.

[11] TONG, Y. L. (1969). On partitioning a set of normal populations by their locations with respect to a control. *Ann. Math. Statist.* **40** 1300–1324.

[12] TONG, Y. L. (1977). An ordering theorem for conditionally independent and identically distributed random variables. *Ann. Statist.* **5** 274–277.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEBRASKA
LINCOLN, NEBRASKA 68588