

A Class of Sequential Multiple Decision Procedures*

by

Austin M. Barron

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 169

September 1968

* This research was partly supported by contract AF 33(615)69C1244 with the Aerospace Research Laboratories, and partly by GPO6073 from the National Science Foundation.

INTRODUCTION

One problem in statistics is concerned with the ranking and selection of k populations or categories. Each of the k populations Π_i , $i = 1, 2, \dots, k$ has an observable random variable X_i whose distribution depends on some parameter θ_i . The parameters θ_i usually represent an important attribute of the distribution such as the mean or variance. In general, the value of the parameter for any given population is unknown. That is either all the θ_i 's are unknown, or if they are known the exact population from which a given θ comes is unknown. The "best" population is designated to be that one which has the largest parameter value θ_i (or equivalently, the smallest.) Then the fixed sample-size multiple decision problem is to find "optimal" procedures, based on random samples of given size from each population, which will select a subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ of fixed or random size containing the "best" population. The formulation of the problem as such can be found in Bechhofer [2], Lehmann [26], and Gupta [15], [22].

The class of multiple decision rules (or selection and ranking procedures) evolved from short comings in the classical tests of homogeneity of parameters. Testing the hypothesis of equality of means does not answer questions that an experimenter seeks. Particularly if the hypothesis of homogeneity is rejected the experimenter is faced with the situation of having no statistical procedure available to rank the populations.

A more meaningful approach to this type of question is to view it as a multiple decision problem. Early contributions to this more realistic approach were made by Mosteller [27], who considered slippage configurations; Paulson [28] who considered multiple decision problems in the analysis of variance; and Bahador [1] who considered the theory of the k-sample problems. Many authors have made various contributions and modifications to the basic problem, and additional references can be found in Bechhofer [2], Gupta [15, [22], Gupta and Sobel [16], and Lehmann [26]. Most of these contributions fit into two classes generally called (1) the "indifference zone" approach and (2) "subset selection" approach.

In the "indifference zone" approach a population is selected so as to guarantee with a prescribed probability P^* that it is the "best" population whenever the parameters obey some other condition. This other condition is thought of as an "indifference zone" in the parameter space (see Bechhofer [2]). For example, in the problem of selecting the population with the largest mean of k normal populations with common known variance the P^* condition is satisfied whenever the difference between the largest mean and the next largest mean is greater than δ^* , another prescribed constant. Other contributions to this problem can be found in Bechhofer and Sobel [3], Bechhofer, Dunnett and Sobel [4] and Sobel and Huyett [35].

The second approach requires only that the "best" population be included in the selected subset with a prescribed probability P^* for any configuration of the parameters. Thus while with the "indifference zone" approach the experimenter must be given two a priori constants P^* and δ^* , in the "subset selection" approach only one constant P^* need be

given. The size of the selected subset is of course a random variable and so in comparing two procedures both meeting the P^* condition the one which gives a smallest expected subset size is in some sense more desirable. Expected minimal rank and expected sum of ranks of the populations selected in the subset are other performance criteria used in comparing procedures. Contributions to this aspect of the problem can be found in Paulson [28], Gupta [15], [20], [21], Gupta and Sobel [16], [17], [18], [19], Rizvi [33], and Seal [35]. Recent application of the subset selection approach applied to selection and ranking of multivariate normal population can be found in Gupta [23] and Gupta and Panchapakesan [24].

A Bayesian approach to the problem of selection and ranking has been worked on by Dunnett [11], Guttman and Tiao [25], and Deely and Gupta [10] where apriori distributions on the parameter space are assumed. Deely [9] approached the problem through an empirical Bayes study in which he assumes only the existence of an apriori distribution, the exact distribution itself remaining unknown.

A natural modification to the multiple decision problem is the use of sequential and multistage procedures. That is procedures for which the number of observations is not determined in advance but is dependent on the outcome of the observations as they are made. Nearly all of the work in sequential and multistage selection and ranking procedures has been through the "indifference zone" approach. Bechhofer, Dunnett, and Sobel [4] investigated a two-stage multiple decision procedure for the problem of selecting the population with the largest mean of k normal populations with known variances. The problem of sequentially selecting the "best" one of several normal populations with common unknown variances

was worked on by Bechhofer [5] and Bechhofer and Blumenthal [6]. Paulson [29], [30], [31], and [32] has proposed different solutions to the same problem as well as the problem of sequentially selecting the "best" of k binomial populations. Bechhofer, Kiefer and Sobel [7] in a recent monograph have considered sequential procedures for selecting the best of k Koopman-Darmois populations, again using an "indifference zone" approach.

The present thesis deals with sequential and multi-stage procedures using the "subset selection" approach. In this approach no conditions are imposed on the parameter space. The main problem is that of finding a subset of k normal populations, with common known variance, which contains the "best" population with a prescribed probability P^* . The means will be assumed known but not the correct pairing of populations to means. Sequential procedures based on a fixed sample-size type (see Gupta [15], [22] and Gupta and Sobel [17]) will be discussed. The first class of procedures consists of rules of a non-eliminating type; a rule belonging to this class selects and rejects populations at various stages but continues taking samples from all populations until the procedure terminates. At every stage of this above procedure a sample of one is obtained from each of the k populations and a multiple decision rule is employed which determines which of the populations would be selected and which rejected on the basis of that one sample. The number of times a population would be selected in m independent stages then determines whether the experimenter accepts or rejects the population or if another trial must be performed to reach a decision. The probability that the "best" population would be accepted at any individual stage on the basis of the one sample

multiple decision rule is at least P^* . The second class of procedures is derived from the first and is of an eliminating type in that it stops sampling from a population once a decision about it has been reached.

Chapter I is concerned with a class of non-eliminating procedures. Some general monotonic properties of this class of procedures are derived. A specific subclass of these procedures is investigated in detail in Section 1.3 and expressions for the probability of selecting a population and the expected number of stages to come to a decision are found. Section 1.4 deals with approximations to the various probabilities and expectations found in Section 1.3, using a difference equation approach. An approximate minimax rule for choosing a specific procedure that minimizes the maximum number of samples needed to make a decision on each population, among all procedures guaranteeing certain probability conditions, is discussed in Section 1.5. Finally some comparisons to the fixed-sample size procedure for the slippage problem and the equally spaced means problem are offered.

Chapter II discusses the eliminating rule and some monotonic properties of it. Section 2.2 contains a theorem that allows a direct comparison between the probability of selecting a population using the non-eliminating rule, and the probability of selecting a population using the eliminating rule. Section 2.3 presents the results of a Monte-Carlo study of the performance of the second procedure for the slippage problem.

The final chapter offers some generalizations of the procedures discussed in Chapters I and II to more general classes of distributions. Examples using the binomial and gamma populations are shown. The procedures are applied to the problem of finding those populations whose parameters

are "better than a standard". Finally a ranking and selection procedure for k independent binomial populations is discussed.

CHAPTER I
 A CLASS OF NONELIMINATING SEQUENTIAL PROCEDURES
 FOR SELECTING THE BEST OF k NORMAL
 POPULATIONS WITH COMMON KNOWN VARIANCE

1.1 Definition of the General Class of Procedures

In this section the general nature of a non-eliminating sequential multiple decision procedure will be outlined. Let $\Pi_1, \Pi_2, \dots, \Pi_k$ denote k given normal populations with means $\theta_1, \theta_2, \dots, \theta_k$ respectively and common known variance σ^2 . Let $\theta_{[1]} \leq \theta_{[2]} \leq \dots \leq \theta_{[k]}$ be the ranked means, and let $\Pi_{(j)}$ (unknown) be the population with mean $\theta_{[j]}$. The object is to select a small subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ so as to guarantee, with a prescribed probability P^* , that the population with the largest (or equivalently the smallest) mean is included in the selected subset. We denote this event by CS (correct selection). The sequential procedure will be a modification of the following (see Gupta [15], [22] and Gupta and Sobel [17]) fixed sample-size one.

$R(n)$: Take a sample of size n from each of the k populations.

$\Pi_i, i = 1, 2, \dots, k$, and select populations Π_i if and only

if $\bar{X}_i \geq \bar{X}_{\max} - \sigma d / \sqrt{n}$ where d is chosen such that

$\inf_{\Omega} P\{CS|R(n)\} = P^*$ and $\Omega = \{\theta: \theta = (\theta_1, \dots, \theta_k) - \infty < \theta_i < \infty,$

$i = 1, \dots, k\}$. $\bar{X}_i, i = 1, 2, \dots, k$, denotes the sample mean

from Π_i and $\bar{X}_{\max} = \max_{1 \leq i \leq k} \bar{X}_i$.

It has been shown (Gupta [22]) that under this formulation,

$$(1.1.1) \quad p_i = P(\text{selecting } \Pi_i) = \int_{-\infty}^{\infty} \left[\prod_{\substack{j=1 \\ j \neq i}}^k \Phi(x+d+(\theta_{[i]}-\theta_{[j]})\frac{\sqrt{n}}{\sigma}) \right] \varphi(x) dx$$

where $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ and $\Phi(t) = \int_{-\infty}^t \varphi(x) dx$. Therefore, $P\{CS|R(n)\}$

$$= P(X_{(k)} \geq X_{\max} - \frac{\sigma d}{\sqrt{n}}) = \int_{-\infty}^{\infty} \left[\prod_{j=1}^{k-1} \Phi(x+d+(\theta_{[k]}-\theta_{[j]})\frac{\sqrt{n}}{\sigma}) \right] \varphi(x) dx \text{ and thus}$$

the infimum of the probability of a correct selection occurs when $\theta_1 = \dots = \theta_k = \theta$ and is independent of the common value θ . Hence d is defined

by $\int_{-\infty}^{\infty} \Phi^{k-1}(x+d)\varphi(x) dx = P^*$, and it is at once clear that d is independent of n the common sample-size. It is also clear that if $i \geq j$ then $p_i \geq p_j$, for $(\theta_{[i]} - \theta_{[l]}) - (\theta_{[i]} - \theta_{[j]}) = (\theta_{[j]} - \theta_{[l]}) \geq 0$ for all $l = 1, 2, \dots, k$.

For the following procedure it is assumed that the parameters $\theta_1, \theta_2, \dots, \theta_k$ are known, but that the exact pairing of the i th ranked mean $\theta_{[i]}$ to the population it came from is unknown for all $i = 1, 2, \dots, k$. In fact, equation (1.1.1) shows that p_1, p_2, \dots, p_k depend on $\theta_1, \theta_2, \dots, \theta_k$ through the differences $(\theta_{[i]} - \theta_{[j]})$ only, so that many sets of means $(\theta_1, \theta_2, \dots, \theta_k)$ may give rise to the same set of p_1, p_2, \dots, p_k . Thus the experimenter need only have an a priori knowledge of p_1, p_2, \dots, p_k and the results of the chapter will apply for any set $(\theta_1, \theta_2, \dots, \theta_k)$ giving rise to these p_i 's. At the first stage perform $R_{(1)}$ for each of the k populations. For population Π_i define a success if $X_i \geq X_{\max} - \sigma d$, that is if Π_i is accepted under $R_{(1)}$. Similarly,

define a failure if $X_i < X_{\max}$ - od thus rejecting Π_i under $R_{(1)}$. Now define random variables $Y_{11}, Y_{21}, \dots, Y_{k1}$ where $Y_{i1} = 0$ or 1 according as a failure or a success occurs, with population $\Pi_i, i = 1, \dots, k$.

Next draw an additional sample of one from each of the k populations independently of the first sample and perform $R_{(1)}$ on the new sample defining $Y_{12}, Y_{22}, \dots, Y_{k2}$ where $Y_{i2} = 0$ or 1 as a failure or success is gotten from population $\Pi_i, i = 1, 2, \dots, k$, on the second sample. Continuing in this manner, after the m th sample of one drawn from each of the k populations independently of the previous $m-1$ samples perform $R_{(1)}$ and define $Y_{1m}, Y_{2m}, \dots, Y_{km}$ where $Y_{im} = 0$ or 1 according as a failure or a success is gotten from Π_i on the m th sample.

By construction $Y_{1m}, Y_{2m}, \dots, Y_{km}$ are Bernoulli random variables with $P(Y_{im} = 0) = q_i, P(Y_{im} = 1) = p_i, p_i + q_i = 1, i = 1, 2, \dots, k, m \geq 1$. For each $m \geq 1, Y_{1m}, Y_{2m}, \dots, Y_{km}$ are dependent random variables. However, for any fixed population $\Pi_i, Y_{i1}, Y_{i2}, \dots, Y_{im}$ are independent by the nature of the sampling. So if for every $i = 1, 2, \dots, k$ we define

$S_{im} = \sum_{j=1}^m Y_{ij}$ it follows that S_{im} is distributed as a binomial random

variable with parameters m and p_i . It should be pointed out that

$S_{im}, i = 1, 2, \dots, k$ are dependent binomial random variables.

For each Π_i, S_{im} counts the number of successes observed in m independent trials, that is the number of times Π_i is accepted by $R_{(1)}$ with the probability of being accepted at least P^* . Since $p_i > p_j$ for $i > j$ the expected number of successes from $\Pi_{(i)}$ is greater than the expected number of successes from $\Pi_{(j)}$. Therefore $\Pi_{(k)}, (\Pi_{(1)})$ would have the greatest (smallest) expected number of successes in m trials. It is

based on this fact that the procedure \mathcal{A}_1 will be defined.

Let $\{a_j\}$, and $\{b_j\}$ be two sequences of real numbers monotonically increasing to infinity such that

$$(1.1.2) \quad (i) \quad a_1 < 0 < b_1, \quad a_j < b_j, \quad j = 2, 3, \dots$$

$$(ii) \quad P\left\{ \bigcap_{m=1}^{\infty} [a_m < S_{im} < b_m] \right\} = 0, \quad i = 1, 2, \dots, k.$$

The existence of such sequences is guaranteed by the following theorem (see Feller [12]).

Law of the Iterated Logarithm: For all $i = 1, 2, \dots, k$ with probability

$$\text{one we have } \limsup_{m \rightarrow \infty} \frac{\frac{S_{im} - mp_i}{\sqrt{mp_i q_i}}}{(2 \log \log m)^{\frac{1}{2}}} = 1. \text{ Alternatively,}$$

$$(1.1.3) \quad P\{S_{im} > mp_i + \tau(2m p_i q_i \log \log m)^{\frac{1}{2}} \text{ i.o.}\} = 1, \quad 0 < \tau < 1 \quad \text{and}$$

$$P\{S_{im} > mp_i + \tau(2m p_i q_i \log \log m)^{\frac{1}{2}} \text{ for only finitely many } m\} = 1,$$

$$\tau > 1.$$

Then if $b_m = mp_{(1)} + \tau(2mpq \log \log m)^{\frac{1}{2}}$ where $pq = \min_{1 \leq i \leq k} p_i q_i$, $m \geq 1$; $0 < \tau < 1$ and a_m is any monotonically increasing sequence going to ∞ with $a_1 < 0$ and such that $a_m < b_m$, $m \geq 1$ (1.1.2) will be satisfied. Similarly we can choose an appropriate sequence for $\{a_m\}$ by using the second equation in (1.1.3). These choices are actually stronger than is actually needed but are sufficient for showing the existence of the two sequences.

The sequential procedure \mathcal{S}_1 can now be defined. Since the problem is sequential, if no decision has been made about a population Π_i by the m th stage $m \geq 1$ there are three possible choices for the experimenter at the $(m+1)$ st stage: (1) accept Π_i , that is choose Π_i as one of the subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ which will contain the "best" with probability at least P^* , (2) reject Π_i , that is do not include it in the selected subset, or (3) make no decision concerning Π_i and continue onto the $(m+2)$ nd stage. Since the procedure is non-eliminating samples are taken from all populations until all have been either accepted or rejected. A population will be called tagged whenever it falls into the acceptance or rejection regions. The procedure \mathcal{S}_1 is as follows:

\mathcal{S}_1 : "Tag population $\Pi_i, i = 1, 2, \dots, k$ at the very first stage $m \geq 1$ such that $S_{im} \notin (a_m, b_m)$ and mark it "rejected" if $S_{im} \leq a_m$ and "accepted" if $S_{im} \geq b_m$. Continue sampling from all k populations until each one has been tagged, then accept those marked "accepted" and reject those marked "rejected".

If at stage m_i population Π_i is tagged, that is m_i is the first stage such that $S_{im_i} \notin (a_{m_i}, b_{m_i})$, then the experimenter will no longer define $Y_{ij}, j > m_i$ or hence $S_{ij}, j > m_i$. In a sense the procedure ends for Π_i at stage m_i although further samples are taken from it in order to complete the test for the remaining populations.

Condition (1.1.2) guarantees that $m_i < \infty$ for all $i = 1, 2, \dots, k$. The number of samples used in \mathcal{S}_1 is then $k(\max_{1 \leq i \leq k} m_i)$. Since at a finite m_i a decision is reached about Π_i it then follows that for all $i = 1, 2, \dots, k$:

$$(1.1.4) \quad P(\Pi_i \text{ is selected } | \mathcal{L}_1) + P(\Pi_i \text{ is rejected } | \mathcal{L}_1) = 1.$$

Then in general,

$$(1.1.5) \quad P\{CS | \mathcal{L}_1\} = \sum_{m=1}^{\infty} P\left\{ \bigcap_{r=1}^{m-1} [a_r < S_{(k)r} < b_r] \cap [S_{(k)m} \geq b_m] \right\}$$

where $S_{(j)r}$ (unknown) is the sum that corresponds to the population with $\theta_{[j]}$, $j = 1, 2, \dots, k$.

1.2 Some Monotone Properties of \mathcal{L}_1

In the previous section it was shown that each population Π_i gave rise to a sequence of zeros and ones which were summed to provide the test statistics. Consider a genetic population Π , and a sequence of independent random variables Y_1, Y_2, \dots such that $P(Y_i = 0) = 1 - p$, $P(Y_i = 1) = p$, $i = 1, 2, \dots$. Define $S_m = \sum_{i=1}^m Y_i$ for $m = 1, 2, \dots$, and consider pairs of sequences $\underline{c} = (\{a_m\}, \{b_m\})$ such that

$$(1.2.1) \quad \begin{aligned} & \text{(i)} \quad a_m \leq a_{m+1}, \quad b_m \leq b_{m+1}, \quad m = 1, 2, \dots \\ & \text{(ii)} \quad \lim_{m \rightarrow \infty} a_m = \lim_{m \rightarrow \infty} b_m = \infty \\ & \text{(iii)} \quad P\left(\bigcap_{m=1}^{\infty} [a_m < S_m < b_m] \right) = 0 \end{aligned}$$

Def. (1.2.1) Two pairs of sequences \underline{c} and \underline{c}' are said to be ordered if $a'_m \leq a_m$, $b'_m \leq b_m$ (or $a_m \leq a'_m$, $b_m \leq b'_m$) for all $m \geq 1$. We denote this by $\underline{c}' \leq \underline{c}$ (or $\underline{c} \leq \underline{c}'$) and say \underline{c}' is ordered less (greater) than \underline{c} .

The first theorem of this section shows the behavior of (1.1.5) with respect to ordered pairs of sequences. It is an immediate result of the following lemma.

Lemma (1.2.1). Let $\{S_m, m \geq 1\}$ be defined as above and for all \underline{c} satisfying (1.2.1) set $A_m(\underline{c}) = \left\{ \bigcap_{v=1}^{m-1} [a_v < S_v < b_v] \cap [S_m \geq b_m] \right\}$ and

$$P_{\underline{c}}(A) = \sum_{m=1}^{\infty} P(A_m(\underline{c})). \text{ Then if } \underline{c}' \leq \underline{c}, P_{\underline{c}'}(A) \geq P_{\underline{c}}(A).$$

Proof. Since $\underline{c}' \leq \underline{c}$ by definition 1.2.1 $a'_m \leq a_m, b'_m \leq b_m$. Then

clearly $[S_m \geq b_m] \subset [S_m \geq b'_m]$ and also $\bigcap_{r=1}^{m-1} [a_r < S_r < b_r] \subset \bigcap_{r=1}^{m-1} [a'_r < S_r < b'_r]$.

But then this implies that either $\bigcap_{r=1}^{m-1} [a'_r < S_r < b'_r]$ holds or there exists

an $n \leq m-1$ such that $\bigcap_{r=1}^{n-1} [a'_r < S_r < b'_r] \cap [S_n \geq b'_n]$ holds. Thus it

follows that $A_m(\underline{c}) \subset A_n(\underline{c}')$ for some $n = 1, 2, \dots, m$ and so $A_m(\underline{c}) \subset$

$\bigcup_{n=1}^m A_n(\underline{c}')$ and therefore $\bigcup_{m=1}^{\infty} A_m(\underline{c}) \subset \bigcup_{m=1}^{\infty} A_m(\underline{c}')$. It is clear that $A_m(\underline{c})$

$\cap A_n(\underline{c}) = \emptyset$ for $m \neq n$ and for all \underline{c} satisfying (1.2.1) since $A_m(\underline{c})$

is the first time the sequence $\{S_m, m \geq 1\}$ leaves the bounds (a_m, b_m)

and crosses the upper one. Therefore $P(\bigcup_{m=1}^{\infty} A_m(\underline{c})) = \sum_{m=1}^{\infty} P(A_m(\underline{c})) = P_{\underline{c}}(A)$

and from the previous implication $P_{\underline{c}}(A) \leq P_{\underline{c}'}(A)$ which completes the proof.

Theorem (1.2.1). If $\underline{c}' \leq \underline{c}$ then $P\{\text{selecting } \Pi_i | \mathcal{L}_1(\underline{c}')\} \geq$

$P\{\text{selecting } \Pi_i | \mathcal{L}_1(\underline{c})\}$ and $P\{\text{rejecting } \Pi_i | \mathcal{L}_1(\underline{c}')\} \leq P\{\text{rejecting } \Pi_i | \mathcal{L}_1(\underline{c})\}$

for $i = 1, 2, \dots, k$. In particular $P\{CS | \mathcal{L}_1(\underline{c}')\} \geq P\{CS | \mathcal{L}_1(\underline{c})\}$ where

$\mathcal{L}_1(\underline{c})$ is the procedure \mathcal{L}_1 using the sequences $\underline{c} = (\{a_m\}, \{b_m\})$.

Proof. For any fixed $i = 1, 2, \dots, k$ $P(\text{Selecting } \Pi_i | \mathcal{L}_1(\underline{c})) = P_{\underline{c}}(A)$ as

defined in Lemma 1.2.1 and so the first statement of the theorem follows.

From (1.1.4) $P(\text{rejecting } \Pi_1 | \mathcal{L}_1(\underline{c})) = 1 - P(\text{selecting } \Pi_1 | \mathcal{L}_1(\underline{c})) = 1 - P_{\underline{c}}(A)$.

Since $P_{\underline{c}}(A) \leq P_{\underline{c}'}(A)$ for $\underline{c}' \leq \underline{c}$ it then follows that

$P(\text{rejecting } \Pi_1 | \mathcal{L}_1(\underline{c})) \geq P(\text{rejecting } \Pi_1 | \mathcal{L}_1(\underline{c}'))$. Applying the Lemma 1.2.1

$\Pi_{(k)}$ gives the result for a correct selection and completes the proof.

Def. (1.2.2). A set of pairs of sequences \mathcal{C} will be called an ordered

class if for every \underline{c} and \underline{c}' belonging to \mathcal{C} either

$$\underline{c} \leq \underline{c}' \text{ or } \underline{c}' \leq \underline{c}.$$

Thus for an ordered class \mathcal{C} if there exists a $\underline{c} \in \mathcal{C}$ such that $\underline{c} \leq \underline{c}'$ for all $\underline{c}' \in \mathcal{C}$, then the procedure $\mathcal{L}_1(\underline{c})$ maximizes the probability of a correct selection over the family. If no minimal \underline{c} exists in \mathcal{C} the probability of a correct selection can be increased by choosing smaller members of \mathcal{C} at the cost of increasing the chance of selecting "poor" populations.

Now consider another genetic population Π' and the sequence Y'_1, Y'_2, \dots of independent random variables generated by it. Suppose

$P\{Y'_j = 0\} = 1 - p'$ and $P\{Y'_j = 1\} = p'$, $j = 1, 2, \dots$, where $p' < p$. De-

fine $S'_m = \sum_{i=1}^m Y'_i$, $m \geq 1$, and following the notation of Lemma (1.2.1)

set $A'_m(\underline{c}) = \left\{ \bigcap_{r=1}^{m-1} [a_r < S'_r < b_r] \cap [S'_m \geq b_m] \right\}$ and $P_{\underline{c}}(A') = \sum_{m=1}^{\infty} P(A'_m(\underline{c}))$

where $\underline{c} = (\{a_m\}, \{b_m\})$ satisfies (1.2.1). To compare $P_{\underline{c}}(A)$ and $P_{\underline{c}}(A')$ we need the following two lemmas.

Lemma (1.2.2). There exists a sequence of independent identically distributed random variables $\{U_m, m \geq 1\}$ such that

- (i) $P(U_m \leq u) = P(Y'_m \leq u)$ for all u and $m \geq 1$, and
- (ii) $P(U_m \leq Y'_m) = 1$, $m \geq 1$.

Proof. Define a sequence of independent random variables $Z_m = 0$ or 1 such that Z_m is independent of Y_j , $j \neq m$ and $P(Z_m = 1 | Y_m = 1) = p'/p$, $P(Z_m = 0 | Z_m = 0) = 1$ so that $\{Y_m Z_m, m \geq 1\}$ is a sequence of independent and identically distributed random variables. Then let $U_m = Y_m Z_m, m \geq 1$. $P(U_m = 1) = P(Y_m = 1, Z_m = 1) = P(Z_m = 1 | Y_m = 1)P(Y_m = 1) = p'/p$ $p = p'$, and therefore $P(Y_m = 0) = 1 - p'$, $m = 1, 2, \dots$. Clearly then the U_m 's have the same distribution as the sequence of Y_m 's and clearly $P(U_m \leq Y_m) = 1$, which completes the proof.

For the sequence U_1, U_2, \dots as defined in Lemma (1.2.2) let

$$T_m = \sum_{i=1}^m U_i, m \geq 1 \text{ and for every } \underline{c} \text{ satisfying (1.2.1) set}$$

$$B_m(\underline{c}) = \left\{ \bigcap_{r=1}^{m-1} [a_r < T_r < b_r] \cap [T_m \geq b_m] \right\} \text{ and } P_{\underline{c}}(B) = \sum_{m=1}^{\infty} P(B_m(\underline{c})).$$

$P_{\underline{c}}(A)$ and $P_{\underline{c}}(A')$ can now be compared.

Lemma (1.2.3). For every \underline{c} satisfying (1.2.1), $P_{\underline{c}}(A') \leq P_{\underline{c}}(A)$.

Proof. $P_{\underline{c}}(A') = P_{\underline{c}}(B)$ from Lemma (1.2.2) since the distribution of the U_i 's is identical with the distribution of the Y_i 's and both sequences are independent. However the U_i 's are bounded above by the Y_i 's thus for all $m \geq 1$ $P(T_m \leq S_m) = 1$. Therefore, $[T_m \geq b_m] \subset [S_m \geq b_m]$ and also $\bigcap_{v=1}^{m-1} [a_v < T_v < b_v]$ implies either $\bigcap_{v=1}^{m-1} [a_v < S_v < b_v]$ holds or there exists an $n \leq m-1$ such that $\bigcap_{v=1}^{n-1} [a_v < S_v < b_v] \cap [S_n > b_n]$. It then

follows that $B_m(\underline{c}) \subset A_n(\underline{c})$ for some $n = 1, 2, \dots, m$ and so

$$\bigcup_{m=1}^{\infty} B_m(\underline{c}) \subset \bigcup_{m=1}^{\infty} A_m(\underline{c}).$$

As in the proof of Lemma (1.2.1) for $m \neq n$

$A_m(\underline{c}) \cap A_n(\underline{c}) = \emptyset$ and similarly $B_m(\underline{c}) \cap B_n(\underline{c}) = \emptyset$. Thus from the previous implication $P_{\underline{c}}(A') = P_{\underline{c}}(B) \leq P_{\underline{c}}(A)$, and completes the proof.

If we define $r_i = P(\text{selecting } \Pi_{(i)} | \mathcal{J}_1(\underline{c}))$, $i = 1, 2, \dots, k$ we get the following theorem.

Theorem (1.2.2). $r_k \geq r_{k-1} \geq \dots \geq r_1$.

Proof. It will suffice to show that $r_1 \leq r_2$. Now

$$r_i = \sum_{m=1}^{\infty} P\left\{ \bigcap_{v=1}^{m-1} [a_v < S_{(i)v} < b_v] \cap [S_{im} \geq b_m] \right\} \text{ where } S_{im} \text{ is a binomial}$$

random variable with parameters p_i and m . Lemma (1.2.3) shows that r_i is a monotonically increasing function of p_i . In Section 1.1 it was noted that the mean $\theta_{[1]}$ from $\Pi_{(1)}$ is less than $\theta_{[2]}$ from $\Pi_{(2)}$ so that $p_2 > p_1$ and therefore $r_2 > r_1$ and the proof is complete.

The above property describes the monotonicity of the selection probabilities. Another property which follows is given in the next corollary.

Corollary (1.2.1). The procedure $\mathcal{J}_1(\underline{c})$ is unbiased.

Proof. It follows from (1.1.4) and Theorem (1.2.2) that the probability of rejecting the population with the largest mean is less than or equal to the probability of rejecting any other population. Hence $\mathcal{J}_1(\underline{c})$ is unbiased. (See Gupta [22]).

1.3 A Specific Class of Ordered Pairs of Sequences

A: Introduction, Notations, and Definitions

This and the remaining sections of Chapter I will investigate the procedure $\mathcal{J}_1(\underline{c})$ using the following ordered class of pairs of sequences C_1 . Let $a_m = cm - D_1$, and $b_m = cm + D_2$ where $0 < c < 1$ and $D_1, D_2 > 0$. Since $cm + D \leq c'm + D$ for $c \leq c'$ and any fixed D , the class $C_1 = (\underline{c})$ where $\underline{c} = (\{cm - D_1\}, \{cm + D_2\})$ for fixed D_1, D_2 is a ordered class in $c \in (0, 1)$. In fact we will assume c rational in $(0, 1)$ and the actual

probabilities involved will be on the sub-family $C_1^c \subset C_1$ where c is rational.

It will be shown that with a simple change of variables $P\{CS|\mathcal{A}_1(\underline{c})\}$ can be evaluated as a one-dimensional random walk on a finite interval. This section will deal with exact expressions for the probability of selecting or rejecting a population at a given stage of the procedure, and the overall probability of selecting or rejecting a population. Also found are expressions for $E\{M_1|\mathcal{A}_1(\underline{c})\}$ the expected number of stages until population Π_i , $i = 1, 2, \dots, k$ is tagged. Section 1.4 will deal with bounds on the various probabilities and expectations. These bounds are often easier to compute than the corresponding exact expressions and in addition give a better insight into the nature of the procedure.

In what follows several definitions and properties concerning random walk are needed. Most of these properties are well known and can be found in Spitzer's [36] book on random walk. The random walk used here will be defined, on the state space $L = \{x|x = 0, \pm 1, \pm 2, \dots\}$, by a transition function $P(x,y)$ equal to the probability of starting at x and going to y in one step. The main properties of $P(x,y)$ are

1) $\sum_{y \in L} P(x,y) = 1$, 2) $P(x,y) = P(0,y-x)$. The latter permits a description of the walk by $P(0,y)$, $y \in L$. Let $P_0(x,y) = \delta(x,y)$, $P_1(x,y) = P(x,y)$ and $P_n(x,y)$ equal the probability of going from x to y in exactly n steps, where x and y are in L and $\delta(x,y) = 0$ unless $x = y$ when $\delta(x,y) = 1$.

Consider a sequence of discrete, independent, identically distributed random variables $\{y_i, i \geq 1\}$ where $P(y_1 = y) = P(0,y)$ and define

$S_m = \sum_{i=1}^m y_i$, $m \geq 1$. Then $\{y_i, i \geq 1\}$ can be said to generate the random walk. If the walk starts at $x \in L$ then $x + S_m$ is the state of the walk after the m th step. For positive integers d_1, d_2 , define:

$$(1.3.1) \quad B_1 = (-\infty, -sd_1], \quad B_2 = [sd_2, \infty), \quad B = B_1 \cup B_2$$

$$(1.3.2) \quad M_B = \min \{m \geq 1 \mid S_m \in B\}$$

$$(1.3.3) \quad Q_n(x, y) = P\{[S_n = y] \cap [M_B > n]\}, \quad x, y \in L-B, \quad n \geq 0$$

$$(1.3.4) \quad H_B^{(n)}(x, y) = P\{[S_n = y] \cap [M_B = n], \quad x \in L-B, \quad y \in B, \quad n \geq 1\}.$$

It is clear that M_B is the stopping time of the walk. In terms of the procedure $\mathcal{L}_1(\underline{c})$, M_B is the stage at which a given population Π_i is tagged. $Q_n(x, y)$ is the probability of going from x to y in n steps without leaving $L-B$, and $H_B^{(n)}$ is the probability of starting at $x \in L-B$ and leaving $L-B$ at the n th step entering B at y . Analytically, (1.3.3) and (1.3.4) can be described as follows (See Spitzer [36]):

$$(1.3.5) \quad Q_0(x, y) = \delta(x, y), \quad Q_1(x, y) = P(x, y)$$

$$Q_{n+1}(x, y) = \sum_{t \in L-B} Q_n(x, t) Q(t, y) \quad x, y \in L-B, \quad n \geq 1$$

$$(1.3.6) \quad H_B^{(n+1)}(x, y) = \sum_{t \in L-B} Q_n(x, t) P(t, y), \quad x \in L-B, \quad y \in B, \quad n \geq 1$$

As in Section 1.1, for a fixed population Π_i , set $S_{im} = \sum_{j=1}^m y_{ij}$, $m \geq 1$ where y_{ij} , $j \geq 1$ is a sequence of independent identically distributed random variables with $P(y_{ij} = 0) = 1 - p_i$, and $P(y_{ij} = 1) = p_i$.

Define $Z_{ij} = Y_{ij} - c$ for $j \geq 1$ and $i = 1, 2, \dots, k$. Then

$$R_{im} = \sum_{j=1}^m Z_{ij} = \sum_{j=1}^m (y_{ij} - c) = S_{im} - cm, \text{ and hence } [cm - D_1 < S_{im} < cm + D_2] =$$

$$[-D_1 < R_{im} < D_2], \quad [S_{im} \geq cm + D_2] = [R_{im} \geq D_2], \text{ and } [S_{im} \leq cm - D_1] = [R_{im} \leq -D_1].$$

By this transformation the problem of tagging a population Π_i , as it crosses bounds depending on the stage of the procedure is converted to one of a one-dimension random walk $R_{im} = \sum_{j=1}^m Z_{ij}$ where $P(Z_{ij} = 1-c) = p_i$ and $P(Z_{ij} = -c) = 1-p_i$ $j \geq 1$, on the interval $(-D_1, D_2)$. Here the bounds used to reach a decision are independent of the stage of the procedure.

Suppose further that $c = r/s, d_1 = [D_1] + 1, d_2 = [D_2] + 1$, where r, s, d_1 , and d_2 are positive integers. In addition assume that r and s are relatively prime with $r < s$. Then the state space of the walk is all points of the form $\frac{Ns - Mr}{s}$ for all $M > N > 0$. It is a well known theorem of number theory that $xs - yr = I$ has non-negative integer solutions, $x = N, y = M$ with $M > N$, for any integer I provided r and s are relatively prime. In general then the state space is at the form $I/s, I$ an integer. Thus the correspondence $I/s \leftrightarrow I$ enables one to consider the walk on the integer space L via $P(Z_{ij} = -r) = 1-p_i, P(Z_{ij} = s-r) = p_i$, with bounds $-sd_1$, and sd_2 .

The subscript i will now be dropped on p_i and a general population Π with acceptance probability p in using $R(1)$ will be considered. Then the translation function $P(0, y) = \begin{cases} p & \text{if } y = s-r \\ 1-p & \text{if } y = -r \end{cases}$ describes the walk on L . After n steps the walk can only be at points in L of the form $y = (s-r)j - r(n-j)$ for $j = 0, 1, \dots, n$. For fixed $0 \leq j \leq n$

the probability of any such point must then be $p^j(1-p)^{n-j}$, and there are clearly $\binom{n}{j}$ number of ways to reach this point, thus for $n \geq 1$

$$(1.3.7) \quad P_n(0,y) = \begin{cases} \binom{n}{j} p^j(1-p)^{n-j} & \text{if } y = (s-r)j - r(n-j) \text{ } j=0,1,\dots,n, n \geq 1 \\ 0 & \text{elsewhere} \end{cases}$$

The fact that all members of the class C'_1 satisfy (iii) of (1.2.1) is well known and a proof for a more general random walk can be found in Spitzer (1964). A proof for the case $p \neq r/s$ will appear as Corollary (1.4.1) in Section 1.4.

Referring now to (1.3.5) and noting that $L-B = \{-sd_1+1, \dots, sd_2-1\}$ we can express $Q_n(x,y)$ as the n th power of an $N \times N$ matrix $Q = (q_{ij})$ where $N = s(d_1+d_2)-1$ and $q_{ij} = P(i-sd_1, j-sd_1)$ for $i - sd_1, j - sd_1, \in L-B$. Equation (1.3.5) expresses the fact that $Q_n(i,j) = (q_{i+sd_1, j+sd_1}^{(n)})$ where $q_{i,j}^{(n)}$ is the (i,j) entry in Q^n . To insure the procedure $\mathcal{A}'_1(\underline{c})$ will not end on or before a given stage with probability one, the condition $D_1 + D_2 > 1$ will be imposed. This guarantees that any vertical line segment joining the parallel lines $y = r/s x - D_1$ and $y = r/s x + D_2$ will have length greater than one. In other words, at any stage m of the procedure in its original form, Y_j where $Y_j = 0$ or 1 for all $j \geq 1$, there is always a state available to the walk with the bounds. In particular if $D_1 = D_2 = \frac{1}{2}$ and $s = 2$ and $r = 1$, then the procedure $\mathcal{A}'_1(\underline{c})$ reduces to the fixed sample-size procedure $R(1)$. Note that when no ambiguity exists we will write $\mathcal{A}'_1(\underline{c})$ or $\mathcal{A}'_1(r/s)$ for $\mathcal{A}'_1(\underline{c})$ when $\underline{c} = (\{cm-D_1\}, \{cm+D_2\})$ for fixed D_1, D_2 .

B: $P\{\text{selecting } \Pi \text{ at Stage } m | \mathcal{L}_1(r/s)\}$

$P\{\text{rejecting } \Pi \text{ at Stage } m | \mathcal{L}_1(r/s)\}$

Since at any step the random walk R_m can only move $t=s-r$ steps to the right or r steps to the left there are only a finite number of points a population can be absorbed (selected or rejected) at. Therefore, we can define subsets of B_1 and B_2 in (1.3.1) as follows:

$$(1.3.8) \quad B_1' = [-sd_1-r+1, -sd_1-r+2, \dots, -sd_1] \subset B_1$$

$$(1.3.9) \quad B_2' = [sd_2, sd_2+1, \dots, sd_2+s-r-1] \subset B_2$$

Then the points of $B_1' \cup B_2'$ are the actual absorption states of the walk. Moreover, by the nature of the walk absorption can take place at $y \in B_2'$ at stage m if and only if at stage $m-1$, $y-(s-r) \in L-B$. Similarly, absorption at $y \in B_1'$ at stage m can take place if and only if at stage

$$m-1, y+r \in L-B. \text{ From (1.3.6) } H_B^{(n)}(x,y) = \sum_{t=sd_1+1}^{sd_2-1} Q_{n-1}^{(x,t)} P(t,y)$$

for $y \in B_1' \cup B_2'$. However as stated above if $y \in B_2'$ then $P(t,y) > 0$ if and only if $t = y-(s-r)$, and if $y \in B_1'$ then $P(t,y) > 0$ if and only if $t = y+r$. Therefore,

$$(1.3.10) \quad H_B^{(n)}(x,y) = \begin{cases} p q_{x+sd_1, y-(s-r)+sd_1}^{(n-1)} & \text{for } y \in B_2' \\ p q_{x+sd_1, y+r+sd_1}^{(n-1)} & \text{for } y \in B_1' \end{cases}$$

and we get the following theorem.

Theorem (1.3.1). $P\{\text{Selecting } \Pi \text{ at stage } m | \mathcal{L}'_1(r/s)\}$

$$= p \sum_{j=N-(s-r)-1}^N q_{sd_1, j}^{(m-1)}$$

or p times the sum of the last $s-r$ elements in the sd_1 st row of $Q^{(m-1)}$, and similarly $P\{\text{rejecting } \Pi \text{ at stage } m | \mathcal{L}'_1(r/s)\}$

$$= q \sum_{i=1}^r q_{sd_1, j}^{(m-1)}$$

or q times the sum of the first r elements in the sd_1 st row of $Q^{(m-1)}$.

Proof. From (1.3.4) and (1.3.9),

$$P\{\text{selecting } \Pi \text{ at stage } m | \mathcal{L}'_1(r/s)\} = \sum_{y \in B'_2} H_B^{(m)}(0, y)$$

since we select Π only at points of B'_2 . Substituting from (1.3.10) then gives the first result. Similarly, for the second result

$$P\{\text{rejecting } \Pi \text{ at stage } m | \mathcal{L}'_1(r/s)\} = \sum_{y \in B'_1} H_B^{(m)}(0, y),$$

and substituting from (1.3.10) produces the second probability of the theorem which completes the proof.

Corollary (1.3.1).

$$(1.3.11) \quad P\{\text{tagging } \Pi \text{ at stage } m | \mathcal{L}'_1(r/s)\} = p \sum_{j=N-(s-r)+1}^N q_{sd_1, j}^{(m-1)} \\ + q \sum_{j=1}^r q_{sd_1, j}^{(m-1)}$$

Proof. The corollary follows from Theorem (1.3.1) and the fact that tagging Π at stage m means either accepting Π or rejecting Π at stage m , which completes the proof.

The acceptance and rejection probabilities at stage n are of practical interest. Table A1 in the appendix gives these probabilities for selected values of $r, s, d = d_1 = d_2$.

C: $P\{\text{Selecting } \Pi | \mathcal{L}_1(r/s)\}, P\{\text{Rejecting } \Pi | \mathcal{L}_1(r/s)\}$

We now define for $x \in L-B$ and $y \in B$, $H_B(x, y) = \sum_{m=1}^{\infty} H_B^{(m)}(x, y)$,

the probability of starting at x and being absorbed at y . From (1.3.10) for $x \in L-B$

$$(1.3.12) \quad H_B(x, y) = \begin{cases} p \sum_{m=1}^{\infty} q_{x+sd_1, y-(s-r)+sd_1}^{(m-1)} & \text{for } y \in B_2' \\ q \sum_{m=1}^{\infty} q_{x+sd_1, y+r+sd_1}^{(m-1)} & \text{for } y \in B_1' \end{cases}$$

The matrix $Q = (q_{ij})$ is the transition matrix of the random walk restricted to states in $L-B$. It has the following elements $q_{i, i+s-r} = p$ for $i = 1, 2, \dots, N-s+r$, $q_{i, i+s-r} = q$ for $i = s-r+1, \dots, N$ and $q_{ij} = 0$ elsewhere. Thus Q is a substochastic matrix with the following form.

$$(1.3.13) \quad Q = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & p & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & p & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & & \vdots & & & & \vdots & & & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & p & 0 & 0 & 0 & \dots & 0 \\ q & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & p & 0 & 0 & \dots & 0 \\ 0 & q & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & p & 0 & \dots & 0 \\ \vdots & & & & & \vdots & & & & \vdots & & & & \vdots & & \\ 0 & & & \dots & 0 & 0 & q & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & & & \dots & 0 & 0 & 0 & q & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

We can bound Q by a matrix $P = (p_{ij})$, so that P is an irreducible stochastic matrix with $q_{ij} \leq p_{ij}$, $i, j = 1, 2, \dots, N$, by completing the row sums of Q so that they add up to one giving equal weight to each non-zero element in a given row. We make use of the following two lemmas and a remark from Gantmacher [13], [14].

Lemma (1.3.1). (Gantmacher). If $A = (a_{ij})$ and $C = (c_{ij})$ are square matrices of the same order n , where A is irreducible and $C^+ \leq A$, for $C^+ = (|c_{ij}|)$, then for every characteristic g of C and the maximal characteristic root h of A we have $|g| \leq h$, where equality holds if and only if $C^+ = A$.

Lemma (1.3.2) (Gantmacher). If the function $f(x)$ can be expanded in a power series $f(x) = \sum_{p=1}^{\infty} a_p x^p$ in the circle $|x| < h$, then this expression is valid when the scalar argument x is replaced by a matrix A whose characteristic roots all lie within the circle of convergence.

Remark (1.3.1) (Gantmacher). For an irreducible matrix $A \geq 0$ the maximal characteristic root h equals the row sums of A when all the row sums of A are equal.

The remark gives us immediately that the maximal characteristic value of P , being a stochastic matrix, is one. By construction $Q \leq P$ but $Q \neq P$ and hence Lemma (1.3.1) guarantees that $|g| < 1$, where g is any characteristic root of Q . The following theorem can now be proved.

Theorem (1.3.2).

$$(1.3.14) \quad P\{\text{selecting } \Pi | \mathcal{L}'_1(r/s)\} = p \sum_{j=N-s+r+1}^N q^{sd_{1,j}},$$

and

$$(1.3.15) \quad P\{\text{rejecting } \Pi | \mathcal{L}'_1(r/s)\} = q \sum_{j=1}^r q^{sd_{1,j}}$$

where $q^{i,j}$ is the (i,j) entry in $[I-Q]^{-1}$.

Proof. Since, as shown above, all the characteristic roots of Q lie inside the unit circle, Lemma (1.3.2) allows us to form the sum

$$\sum_{m=0}^{\infty} Q^{(m)} = [I-Q]^{-1}. \text{ Apply this to (1.3.12) we get}$$

$$(1.3.16) \quad H_B(x,q) = \begin{cases} p q^{x+sd_1, y+s-r+sd_1} & \text{for } y \in B'_2 \\ q q^{x+sd_1, y+r+sd_1} & \text{for } y \in B'_1 \end{cases}$$

However $P\{\text{selecting } \Pi | \mathcal{L}'_1(r/s)\} = \sum_{y \in B'_2} H_B(0,y) = p \sum_{j=N-s+r+1}^N q^{sd_{1,j}}$, and

similarly $P\{\text{rejecting } \Pi | \mathcal{L}'_1(r/s)\} = q \sum_{j=1}^r q^{sd_{1,j}}$ which completes the proof.

Corollary (1.3.2).
$$p \sum_{j=N-s+r+1}^N q^{sd_{1,j}} + q \sum_{j=1}^r q^{sd_{1,j}} = 1.$$

Proof. Since it is known that with this random walk one must accept or reject Π with probability one, the corollary follows from Theorem (1.3.2).

A more explicit form for the $P\{\text{selecting } \Pi | \mathcal{L}_1(r/s)\}$ can be obtained when $s = r+1$, or $r = 1$. If $s = r+1$ then $r/s = r/r+1$, $r = 1, 2, \dots$, from (1.3.9) $B_2' = [d_2]$ and from (1.3.14) $P\{\text{selecting } \Pi | \mathcal{L}_1(r/s)\} = p q^{sd_1, N}$. Thus we need the $(sd_1, s(d_1+d_2)-1)$ entry in $[I-Q]^{-1}$. Thus,

$$(1.3.17) \quad P\{\text{selecting } \Pi | \mathcal{L}_1(r/r+1)\} = \frac{p q^{sd_1, N}}{|I-Q|}$$

where $q_{sd_1, N}^*$ is the cofactor of the $(s(d_1+d_2)-1, sd_1)$ entry in $[I-Q]$.

Define Δ_ℓ to be the determinant of the $\ell \times \ell$ matrix $B = (b_{ij})$ where $b_{ii} = 1$ for $i = 1, \dots, \ell$, $b_{i, i+1} = -p$ for $i = 1, \dots, \ell-1$, $b_{1, i-r} = -q$ for $i = r+1, \dots, \ell$ and $b_{ij} = 0$ elsewhere. Then

$$(1.3.18) \quad B = \begin{bmatrix} 1 & -p & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -p & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & & & & & & & \vdots & & & & & \vdots & \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -p & 0 & 0 & 0 & \dots & 0 & 0 \\ -q & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & -p & 0 & \dots & 0 & 0 \\ 0 & -q & 0 & 0 & & & & & & & & & & \\ \vdots & & & & & & & \vdots & & & & & \vdots & \\ 0 & 0 & 0 & 0 & \dots & -q & 0 & 0 & 0 & 0 & 0 & \dots & -p & \\ 0 & 0 & 0 & 0 & \dots & 0 & -q & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

expanding by minors using the first row, whose only non-zero elements are $b_{11} = 1$ and $b_{12} = -p$, it is easily seen that Δ_ℓ obeys the following

difference equation;

$$(1.3.19) \quad \Delta_\ell = \Delta_{\ell-1} - \delta(r) \Delta_{\ell-(r+1)} \quad \ell \geq r+1 \quad \text{where}$$

$$\Delta_0 = \Delta_1 = \dots = \Delta_r = 1, \text{ and } \delta(r) = pq^r, \quad r \geq 1.$$

Lemma (1.3.3). For Δ_ℓ as defined above, $\Delta_\ell = \sum_{i=0}^n (-1)^i \binom{\ell-ir}{i} \delta^i(r)$ where $\ell = n(r+1), \dots, n(r+1)+r$, $n \geq 0$.

Proof. The proof is based on mathematical induction. It is clear that for $\ell = 0, 1, \dots, r$ $n = 0$ and $\Delta_\ell = 1$. Suppose the lemma holds for all ℓ' such that $\ell \leq \ell'$ and consider $\Delta_{\ell+1}$. From (1.3.19)

$$\begin{aligned} \Delta_{\ell+1} &= \Delta_\ell - \delta(r) \Delta_{\ell-r} \\ &= \sum_{i=0}^n (-1)^i \binom{\ell-ir}{i} \delta^i(r) - \delta(r) \sum_{i=0}^n (-1)^i \binom{\ell-r-ir}{i} \delta^i(r) \\ &= \sum_{i=0}^n (-1)^i \binom{\ell-ir}{i} \delta^i(r) - \sum_{i=0}^n (-1)^i \binom{\ell-r-ir}{i} \delta^{i+1}(r) \\ &= \sum_{i=0}^n (-1)^i \binom{\ell-ir}{i} \delta^i(r) + \sum_{i=1}^{n+1} (-1)^i \binom{\ell-ir}{i-1} \delta^i(r) \\ &= 1 + \sum_{i=1}^n (-1)^i \left\{ \binom{\ell-ir}{i} + \binom{\ell-ir}{i-1} \right\} \delta^i(r) + (-1)^{n+1} \binom{\ell-(n+1)r}{n} \delta^{n+1}(r). \end{aligned}$$

Now $\binom{\ell-ir}{i} + \binom{\ell-ir}{i-1} = \binom{\ell+1-ir}{i}$ for $i = 1, \dots, n$ and thus

$$\Delta_{\ell+1} = \sum_{i=0}^n (-1)^i \binom{\ell+1-ir}{i} \delta^i(r) + (-1)^{n+1} \binom{\ell-(n+1)r}{n} \delta^{n+1}(r). \text{ If } \ell \text{ is not}$$

of the form $n(r+1)+r$, that is $\ell = n(r+1)+j$, $0 \leq j < r$, then $n(r+1)+j - (n+1)r = n-(r-j) < n$ or $\binom{\ell-(n+1)r}{n} = 0$. If however

$$\ell = n(r+1)+r \quad \binom{\ell-(n+1)r}{n} = \binom{n}{n} = 1, \text{ and in both cases } \Delta_{\ell+1} =$$

$$\sum_{i=0}^m (-1)^i \binom{\ell-ir}{i} \delta^i(r) \text{ where } \ell+1 = m(r+1), \dots, m(r+1)+r \text{ and } m = n \text{ or}$$

$n+1$.

Lemma (1.3.4). For the matrix B in (1.3.14) $b_{i,\ell}^* = p^{\ell-i} \Delta_{i-1}$, for $i = 1, 2, \dots, \ell$.

Proof. $b_{i,\ell}^* = (-1)^{i+\ell}$ (cofactor of $b_{\ell,i}$), where the cofactor of $b_{\ell,i}$ is the determinant of the matrix B with the last row and i th column deleted. Eliminating the last row and the i th column of B we get a matrix B' , of the form below, where B_1 in the upper left corner

$$B' = \begin{pmatrix} B_1 & 0 \\ B_2 & B_3 \end{pmatrix}$$

is an $i-1 \times i-1$ block matrix of the same form as B (1.3.14), B_3 is an $\ell-i \times \ell-i$ diagonal matrix with $-p$ for its diagonal elements, and B_2 is an $i-1 \times \ell-i$ block matrix. Thus $|B'|$ can be gotten by expanding consecutively by columns starting with the ℓ -1st row. Thus

$$|B'| = |B_3| |B_1| \text{ or } |B'| = (-p)^{\ell-i} |B_1|. \text{ But by the definition of } \Delta_{\ell},$$

$$|B_1| = \Delta_{i-1}, \text{ and therefore } b_{i,\ell}^* = (-1)^{i+\ell} (-1)^{\ell-i} p^{\ell-i} \Delta_{i-1} = p^{\ell-i} \Delta_{i-1},$$

$i = 1, \dots, \ell$ which completes the proof.

Theorem (1.3.3). If $s = r+1$, $P(\text{selecting } \Pi|_{\mathcal{A}_1}(r/s)) = \frac{p^{sd_2} \Delta_{N-sd_2}}{\Delta_N}$

where $N = s(d_1+d_2)-1$.

Proof. Applying Lemma (1.3.3) to the matrix $(I-Q)$ it is easily seen

that $|I-Q| = \Delta_N = \sum_{i=0}^n (-1)^i \binom{N-ri}{i} \delta^i(r)$ where $n = \lfloor \frac{N}{r+1} \rfloor$. By Lemma (1.3.4) for $\ell = N$ and $i = sd_1$ noting $N-sd_1 = sd_2-1$ and $sd_1-1 = N-sd_2$ we get $q_{sd_1, N}^* = p^{sd_2-1} \Delta_{N-sd_2}$. Substituting into (1.3.17) gives the desired result and completes the proof.

Corollary (1.3.3). If $r = 1$, $P(\text{selecting } \Pi|_{\mathcal{A}_1}(r/s)) = 1 - \frac{q^{sd_1} \bar{\Delta}_{N-sd_1}}{\bar{\Delta}_N}$

where $\bar{\Delta}_\ell$ equals Δ_ℓ with p and q interchanged.

Proof. If $r = 1$ Theorem (1.3.2) says that, $P(\text{rejecting } \Pi|_{\mathcal{A}_1}(r/s)) =$

$q q^{sd_1, 1} = \frac{q q_{sd_1, 1}^*}{\det(I-Q)}$ where $Q = (q_{ij})$ has $q_{ii} = 1, i = 1, \dots, N$
 $q_{i, i-1} = -q, i = 2, \dots, N, q_{i, i+t} = -p, i = 1, \dots, N-t$ and $q_{ij} = 0$
 elsewhere. Note first that $|I-Q| = |(I-Q)^t|$, but that $(I-Q)^t$ is the
 matrix described in Lemma (1.3.3) with p and q interchanged, thus
 $|I-Q| = \bar{\Delta}_N$. Now observe that $[(I-Q)^{-1}]^t = [(I-Q)^t]^{-1}$ and so $q_{sd_1, 1}^*$
 of $(I-Q)$ is q_{1, sd_1}^{*t} of $(I-Q)^t$. Thus from $(I-Q)^t$ if we delete the
 first column and the i th row $i = 1, \dots, N$ we get a matrix B'' , where

$$B'' = \begin{pmatrix} B_4 & 0 \\ & B_6 \\ B_5 & \end{pmatrix}$$

B_4 is an $i-1 \times i-1$ diagonal matrix with $-q$ as its diagonal elements, B_6 is a $\beta-i \times \beta-i$ matrix of the same form as $(I-Q)^t$, and B_5 is an $\beta-i \times i-1$ block matrix. Expanding by the first column of B'' we get

$$|B''| = |B_4| |B_6| = (-q)^{i-1} |B_6|, \text{ but } |B_6| = \bar{\Delta}_{\beta-i} \text{ so that } |B''| =$$

$$(-q)^{i-1} \bar{\Delta}_{\beta-i}, \quad i = 1, \dots, N. \text{ Thus for } i = sd_1, \quad q_{sd_1,1}^* = q_{1, sd_1}^{*t} =$$

$$(-1)^{sd_1+1} (-1)^{sd_1-1} q^{sd_1-1} \bar{\Delta}_{N-sd_1} = q^{sd_1-1} \bar{\Delta}_{N-sd_1}. \text{ Therefore}$$

$$P(\text{rejecting } \Pi | \mathcal{L}_1(1/s)) = \frac{q^{sd_1-1} \bar{\Delta}_{N-sd_1}}{\bar{\Delta}_N} \text{ and since } P(\text{selecting } \Pi | \mathcal{L}_1(1/s)) =$$

$1 - P(\text{rejecting } \Pi | \mathcal{L}_1(1/s))$ the corollary follows.

D: $E(M_B | \mathcal{L}_1(r/s))$, the Expected Number of Steps

Until a Population is Tagged

$S_m = \sum_{i=1}^m Z_i$ is the sum of independent identically distributed random variables with $EZ_1 = (s-r)p - rq$ $EZ_1^2 = (s-r)^2 p + r^2(1-p)$. Thus if $EZ_1 \neq 0$, then $ES_{M_B} = EZ_1 \cdot EM_B$, and if $EZ_1 = 0$ $ES_{M_B}^2 = EZ_1^2 EM_B$.

Thus we prove the following theorem.

Theorem (1.3.4).

$$(1.3.20) \quad E(M_B | \mathcal{L}_1(r/s)) = \frac{1}{(s-r)p - rq} \left[p \sum_{j=N-s+r+1}^N (j+s-r-sd_1) q^{sd_1, j} \right. \\ \left. + q \sum_{j=1}^r (j-r-sd_1) q^{sd_1, j} \right] \text{ if } p \neq r/s$$

and

$$(1.3.21) \quad E(M_B | \mathcal{F}_1^s(r/s)) = \frac{1}{(s-r)r} \left[p \sum_{j=N-s-r+1}^N (j+s-r-sd_1)^2 q^{sd_1, j} \right. \\ \left. + q \sum_{j=1}^r (s-r-sd_1)^2 q^{sd_1, j} \right] \quad \text{if } p = r/s .$$

Proof. If $p \neq r/s$ $E(M_B | \mathcal{F}_1^s(r/s)) = (EZ_1)^{-1} ES_{M_B}$. However $P(S_{M_B} = y) = H_B(0, y)$ for $y \in B$, thus $ES_{M_B} = \sum_{y \in B} y H_B(0, y) = \sum_{y \in B_1' \cup B_2'}$ and from (1.3.16)

$$ES_{M_B} = p \sum_{y \in B_2'} y q^{sd_1, y-s+r+sd_1} + q \sum_{y \in B_1'} y q^{sd_1, y+r+sd_1} \\ = \sum_{j=N-s+r+1}^N (j+s-r-sd_1) q^{sd_1, j} + q \sum_{j=1}^r (j-r-sd_1) q^{sd_1, j}$$

and (1.3.20) follows. For $p = r/s$, $EZ_1 = 0$, and $EZ_1^2 = (s-r)r$ and again using (1.3.16) one can write $ES_{M_B}^2$ from which (1.3.21) follows and the proof is complete.

It is noted here that another expression for $E(M_B | \mathcal{F}_1^s(r/s))$ can be derived using the fact that $P(M_B = m) = \sum_{y \in B} H_B^{(m)}(0, y)$. Thus from (1.3.10);

$$\begin{aligned}
E(M_B | \mathcal{L}_1(r/s)) &= \sum_{y \in B} \sum_{n=1}^{\infty} n H_B^{(n)}(0, y) \\
&= \sum_{y \in B_1} p \sum_{n=1}^{\infty} n q_{sd_1, y+r+sd_1}^{(n-1)} \\
&\quad + \sum_{y \in B_2} \sum_{n=1}^{\infty} n q_{d_1, y-s+r+sd_1}^{(n-1)}.
\end{aligned}$$

Since it was shown that $\sum_{n=0}^{\infty} Q^n = (I-Q)^{-1}$ it follows that $\sum_{n=1}^{\infty} n Q^{n-1} =$

$(I-Q)^{-2}$, therefore:

$$\begin{aligned}
(1.3.22) \quad E(M_B | \mathcal{L}_1(r/s)) &= \sum_{y \in B_1} \bar{q}^{sd_1, y+r+sd_1} + \sum_{y \in B_2} \bar{q}^{sd_1, y-t+sd_1} \\
&= \sum_{j=1}^r \bar{q}^{sd_1, j} + \sum_{j=N-s+r+1}^N \bar{q}^{sd_1, j}
\end{aligned}$$

where $\bar{q}^{i,j}$ is the (i,j) entry in $(I-Q)^{-2}$. In general, however,

(1.3.22) is not easier to compute than (1.3.20) or (1.3.21).

1.4 Bounds and Approximations

In this section we again consider the procedure $\mathcal{L}_1(r/s)$ as a random walk R_m on the integer space $L-B = [-sd_1+1, \dots, sd_2-1]$ defined by (1.3.7). Bounds and approximations for the probability of a more general random walk leaving $L-B$ at one end and the expected number of stages to do so have been discussed by Wald [37], and Feller [12] as well as others.

A method for getting these bounds following methods of Feller and Wald will be discussed. Applications of these bounds and approximations will enable us to state some theorems concerning the behavior of the procedure $\rho_1(r/s)$, and allow us to choose values of $c = r/s$, d_1 , and d_2 so as to attain a given value of $P(CS | \rho_1(r/s))$ while keeping the expected selected subset size small. Let $t = s-r$.

Let $U_{r/s}(x)$ be the probability of the walk starting at $x \in L-B$ and reaching or crossing sd_2 before $-sd_1$. Then $P(\text{selecting } \Pi | \rho_1(r/s)) = U_{r/s}(0)$. Conditioning on the first step, $U(x) = U_{r/s}(x)$ satisfies the following homogeneous difference equation and boundary conditions:

$$(1.4.1) \quad U(x) = p U(x+t) + q U(x-r), \quad -sd_1 < x < sd_2$$

$$U(x) = 0 \quad \text{for } x \leq -sd_1, \quad U(x) = 1, \quad x \geq sd_2$$

Feller (1957) has shown that there exists a unique solution for (1.4.1).

Since the actual absorbing states of the walk are the s points of $B_1 \cup B_2 = [-sd_1-r+1, \dots, -sd_1] \cup [sd_2, \dots, sd_2+t-1]$ we can replace (1.4.1) by

$$(1.4.2) \quad U(x) = p U(x+t) + q U(x-r), \quad -sd_1 < x < sd_2$$

$$U(x) = 0 \quad x = -sd_1-r+1, \dots, -sd_1$$

$$U(x) = 1 \quad x = sd_2, \dots, sd_2+t-1$$

The characteristic equation of the generating random variable of this walk is $f(x) = (1-p)x^{-r} + px^{s-r}$, setting $f(x) = 1$ we get

$$(1.4.3) \quad px^s - x^r + 1-p = 0$$

Every root y of (1.4.3) leads to a formal solution of the difference equation in (1.4.2), via $U(x) = Ay^x$, which can not satisfy the s boundary conditions. If (1.4.3) has s distinct roots y_1, \dots, y_s then $U(x) = \sum_{i=1}^s A_i y_i^x$ is again a formal solution of (1.4.2), and we can choose A_1, \dots, A_s so that it satisfies the boundary conditions. If one of the roots of (1.4.3) y_i has multiplicity $j > 1$ we still have s constants A_1, \dots, A_s since $y_i^x, xy_i^x, \dots, x^{j-1} y_i^x$ are all formal solutions of (1.4.2).

Suppose $p \neq r/s$. Equation (1.4.3) has unity as a single root, and exactly one more positive root y . For if we consider $f(x) = px^s - x^r + 1-p$, then $f'(x) = x^{r-1}(psx^{s-r} - r)$ so that for $x > 0$ $f(x)$ is a decreasing function on $(0, [\frac{r}{ps}]^{\frac{1}{s-r}})$ and an increasing function on $([\frac{r}{ps}]^{\frac{1}{s-r}}, \infty)$. Further $f(0) = 1-p > 0$ and $f(1) = 0$ hence if $p > r/s$, $f(x)$ crosses the X -axis at $0 < y < 1$, and if $p < r/s$ $f(x)$ crosses the X -axis at $y > 1$.

Define $g(x) = A_1 + A_2 y^x$ where $y \neq 1$ is a positive root of (1.4.3) and $g(-sd_1 - r + 1) = 0$, $g(sd_2) = 1$. Then $g(x)$ is a formal solution of the difference equation in (1.4.2) and $g(x) \geq 0$ for $x = -sd_1 - r + 1, \dots, -sd_1$ and $g(x) \geq 1$ for $x = sd_2, \dots, sd_2 + t - 1$. Thus $g(x) - U(x)$ is a formal solution of the difference equation in (1.4.2) with non-negative

boundary conditions, hence $U(x) \leq g(x)$. In a similar manner define $h(x) = A_3 + A_4 y^x$ where $h(-sd_1) = 0$ and $h(sd_2 + \epsilon - 1) = 1$. Then $U(x) - h(x)$ is again a formal solution of (1.4.2) with non-negative boundary conditions, thus $U(x) \geq h(x)$. Solving for A_1, A_2, A_3, A_4 , with $p \neq r/s$,

$$\frac{y^x - y^{sd_1}}{s(d_1+d_2)+t-1} \leq U(x) \leq \frac{1-y^{sd_1+r-1}}{s(d_1+d_2)+r-1}$$

Then,

$$(1.4.4) \quad \frac{1-y^{sd_1}}{s(d_1+d_2)+t-1} \leq U(0) \leq \frac{1-y^{sd_1+r-1}}{s(d_1+d_2)+r-1}$$

If $p = r/s$ differentiation of (1.4.3) shows that $y = 1$ is a double root, and so $A_5 + A_6 x$ is a formal solution of the difference equation in (1.4.2). Repeating the above arguments we find for $p = r/s$,

$$(1.4.5) \quad \frac{sd_1}{s(d_1+d_2)+t-1} \leq U(0) \leq \frac{sd_1+r-1}{s(d_1+d_2)+r-1}$$

If we make the assumption that $sd_1 \gg r$, and $s(d_1+d_2) \gg t$ we can write (1.4.4) and (1.4.5) as:

$$(1.4.6) \quad P(\text{selecting } \Pi | \mathcal{L}_1(r/s)) \approx \frac{1-y^{sd_1}}{1-y^{s(d_1+d_2)}} \quad \text{if } p \neq r/s$$

$$(1.4.7) \quad P(\text{selecting } \Pi | \mathcal{L}_1(r/s)) \approx \frac{d_1}{d_1+d_2} \quad \text{if } p = r/s$$

These formulae can also be derived using Wald's identity and assuming the walk ends exactly at one of the boundaries. Using (1.4.6) and (1.4.7) we can get approximations for $E(M_B | \mathcal{L}_1(r/s))$. For $p \neq r/s$ $E(M_B | \mathcal{L}_1(r/s)) = \frac{1}{EZ_1} ES_{M_B}$, and if it is assumed we leave L-B at the boundary points $-sd_1$ or sd_2 we get

$$\begin{aligned} ES_{M_B} &\approx sd_2 P(\text{accepting } \Pi | \mathcal{L}_1(r/s)) - sd_1 P(\text{rejecting } \Pi | \mathcal{L}_1(r/s)) \\ &= \frac{s(d_1+d_2)(1-y)^{sd_1} - sd_1(1-y)^{s(d_1+d_2)}}{s(d_1+d_2)(1-y)} \end{aligned}$$

Thus if $p \neq r/s$,

$$(1.4.8) \quad E(M_B | \mathcal{L}_1(r/s)) \approx \frac{s((d_1+d_2)(1-y)^{sd_1} - d_1(1-y)^{s(d_1+d_2)})}{(ps-r)(1-y)^{s(d_1+d_2)}}$$

For $p = r/s$ we use $E(M_B | \mathcal{L}_1(r/s)) = (EZ_1^2)^{-1} ES_{M_B}^2$ which with (1.4.7) gives

$$(1.4.9) \quad E(M_B | \mathcal{L}_1(r/s)) \approx \frac{s^2 d_1 d_2}{(s-r)r}$$

The next lemma demonstrates the well known fact (see Cox and Miller [8]) that the probability a population is not yet absorbed, by the n th stage tend to zero geometrically fast, i.e. $P(M_B > n) = O(\rho^n)$ for some $(0 < \rho < 1)$. For the unrestricted random walk defined by (1.3.7) define:

$$(1.4.10) \quad g(\theta) = \sum_{x \in L} e^{-\theta x} P(0, x)$$

and

$$(1.4.11) \quad g_n(\theta) = \sum_{x \in L} e^{-\theta x} P_n(0, x) \quad n > 1 .$$

Then it is well known that $g_n(\theta) = (g(\theta))^n$.

Lemma (1.4.1).

$$(1.4.12) \quad P(M_B > m) \leq \begin{cases} e^{\theta_1 d_2} g^m(\theta_1) & \text{if } EZ_1 > 0 \\ e^{-\theta_1 d_1} g^m(\theta_1) & \text{if } EZ_1 < 0 \end{cases}$$

where θ_1 is the minimum point of $g(\theta)$, i.e. $e^{-\theta_1} = \left(\frac{r}{s-r} \frac{1-p}{p}\right)^{1/s}$.

Proof. From (1.3.3),

$$P(M_B > m) = \sum_{x \in L-B} Q_m(0, x) \leq \sum_{x \in L-B} P_m(0, x)$$

since $P_n(0, x)$ accounts for all walks from 0 to x in n steps, while $Q_n(0, x)$ accounts for only those which do not leave the boundaries.

Suppose $\theta > 0$ and note d_2 is greater than all points of $L-B$. Then

$$\begin{aligned} \sum_{x \in L-B} P_m(0, x) &\leq \sum_{x \in L-B} e^{-\theta(x-d_2)} P_m(0, x) \leq e^{\theta d_2} \sum_{x \in L} e^{-\theta x} P_m(0, x) \\ &= e^{\theta d_2} g_m(\theta) = e^{\theta d_2} g^m(\theta) . \end{aligned}$$

Thus we get

$$P(M_B > m) \leq e^{\theta d_2} g^m(\theta)$$

which is independent of $\theta > 0$ and if $EZ_1 > 0$ it follows that the minimum point θ_1 of $g(\theta)$ is greater than zero, hence the first inequality of the lemma. For $EZ_1 < 0$ a similar argument produces the second inequality which completes the proof.

It is easily seen that for $\theta_1 = +\ln\left(\frac{r}{s-r} \frac{1-p}{p}\right)^{1/s}$, $g(\theta)$ is between 0 and 1. Hence choosing $\rho = g(\theta_1)$ we have that for $EZ_1 \neq 0$

$$P(M_B > m) \leq C\rho^m \text{ for an appropriate constant } C.$$

Corollary (1.4.1). If $p \neq r/s$ equation (iii) of (1.2.1) holds for

$$a_m = rm/s - D_1, \quad b_m = rm/s + D_2.$$

Proof. Equation (iii) of (1.2.1) says $P\left(\bigcap_{m=1}^{\infty} [a_m < s_m < b_m]\right) = 0$. It will

suffice to show $\lim_{m \rightarrow \infty} P\left(\bigcap_{v=1}^m [a_v < s_v < b_v]\right) = 0$, since this implies the

above probability. However for any $m \geq 1$

$$\bigcap_{v=1}^m [a_v < s_v < b_v] = [M_B > m], \text{ therefore}$$

$$P\left(\bigcap_{v=1}^m [a_v < s_v < b_v]\right) = P(M_B > m) \leq C\rho^m \text{ for some } 0 < \rho < 1$$

As $m \rightarrow \infty$ $P\left(\bigcap_{v=1}^m [a_v < s_v < b_v]\right) \rightarrow 0$ which completes the proof.

In the symmetric boundaries case, $d_1 = d_2 = d$, formulae (1.4.6), (1.4.7), (1.4.8), (1.4.9) simplify to produce a more complete theory.

$$(1.4.13) \quad P\{\text{selecting } \Pi_1(r/s)\} \simeq \begin{cases} \frac{1}{sd} & \text{if } p \neq r/s \\ \frac{1}{2} & \text{if } p = r/s \end{cases}$$

$$(1.4.14) \quad E\{M_B | \mathcal{L}_1(r/s)\} \approx \begin{cases} \frac{sd}{ps-r} \cdot \frac{1-y^{sd}}{1+y^{sd}} & \text{if } p \neq r/s \\ \frac{s^2 d^2}{r(s-r)} & \text{if } p = r/s \end{cases}$$

Theorem (1.4.1). Let p be the acceptance probability of any population Π when the rule $R(1)$ is used. Then for the sequential selection procedure $\mathcal{L}_1(\underline{c})$ where $\underline{c} = (\{cm - D\}, \{cm + D\})$ and $c = r/s$, $D > 0$

$$\lim_{D \rightarrow \infty} P(\text{selecting } \Pi | \mathcal{L}_1(\underline{c})) = \begin{cases} 0 & p < r/s \\ \frac{1}{2} & p = r/s \\ 1 & p > r/s \end{cases}.$$

Proof. Suppose $p < r/s$, then from (1.4.4) with $d_1 = d_2 = d$

$$P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \leq \frac{1-y^{sd+r-1}}{1-y^{2sd+r-1}}, \text{ where } y > 1 \text{ is a root of (1.4.3).}$$

Then clearly as $d \rightarrow \infty$, $\frac{1-y^{sd+r-1}}{1-y^{2sd+r-1}} \rightarrow 0$ and thus $P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \rightarrow 0$.

Similarly if $p > r/s$ (1.4.4) gives $P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \geq \frac{1-y^{sd}}{1-y^{2sd+s-r-1}}$,

where $0 < y < 1$ is a root of (1.4.3). As $d \rightarrow \infty$, $y^{sd} \rightarrow 0$ and therefore

$P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \rightarrow 1$. Finally, if $p = r/s$ (1.4.5) shows that

$$\frac{sd}{2sd+s-r-1} \leq P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \leq \frac{sd+r-1}{2sd+r-1}, \text{ hence as } d \rightarrow \infty$$

$P\{\text{selecting } \Pi | \mathcal{L}_1(\underline{c})\} \rightarrow \frac{1}{2}$ which completes the proof.

Theorem (1.4.2). For any population Π under the conditions of the previous theorem, for large d and $p \neq r/s$ we have,

$$(1.4.15) \quad E\{M_B | \mathcal{J}_1(\underline{c})\} \simeq \frac{sd}{|ps-r|} = \frac{d}{|p-r/s|}$$

Proof. It is clear that $\lim_{d \rightarrow \infty} \frac{1-y^{sd}}{1+y^{sd}} = 1$ and that the sequences approaches one through positive numbers if $0 < y < 1$, that is if $ps-r > 0$, and through negative numbers if $y > 1$, that is if $ps-r < 0$. Hence for large d the result follows from (1.4.14). This completes the proof.

Comparisons between exact \mathcal{J}_1 probabilities as given by (1.3.14) and the corresponding approximations in (1.4.13) are found in Table A2 of the appendix for selected values of c and d . Similar comparisons of the exact expression for \mathcal{J}_1 expectations as given by (1.3.20) and the corresponding approximations in (1.4.14) are found in Table A3.

1.5 A Minimax Approach

The problem stated in Section 1.1 was to select a "small" subset of Π_1, \dots, Π_k which contains the population $\Pi_{(k)}$ ($\Pi_{(1)}$) with a probability greater than or equal to P^* , a prescribed probability regardless of the configuration of the means $\theta_1, \dots, \theta_k$. It was stated that $\Pi_{(k)}$ ($\Pi_{(1)}$) had the largest (smallest) acceptance probability p_k (p_1) associated with $R_{(1)}$. A class of procedures $\mathcal{J}_1(\underline{c}) \subseteq C_1'$ has been proposed and certain probabilities and expectations concerning the procedure have been obtained. Consider now the problem of choosing $\Pi_{(k)}$, since choosing $\Pi_{(1)}$ can be converted into this problem by considering $q_1 = 1-p_1 \geq \dots \geq q_k = 1-p_k$. To pick a specific $\underline{c} \in C_1'$ two constants $c = r/s$ and d an integer must be given. Theorem (1.4.1) guarantees

that for any choice of $c \in (p_{k-1}, p_k)$ there exists a $d = d(c, \epsilon)$ such that for any $\epsilon > 0$

$$(1.5.1) \quad (i) \quad P\{CS|\mathcal{A}_1(\underline{c})\} \geq 1-\epsilon$$

$$(ii) \quad P\{\text{selecting } \Pi_{(k-1)}|\mathcal{A}_1(\underline{c})\} \leq \epsilon$$

regardless of the configuration of $p_1 \leq p_2 \leq \dots \leq p_k$ and hence the configuration of $\theta_{[1]} \leq \theta_{[2]} \leq \dots \leq \theta_{[k]}$. So that for a small enough ϵ the P^* condition can always be satisfied by choosing an appropriate $\underline{c} \in \mathcal{C}_1^!$. If we define S to be the size of the selected subset when the procedure terminates, then using $r_i = P\{\text{selecting } \Pi_{(i)}|\mathcal{A}_1(\underline{c})\} i=1,2,\dots,k$ we get from Theorem (1.2.2), $ES = \sum_{i=1}^k r_i \leq 1+(k-1)r_{k-1}$. Thus we can replace (1.5.1) by

$$(1.5.2) \quad (i) \quad P\{CS|\mathcal{A}_1(\underline{c})\} \geq 1-\epsilon$$

$$(ii) \quad 1-\epsilon < ES \leq 1+(k-1)\epsilon$$

regardless of the configuration of the means $\theta_1, \dots, \theta_k$.

Obviously if for a fixed $c \in (p_{k-1}, p_k)$ d is chosen such that (1.5.2) holds then any choice of $d' > d$ will also satisfy (1.5.2). So that the experimenter has for any $c \in (p_{k-1}, p_k)$ a countably infinite number of procedures \underline{c} which guarantee (1.5.2). It is also clear that (1.5.2) are desirable properties in that the larger the bound on $P\{CS|\mathcal{A}_1(\underline{c})\}$ the smaller the expected number of populations selected. Given two procedures $\underline{c}, \underline{c}' \in \mathcal{C}_1^!$ which satisfy (1.5.2), the procedure

which has the smaller expected number of stages is in some sense preferable. Therefore, the experimenter will want to use the $\underline{c} \in \mathcal{C}'_1$ if it exists which minimizes,

$$(1.5.3) \quad \max_{1 \leq i \leq k} E\{M_i | \mathcal{A}_1(\underline{c})\}$$

over the subclass $\mathcal{C}''_1 \subset \mathcal{C}'_1$ of procedures satisfying (1.5.2) and where M_i is the number of stages until population $\Pi_{(i)}$ (unknown) is tagged. In this section we will show two bounds \bar{c} and c^* between which the \underline{c} minimizing the approximation to (1.5.3) given in (1.4.15) is found.

Theorem (1.4.2) shows that $E\{M_i | \mathcal{A}_1(\underline{c})\}$ is asymptotically proportional to d , so that for a given $\underline{c} \in (p_{k-1}, p_k)$ in order to minimize (1.5.3) over all d such that (1.5.3) is satisfied the experimenter would choose the smallest d . Thus the problem is reduced to finding which $\underline{c} \in (p_{k-1}, p_k)$ produces a $\underline{c} \in \mathcal{C}''_1$, that minimizes (1.5.3).

Definition (1.5.1). For any rational $\underline{c} \in (p_{k-1}, p_k)$ let $\underline{d}_1(\underline{c})$ be the first positive integer such that $r_k = 1 - \epsilon$, and $\underline{d}_2(\underline{c})$ be the first positive integer such that $r_{k-1} = \epsilon$. Finally let $d(\underline{c}) = \max(\underline{d}_1(\underline{c}), \underline{d}_2(\underline{c}))$

The existence of $\underline{d}_1(\underline{c})$ and $\underline{d}_2(\underline{c})$ is guaranteed by Theorem (1.4.1). Then we have the following lemma.

Lemma (1.5.1). $\underline{d}_1(\underline{c})$ is a non-decreasing function of \underline{c} , and $\underline{d}_2(\underline{c})$ is a non-increasing function of \underline{c} for \underline{c} rational in (p_{k-1}, p_k) .

Proof. For any fixed d and $\underline{c}' < \underline{c}$, Theorem (1.2.1) implies that $P\{\text{selecting } \Pi_{(i)} | \mathcal{A}_1(\underline{c}')\} \geq P\{\text{selecting } \Pi_{(i)} | \mathcal{A}_1(\underline{c})\}$ for all $i=1,2,\dots,k$. Now for $d = \underline{d}_1(\underline{c})$ $P\{\text{selecting } \Pi_{(k)} | \mathcal{A}_1(\underline{c})\} \geq 1 - \epsilon$, so that

$P\{\text{selecting } \Pi_{(k)} | \mathcal{L}_1(c')\} \geq 1 - \epsilon$. However $d_1(c')$ was defined to be the smallest integer satisfying this inequality therefore $d_1(c') \leq d_1(c)$. Similarly for fixed $d = d_2(c')$ $P\{\text{selecting } \Pi_{(k-1)} | \mathcal{L}_1(c')\} \leq \epsilon$, therefore $P\{\text{selecting } \Pi_{(k-1)} | \mathcal{L}_1(c)\} \leq \epsilon$, but $d_2(c)$ was defined to be the smallest integer satisfying this inequality therefore $d_2(c') \geq d_2(c)$

which completes the proof.

Lemma (1.5.2). For some $c^* \in (p_{k-1}, p_k)$

$$d(c) = \begin{cases} d_1(c), & c \geq c^* \\ d_2(c) & c \leq c^* \end{cases}$$

Proof. Suppose that for some $c \in (p_{k-1}, p_k)$ $d(c) = d_1(c)$. If $c' > c$ then by Lemma (1.5.1) $d_1(c') \geq d_1(c)$ and $d_2(c') \leq d_2(c)$. By assumption $d_1(c) > d_2(c)$ so that $d_1(c') \geq d_1(c) > d_2(c) \geq d_2(c')$ and $d(c') = d_1(c')$. Suppose now that for some $c \in (p_{k-1}, p_k)$ $d(c) = d_2(c)$. If $c'' < c$ then again by Lemma (1.5.1) $d_1(c) \geq d_1(c'')$ and $d_2(c) \leq d_2(c'')$. Since $d(c) = d_2(c)$ it follows that $d_2(c) > d_1(c)$ and $d_2(c'') \geq d_2(c) > d_1(c) \geq d_1(c'')$ and so $d(c'') = d_2(c)$. Thus it has been shown that if $d(c) = d_1(c)$ for $c \in (p_{k-1}, p_k)$ then $d(c') = d_1(c')$ for all $c' \in [c, p_k)$ and if $d(c) = d_2(c)$ for some $c \in (p_{k-1}, p_k)$ then $d(c') = d_2(c')$ for all $c' \in (p_{k-1}, c]$. Now as $c \rightarrow p_k^-$, then from (1.4.10), $P\{\text{selecting } \Pi_{(k)} | \mathcal{L}_1(c)\} \rightarrow \frac{1}{2}$ so that $d_1(c) \rightarrow \infty$ while $d_2(c)$ approaches its minimum for $c \in (p_{k-1}, p_k)$. Therefore, for c near p_k , $d(c) = d_1(c)$. Similarly for c near p_{k-1} $d(c) = d_2(c)$, so there exists a point (or an interval) such that $d_1(c^*) = d_2(c^*)$, and the

llows.

y (1.5.1). $d(c^*) = \min d(c)$ for $c \in (p_{k-1}, p_k)$ where c^* is
t such that $d_1(c^*) = d_2(c^*)$.

At c^* $d_1(c^*) = d_2(c^*)$. Suppose $c > c^*$ then from Lemma (1.4.2),
 $d_1(c)$. But by Lemma (1.4.1) $d_1(c) \geq d_1(c^*)$. Similarly for
 $d(c) = d_2(c) \geq d_2(c^*)$, but $d(c^*) = d_1(c^*) = d_2(c^*)$ so the
y follows.

roximate values for $d_1(c)$ and $d_2(c)$ can be obtained from

$$c y_{k-1} = 1.$$

by using $\frac{1}{1+y_k} = 1-\epsilon$ and $\frac{1}{1+y_{k-1}} = \epsilon$. Thus

$$p_{k-1} y_k^{-s}$$

$$d_1(c) \approx \frac{\ln \frac{\epsilon}{1-\epsilon}}{s \ln y_k(c)} \quad \text{for } c \in (p_{k-1}, p_k)$$

$$d_2(c) = \left(\frac{-\ln y_k(c)}{\ln y_{k-1}(c)} \right) d_1(c) \approx \frac{\ln \frac{1-\epsilon}{\epsilon}}{s \ln y_{k-1}(c)} \quad \text{for } c \in (p_{k-1}, p_k).$$

$$d(c) \approx \begin{cases} \frac{\ln \frac{\epsilon}{1-\epsilon}}{s \ln y_k(c)} & \text{for } c \geq c^* \\ \frac{\ln \frac{1-\epsilon}{\epsilon}}{s \ln y_{k-1}(c)} & \text{for } c \leq c^* \end{cases}$$

roximate unique value of c^* is given in the following lemma.

or

in

Lemma (1.5.4). For $c \in (p_{k-1}, p_k)$ and large $d(c)$,

$$\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} \approx \begin{cases} \frac{d(c)}{c - p_{k-1}} & \text{for } c \leq \bar{c} \\ \frac{d(c)}{p_k - c} & \text{for } c \geq \bar{c} \end{cases}$$

where $\bar{c} = \frac{p_k + p_{k-1}}{2}$.

Proof. From Theorem (1.4.2), for large values of $d(c)$ and

$c \in (p_{k-1}, p_k)$ $E\{M_i | \mathcal{L}_1(c)\} \approx \frac{sd(c)}{|p_i s - r|} = \frac{d(c)}{|p_i - c|}$ where $c = r/s$. It is

clear that since $c > p_{k-1}$ and $p_i \leq p_{k-1}$ for $i = 1, 2, \dots, k-2$ that

$|p_i - c| = c - p_i \geq c - p_{k-1} = |p_{k-1} - c|$. Thus $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} = \max(E\{M_{k-1} | \mathcal{L}_1(c)\}, E\{M_k | \mathcal{L}_1(c)\})$. Now $|p_k - c| = p_k - c \geq c - p_{k-1}$ if and only

if $c \leq \frac{p_k + p_{k-1}}{2} = \bar{c}$. Therefore, for $c \leq \bar{c}$ $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} \approx \frac{d(c)}{c - p_{k-1}}$

and for $c \geq \bar{c}$ $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} \approx \frac{d(c)}{p_k - c}$, which completes the proof.

Lemma (1.5.5). (1) $\frac{d_1(c)}{p_k - c}$ is an increasing function of c , and

(2) $\frac{d_2(c)}{c - p_{k-1}}$ is a decreasing function of $c \in (p_{k-1}, p_k)$.

Proof. Lemma (1.5.1) shows that $d_1(c)$ is a non-decreasing function and $d_2(c)$ is a non-increasing function of $c \in (p_{k-1}, p_k)$. Now $p_k - c$ decreases monotonically to 0 as $c \rightarrow p_k$ and $c - p_{k-1}$ increases monotonically as $c \rightarrow p_k$. Thus in (1) the numerator increases and the denominator decreases hence the fraction increases as c increases; and in (2) the numerator decreases and the denominator increases as c increases hence the fraction decreases as c increases. This completes the proof.

Theorem (1.5.1). For $c \in (p_{k-1}, p_k)$.

$$\min_c \max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} \simeq \begin{cases} \min_{c^* \leq c \leq \bar{c}} \frac{d_1(c)}{c - p_{k-1}} & \text{for } c^* \leq \bar{c} \\ \min_{\bar{c} \leq c \leq c^*} \frac{d_2(c)}{p_k - c} & \text{for } \bar{c} \leq c^* \end{cases}$$

Proof. Suppose $c^* \leq \bar{c}$, then for $c \geq \bar{c}$ Lemma (1.5.2) and Lemma (1.5.4)

show $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} = \frac{d_1(c)}{p_k - c}$. However, by Lemma (1.5.5) this is an increasing

function of c , and thus the minimum for $c \geq \bar{c}$ occurs at $c = \bar{c}$. For $c < c^*$, Lemma (1.5.2) and Lemma (1.5.4) show $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} = \frac{d_2(c)}{c - p_{k-1}}$ which

decreases as c increases, by Lemma (1.5.5), so that the minimum for

$c \leq c^*$ occurs at c^* . Thus it follows that $\min_{p_{k-1} < c < p_k} \max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\}$

occurs for some $c \in [c^*, \bar{c}]$. Since $c \leq \bar{c}$ by Lemma (1.5.4)

$$\min_{c^* \leq c \leq \bar{c}} \max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} = \min_{c^* \leq c \leq \bar{c}} \frac{d(c)}{c - p_{k-1}}, \text{ and since } c > c^*$$

$d(c) = d_1(c)$ and the first approximation of the theorem follows. A similar

argument for $c^* \geq \bar{c}$ will provide the second approximation and hence

prove the theorem.

$$\begin{aligned} \text{Corollary (1.5.2). If } c^* = \bar{c}, \quad \min_{p_{k-1} < c < p_k} \max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\} &= \frac{d_1(c)}{p_k - \bar{c}} \\ &= \frac{d_2(\bar{c})}{\bar{c} - p_{k-1}} \end{aligned}$$

Proof. The proof follows directly from the theorem by allowing $c^* \rightarrow \bar{c}$.

We have shown then that the $c \in (p_{k-1}, p_k)$ which asymptotically minimizes (1.4.18) is found between c^* and \bar{c} , and that $d(c)$ approximated by (1.5.6) then produces a $\underline{c} \in C_1''$, that is a procedure which satisfies (1.5.2). This still leaves the experimenter with the problem of choosing a specific c if $c^* \neq \bar{c}$. It will be shown in Section 1.7 that quite often $c^* \approx \bar{c}$, so that the experimenter will not be "far" from the minimum for any choice of c between \bar{c} and c^* . Numerical evidence indicates that if \bar{c} and c^* are significantly apart, the minimum takes place at or near c^* . Another advantage to using c^* is that the approximation of $d(c^*)$ can be given as a function of p_{k-1} , p_k and ϵ so that the experimenter need not find the roots y_k and y_{k-1} to (1.4.3).

Thus the above discussion suggests that an approximate minimax rule which has certain desirable properties would be $\rho_1(\underline{c}^*)$ where $\underline{c}^* = (\{c^* - D^*\}, \{c^* + d^*\})$ with c^* as defined in (1.5.7) and (1.5.8) and from (1.5.4)

$$(1.5.10). \quad D^* = d(c^*) = (1-c^*) \frac{\ln \frac{\epsilon}{1-\epsilon}}{\ln \frac{p_{k-1}}{p_k}} = \frac{\ln \frac{1-\epsilon}{\epsilon}}{\ln \frac{p_k}{p_{k-1}} \cdot \frac{1-p_{k-1}}{1-p_k}}.$$

This of course is not the only choice of $\underline{c} \in C_1'$ available. It depends on the need of the experimenter who may wish to replace (ii) of (1.4.16) by some other condition such as $r_i \leq \epsilon$ for some $i < k-1$, or he may require less of ES than the bounds given in (ii) of (1.4.7). The use of c^* , d^* is only one suggestion toward meeting practical requirements of a good sequential test.

1.6 The case $c = \frac{1}{2}$

In this section we consider the case $c = \frac{1}{2}$ in the ordered family $\underline{c} = (\{cm-d_1\}, \{cm+d_2\})$ for $c \in (0,1)$. With $c = \frac{1}{2}$ all the bounds and approximations in Section (1.4) are exact expressions, so that an analytic comparison can be made between the results in Section (1.3) and those in Section (1.4). First however, a closed form expression for $P\{\text{accepting } \Pi_i \text{ at stage } m | \underline{c}\}$ is derived following a method found in Feller [12].

For $c = \frac{1}{2}$ the random walk R_{im} defined by (1.3.7) is generated by x_1, x_2, \dots where $P\{x_j = 1\} = p$, $P\{x_j = 0\} = q$, $p+q=1$, $j \geq 1$. Let $u_n(x) = P\{\text{starting at } x, R_{im} \geq 2d_2 \text{ when } -2d_1 < R_{iv} < 2d_2, v = 1, 2, \dots, m-1\}$. By conditioning on the first step and defining the following initial conditions,

$$(1.6.1) \quad u_0(x) = \begin{cases} 0 & x < 2d_2 \\ 1 & x = 2d_2 \end{cases}$$

$$u_n(-2d_1) = u_n(2d_2) = 0, \quad n \geq 1$$

we can write the following difference equation, satisfying (1.6.1),

$$(1.6.2) \quad u_{n+1}(x) = p u_n(x+1) + q u_n(x-1), \quad -2d_1 < x < 2d_2$$

To convert (1.6.2) into a homogeneous difference equation in one variable,

set $U_x(s) = \sum_{n=0}^{\infty} u_n(x) s^n$. It is clear that $|U_x(s)| < \infty$ for $|s| < 1$,

since $|U_x(s)| \leq \sum_{n=0}^{\infty} |u_n(x) s^n| \leq \frac{1}{1-|s|}$. Also $U_x(s)$ obeys the following

conditions,

$$(1.6.3) \quad U_{2d_2}(s) = 1, \quad U_{-2d_1}(s) = 0 .$$

Now multiply both sides of (1.6.2) by s^{n+1} and sum on n to get

$$(1.6.4) \quad U_x(s) = ps U_{x+1}(s) + qs U_{x-1}(s), \quad -2d_1 < 2x < 2d_2 ,$$

a homogeneous difference equation in x satisfying (1.6.3). To solve (1.6.4) we try solutions of the form $U_x(s) = \lambda^x(s)$, which when substituted into (1.6.4) gives $\lambda^x(s) = ps \lambda^{x+1}(s) + qs \lambda^{x-1}(s)$ or,

$$(1.6.5) \quad ps \lambda^2(s) - \lambda(s) + qs = 0 .$$

The roots of (1.6.5) are

$$(1.6.6) \quad \lambda_1(s) = \frac{1 + (1 - 4pqs^2)^{\frac{1}{2}}}{2ps}, \quad \lambda_2(s) = \frac{1 - (1 - 4pqs^2)^{\frac{1}{2}}}{2ps} .$$

Hence $U_x(s) = A(s) \lambda_1^x(s) + B(s) \lambda_2^x(s)$ where $A(s)$ and $B(s)$ are chosen so that $U_x(s)$ obeys (1.6.3). It is easily seen that

$$(1.6.7) \quad A(s) = \frac{-\lambda_2^{-2d_1}(s)}{\lambda_1^{-2d_1}(s)\lambda_2^{2d_2}(s) - \lambda_1^{2d_2}(s)\lambda_2^{-2d_1}(s)}$$

and

$$(1.6.8) \quad B(s) = \frac{\lambda_1^{-2d_1}(s)}{\lambda_1^{-2d_1}(s)\lambda_2^{2d_2}(s) - \lambda_1^{2d_2}(s)\lambda_2^{-2d_1}(s)}$$

Hence,

$$(1.6.9) \quad U_x(s) = \frac{\lambda_1^{-2d_1}(s)\lambda_2^x(s) - \lambda_2^{-2d_1}(s)\lambda_1^x(s)}{\lambda_1^{-2d_1}(s)\lambda_2^{2d_2}(s) - \lambda_1^{2d_2}(s)\lambda_2^{-2d_1}(s)}$$

$$= \frac{\lambda_2^{x+2d_1}(s) - \lambda_1^{x+2d_1}(s)}{\lambda_2^{2(d_1+d_2)}(s) - \lambda_1^{2(d_1+d_2)}(s)}$$

Then $U_x(s)$ is a rational fraction since numerator and denominator are polynomials in s multiplied by $(1-4pqs^2)^{\frac{1}{2}}$. Now define θ by $\frac{1}{\cos \theta} = 2(pq)^{\frac{1}{2}}s$ and note $\lambda_1(s) = \left(\frac{q}{p}\right)^{\frac{1}{2}} e^{i\theta}$, $\lambda_2(s) = \left(\frac{q}{p}\right)^{\frac{1}{2}} e^{-i\theta}$. Therefore we can write,

$$(1.6.10) \quad U_x(s) = \left(\frac{q}{p}\right)^{x-2d_2} \frac{\sin(x+2d_1)\theta}{\sin 2(d_1+d_2)\theta} = \left(\frac{q}{p}\right)^{x-2d_2} \frac{Q(s)}{V(s)}$$

where $Q(s)$ and $V(s)$ are polynomials in s . Now $\sin(2(d_1+d_2)\theta) = 0$

when $\theta = \theta_j = \frac{j\pi}{2(d_1+d_2)}$, $j = 1, 2, \dots$, so that $V(s) = 0$ when

$$s = s_j = \frac{1}{\sqrt{pq} \cos \frac{j\pi}{2(d_1+d_2)}} \quad j = 0, 1, \dots, 2(d_1+d_2) .$$

For $j = 0$ and $2(d_1+d_2)$ we get roots of the numerator. Clearly the term $j=d_1+d_2$ must be disregarded as no number $s_{(d_1+d_2)}$ corresponds to it. All other roots s of $Q(s)$ will be eliminated automatically later. Note that the degree of $Q(s)$ is at most one higher than that of $U(s)$ so

that we can write $\frac{Q(s)}{V(s)} = As + B + \sum_{j=1}^a \frac{p_j}{s-s_j}$ where $a = 2(d_1+d_2)-1$

and Σ' refers to the sum except for $j = d_1+d_2$.

Multiplying $\frac{Q(s)}{V(s)}$ by $s - s_j$ and taking the limit as $s \rightarrow s_j$ we get,

$$(1.6.11) \quad p_j = \lim_{s \rightarrow s_j} (s-s_j) \frac{Q(s)}{V(s)} = \frac{Q(s_j)}{V'(s_j)} .$$

Now $\frac{dV(s)}{ds} = \frac{dV}{d\theta} \frac{d\theta}{ds} = [2(d_1+d_2) \cos 2(d_1+d_2)\theta] \left[\frac{1}{2\sqrt{pq}} \frac{\sin \theta}{\cos^2 \theta} \right]^{-1}$, therefore from (1.6.11) the value of p_j is given by,

$$(1.6.12) \quad p_j = \frac{-\sin(x+2d_1) \frac{j\pi}{2(d_1+d_2)} \sin \frac{j\pi}{2(d_1+d_2)}}{4(d_1+d_2) \sqrt{pq} \cos^2 j\pi \cos^2 \frac{j\pi}{2(d_1+d_2)}} .$$

$$\begin{aligned}
\text{Thus } \left(\frac{q}{p}\right)^{-(z+2d_2)} U_x(s) &= A + BS + \sum_{j=1}^a \frac{p_j}{s-s_j} \\
&= A + BS + \sum_{j=1}^a -\frac{p_j}{s_j} \frac{1}{1 - \frac{s}{s_j}} \\
&= A + BS + \sum_{k=0}^{\infty} s^k \sum_{j=1}^a -\frac{p_j}{s_j^{k+1}}
\end{aligned}$$

Therefore for $n > 1$, and the definition of $U_x(s)$,

$$\begin{aligned}
(1.6.13) \quad u_n(x) &= \frac{\left(\frac{q}{p}\right)^{(z+2d_2)} 2^{2(d_1+d_2)-1}}{y/\sqrt{pq} (d_1+d_2)} \sum_{j=1}^a \\
&= \frac{2^n p^{\frac{n-4d_2-2z}{2}} q^{\frac{n+4d_2-2z}{2}} 2^{2(d_1+d_2)-1}}{2(d_1+d_2)} \sum_{j=1}^a \cos^{n-1} \frac{j\pi}{2(d_1+d_2)} \\
&\quad \sin \frac{(x+2d_1)j\pi}{2(d_1+d_2)} \sin \frac{j\pi}{2(d_1+d_2)} \cos^{n+1} \frac{j\pi}{2(d_1+d_2)} 2^{n+1} (pq)^{\frac{n+1}{2}} \\
&\quad \sin \frac{x+2d_1}{2(d_1+d_2)} j\pi \sin \frac{j\pi}{2(d_1+d_2)}
\end{aligned}$$

setting $x = 0$ in (1.6.13) and substituting p_i for p , $1 - p_i$ for q we get the $P\{\text{selecting } \Pi_{(i)} \text{ at stage } m | \rho_1(\frac{1}{2})\}$ for any $m > 1$. In particular in the symmetric boundary case $d_1 = d_2 = d$,

$$(1.6.14) \quad P\{\text{selecting } \Pi_{(i)} \text{ at stage } m | \mathcal{L}_1(\frac{1}{2})\}$$

$$= \frac{2^n p_i^{\frac{n+4d}{2}} q_i^{\frac{n-4d}{2}}}{4d} \sum_{j=1}^{4d-1} (-1)^j \cos^{m-1} \frac{j\Pi}{4d} \sin \frac{j\Pi}{2} \sin \frac{j\Pi}{4d} .$$

The proof of the equivalence of (1.6.14) and Theorem (1.3.1) can be found for the case $p = \frac{1}{2}$ in Spitzer [36].

Now consider the polynomial in (1.4.3) for $r = 1, s = 2$. The roots are easily seen to be 1 and $\frac{q}{p}$, thus from (1.4.4) and (1.4.5) we get,

$$(1.6.15) \quad P\{\text{selecting } \Pi | \mathcal{L}_1(\frac{1}{2})\} = \frac{1 - (\frac{q}{p})^{2d_1}}{2(d_1 + d_2)} \quad \text{if } p \neq \frac{1}{2} ,$$

$$1 - (\frac{q}{p})$$

and

$$(1.6.16) \quad P\{\text{selecting } \Pi | \mathcal{L}_1(\frac{1}{2})\} = \frac{d_1}{d_1 + d_2} \quad \text{if } p = \frac{1}{2} .$$

Also in this case the sets B'_1 and B'_2 defined in (1.3.8) and (1.3.9) become $[-2d_1]$ and $[2d_2]$ respectively. Thus the expressions for $E\{M_B | \mathcal{L}_1(\frac{1}{2})\}$ given in (1.4.8) and (1.4.9) are exact. Hence for the symmetric boundary problem $d_1 = d_2 = d$ we can write,

$$(1.6.17) \quad P\{\text{selecting } \Pi | \mathcal{L}_1(\frac{1}{2})\} = \begin{cases} \frac{p^{2d}}{p^{2d} + q^{2d}} & \text{if } p \neq \frac{1}{2} \\ \frac{1}{2} & \text{if } p = \frac{1}{2} \end{cases}$$

and

$$(1.6.18) \quad E\{M_B | \mathcal{L}_1(\frac{1}{2})\} = \begin{cases} \frac{d}{p-\frac{1}{2}} \frac{p^{2d-q} 2d}{p^{2d+q} 2d} & \text{if } p \neq \frac{1}{2} \\ 4d^2 & \text{if } p = \frac{1}{2} \end{cases}$$

If the equation in (1.6.4) is to agree with the result in Theorem (1.3.3) for $r = 1$, $s = 2$ then we must have $\Delta_{4d-1} = \Delta_{2d-1}(p^{2d+q} 2d)$ where Δ_j as given in Lemma (1.3.3), is $\sum_{i=0}^n (-1)^i \binom{j-i}{i} \delta^i$ with $\delta = p(1-p)$ and $n = \lfloor \frac{j}{2} \rfloor$. By expanding several terms it can be seen that the equality,

$$(1.6.19) \quad \sum_{i=0}^{2d-1} (-1)^i \binom{4d-1-i}{i} \delta^i = \left[\sum_{i=0}^{d-1} (-1)^i \binom{2d-1-i}{i} \delta^i \right] \left[\sum_{i=0}^d (-1)^i \binom{d-i}{i} \frac{2d}{2d-i} \delta^i \right],$$

holds. The proof of (1.6.19) follows from Theorem (1.6.1) below and the equivalence of the two probabilistic arguments of Sections (1.3) and (1.4).

So for completeness we prove that $\sum_{i=0}^d (-1)^i \binom{d-i}{i} \frac{2d}{2d-i} \delta^i = p^{2d+q} 2d$.

First we need the following lemma.

Lemma (1.6.1).

$$(1.6.20) \quad \sum_j \binom{n}{2j-n} \binom{j}{n-1} = \binom{u+1}{n-1}.$$

Proof. Let $\langle x^a \rangle f(x)$ stand for the coefficient of x^a in the expansion of $f(x)$. Then $\sum_j \binom{n}{2j-u} (-1)^{j-n+1} \binom{n}{u-2j+n} = \langle x^{u-n+2} \rangle (1-x^2)^{-n} (1+x)^n$
 $= \langle x^{u-n+2} \rangle (1-x)^{-n}$
 $= \binom{-n}{u-n+2} (-1)^{u-n+2}$

Now apply the fact that $(-1)^k \binom{-n}{k} = \binom{n+k-1}{k}$ to both sides of the equality and the lemma follows.

Theorem (1.6.1).

$$(1.6.21) \quad \sum_{i=0}^d (-1)^i \binom{2d-i}{i} \frac{2d}{2d-i} (p(1-p))^{i-p} p^{2d} + (1-p)^{2d}, \quad d \geq 1.$$

Proof. It is obvious that the coefficient of p^{2d} on both sides of (1.6.21) is 2, hence it is enough to show that the coefficient of p^a of the left hand side is $(-1)^a \binom{2d}{a}$ $a = 0, 1, \dots, 2d-1$. Now

$$\begin{aligned} \langle p^a \rangle \text{LHS} &= \sum_i (-1)^i \frac{2d}{2d-i} \binom{2d-i}{i} \binom{i}{a-i} (-1)^{a-i} \\ &= (-1)^a 2d \sum_i \frac{(2d-i-1)!}{(2d-2i)!(a-i)!(2i-a)!} \\ &= (-1)^a 2d \sum_i \frac{(2d-a)!}{(2d-2i)!(2i-a)!} \frac{(2d-i-1)!}{(a-i)!(2d-a-1)!} \frac{(2d-a-1)!}{(2d-a)!} \\ &= (-1)^a \frac{2d}{2d-a} \sum_i \binom{2d-a}{2d-2i} \binom{2d-i-1}{2d-a-1}. \end{aligned}$$

Now let $i = 2d-j-1$ and apply (1.6.20) and the theorem follows.

We have showed that $P\{\text{selecting } \Pi | \mathcal{A}_1(\frac{1}{2})\} = \frac{p^{2d} \Delta_{N-2d}}{\Delta_N}$

$$\begin{aligned} & \frac{p^{2d} \sum_{i=0}^{d-1} (-1)^i \binom{2d-1-i}{i} \delta^i}{\sum_{i=0}^{2d-1} (-1)^i \binom{2d-1-i}{i} \delta^i} \\ &= \frac{p^{2d}}{\sum_{i=0}^d (-1)^i \binom{2d-i}{i} \frac{2d}{2d-i} \delta^i} = \frac{p^{2d}}{p^{2d} + q^{2d}} \end{aligned}$$

so that the expression derived in Section 1.3 agrees with that derived in Section 1.4.

1.7 Some Sample Size Comparisons of \mathcal{L}_1 and $R(n)$

A. Introduction

In this section we offer some numerical comparisons between the procedures \mathcal{L}_1 and $R(n)$. Comparisons are difficult in general because analytic expressions involving the two procedures are not available, and because of the small number of tables available on the performance of $R(n)$. Two special configurations of the means $\theta_1, \theta_2, \dots, \theta_k$ will be considered. The first is called the "slippage configuration", that is,

$$(1.7.1) \quad \theta_{[1]} = \theta_{[2]} = \dots = \theta_{[k-1]} = \theta, \quad \theta_{[k]} = \theta + \delta, \quad \delta > 0 .$$

Tables of $P\{\text{selecting } \Pi_1 | R(n)\}$ have been tabulated in this case for selected values of P^* , k , n , and δ by Deely and Gupta [10]. The second configuration called the "equally-spaced means" configuration is,

$$(1.7.2) \quad \theta_{[1]} = \theta, \quad \theta_{[2]} = \theta + \delta, \quad \theta_{[3]} = \theta + 2\delta, \dots, \theta_{[10]} = \theta + (k-1)\delta, \quad \delta > 0 .$$

Tables of $P\{\text{selecting } \Pi_1 | R(n)\}$ have been tabulated in this case for selected values of P^* , k , n , and δ by Gupta [22].

For any multiple-decision rule R , consider the following inequalities, for $0 < \epsilon < 1$,

$$(1.7.3) \quad \begin{aligned} & \text{(i) } P\{\text{CS} | R\} \geq 1 - \epsilon \\ & \text{(ii) } 1 - \epsilon < E\{S | R\} \leq 1 + (k-1)\epsilon . \end{aligned}$$

TABLE 1

Values of \bar{c} and c^* for the Slippage Problem
for the Normal Population: $P^* = .75$

k	δ	0.05	0.10	0.20	0.30	0.40	0.50	0.60	1.00	2.00
2		.7500	.7495	.7480	.7450	.7420	.7370	.7315	.7020	.6060
		.7501	.7499	.7499	.7489	.7489	.7474	.7430	.7421	.7213
3		.7530	.7555	.7595	.7625	.7645	.7650	.7640	.7490	.6580
		.7531	.7558	.7607	.7654	.7699	.7735	.7762	.7841	.7794
4		.7540	.7575	.7640	.7695	.7735	.7765	.7780	.7700	.6865
		.7541	.7577	.7651	.7720	.7781	.7839	.7889	.8018	.8042
5		.7545	.7580	.7675	.7740	.7790	.7830	.7860	.7830	.7060
		.7551	.7592	.7685	.7763	.7832	.7899	.7961	.8127	.8207
6		.7545	.7585	.7680	.7760	.7820	.7870	.7905	.7916	.7200
		.7551	.7597	.7964	.7782	.7861	.7936	.8001	.8201	.8328
7		.7550	.7590	.7700	.7775	.7845	.7900	.7940	.7980	.7305
		.7556	.7602	.7790	.7797	.7885	.7963	.8033	.8254	.8396
8		.7550	.7600	.7705	.7795	.7865	.7925	.7970	.8030	.7400
		.7556	.7612	.7714	.7816	.7904	.7987	.8060	.8246	.8483
9		.7550	.7600	.7710	.7800	.7875	.7940	.7995	.8065	.7470
		.7556	.7612	.7719	.7820	.7913	.8000	.8083	.8327	.8526
10		.7555	.7605	.7700	.7815	.7890	.7955	.8010	.8095	.7535
		.7560	.7617	.7729	.7835	.7927	.8614	.8096	.8350	.8565
25		.7560	.7620	.7745	.7855	.7955	.8035	.8110	.8280	.7965
		.7566	.7632	.7753	.7874	.7989	.8089	.8189	.8505	.8864
50		.7570	.7625	.7760	.7875	.7975	.8070	.8150	.8360	.8190
		.7570	.7637	.7768	.7894	.8008	.8123	.8226	.8573	.9010

TABLE 2

Values of \bar{c} and c^* for the Slippage Problem

for the Normal Population: $P^* = .90$

$k \backslash \epsilon$	0.05	0.10	0.20	0.30	0.40	0.50	0.60	1.00	2.00
2	.9000	.8995	.8980	.8950	.8910	.8860	.8800	.8470	.7215
	.9001	.9000	.8988	.8990	.8982	.8973	.8961	.8903	.8602
3	.9015	.9030	.9045	.9050	.9045	.9035	.9010	.8805	.7760
	.9016	.9033	.9057	.9077	.9094	.9114	.9117	.9124	.8952
4	.9025	.9040	.9075	.9090	.9105	.9105	.9095	.8855	.8040
	.9026	.9043	.9085	.9113	.9147	.9170	.9189	.9223	.9146
5	.9025	.9045	.9085	.9115	.9130	.9140	.9140	.9035	.8210
	.9026	.9047	.9094	.9136	.9167	.9200	.9226	.9274	.9221
6	.9025	.9050	.9095	.9125	.9150	.9160	.9165	.9095	.8335
	.9026	.9052	.9104	.9144	.9186	.9216	.9244	.9324	.9274
7	.9030	.9055	.9100	.9140	.9160	.9180	.9190	.9135	.8430
	.9030	.9057	.9109	.9159	.9194	.9231	.9266	.9348	.9314
8	.9030	.9060	.9105	.9145	.9170	.9195	.9205	.9160	.8505
	.9030	.9062	.9113	.9163	.9202	.9246	.9277	.9363	.9345
9	.9030	.9060	.9110	.9150	.9180	.9205	.9215	.9185	.8505
	.9030	.9062	.9118	.9168	.9212	.9253	.9285	.9378	.9370
10	.9030	.9060	.9110	.9150	.9185	.9210	.9230	.9210	.8615
	.9030	.9062	.9118	.9168	.9216	.9257	.9299	.9401	.9390
25	.9030	.9062	.9115	.9170	.9215	.9245	.9275	.9310	.8950
	.9030	.9062	.9122	.9186	.9243	.9288	.9336	.9496	.9524
50	.9040	.9075	.9135	.9190	.9235	.9275	.9305	.9365	.9130
	.9040	.9077	.9142	.9205	.9262	.9317	.9362	.9510	.9646

For example for $k = 6$ and $\delta = .40$ the c which minimizes M satisfies $c = .7820 \leq c \leq .7861 = c^*$. Numerically, for small δ , $c^* \approx \bar{c}$ while as δ increases c^* becomes larger than \bar{c} . The table indicates that for any $\delta > 0.10$ the maximum difference occurs for $k = 2$, and that this difference increases with δ and decreases with k . Table 2 which gives the same values as Table 1 for $P^* = .90$ shows the same general behavior. It was stated in Section 1.5 that whenever c^* and \bar{c} were significantly apart the procedure using c^* was better than the procedure using \bar{c} in the following sense, that $(c^*, d(c^*))$ gives a smaller value of M than $(\bar{c}, d(\bar{c}))$. This was based on numerical evidence such as given in Table 3. For $P^* = .75$ and $\delta = 1.0$ and 2.0 Table 3 presents a comparison of $\rho_1(\bar{c})$ and $\rho_1(c^*)$ for selected values of k .

The upper number is the approximate value of $\max_{1 \leq i \leq k} E\{M_i | \rho_1(\bar{c})\}$ and the lower number is the approximate value of $\max_{1 \leq i \leq k} E\{M_i | \rho_1(c^*)\}$ both as given in (1.4.14). In each case $\max_{1 \leq i \leq k} E\{M_i | \rho_1(\bar{c})\} \geq \max_{1 \leq i \leq k} E\{M_i | \rho_1(c^*)\}$.

For example for $\delta = 2.0$, $k = 6$ $\max_{1 \leq i \leq k} E\{M_i | \rho_1(c^*)\} = 9.4 < 10.9 = \max_{1 \leq i \leq k} E\{M_i | \rho_1(\bar{c})\}$. Of course both procedures satisfy (1.7.3) with $\epsilon = .001$.

For $P^* = .75$, Table 4 compares M with n when $\rho_1(c^*)$ and $R(n)$ satisfy (1.7.3). The upper value is the expected sample size M while the middle value is the fixed-sample size n . The lower value gives the ratio of M to n . The smaller the ratio the more inclined we are to use $\rho_1(c^*)$ over $R(n)$. The savings in the number of samples needed to achieve (1.7.3) with $\epsilon = .001$ using $\rho_1(c^*)$ over $R(n)$ vary for different values of k from better than 50% to 25% for $\delta \leq .50$ to less of a saving for $.50 \leq \delta \leq 1$. For large values of δ , such as

TABLE 3

Expected Sample Size Comparisons Between $\mathcal{L}_1(\bar{c})$ and $\mathcal{L}_1(c^*)$
for the Normal Population: $P^* = .75$

$\delta \backslash k$	2	3	4	5	6	7	8	9	10	25	50
1.0	16.5	22.8	27.0	30.0	36.9	34.0	35.3	36.3	37.8	44.8	48.4
	16.1	22.1	26.1	29.0	31.2	32.8	34.2	35.2	36.4	43.7	46.7
2.0	5.8	7.2	8.5	9.5	10.4	10.9	11.7	12.0	12.0	16.6	18.4
	5.1	6.7	7.9	8.7	9.4	10.0	10.6	10.9	11.4	14.8	17.2

$\delta = 2$, the fixed sample procedure $R(n)$ requires less samples than $\mathcal{L}_1(c^*)$ to achieve (1.7.3) with $\epsilon = .001$ and so is a more preferable procedure.

As an example for $k = 6$ populations and $\delta = 0.4$, the expected number of samples from each population needed to satisfy (1.7.3) with $\epsilon = .001$ using $\mathcal{L}_1(c^*)$ is 172.5 or a total of 1035 observations, while using $R(n)$ a sample of 248.1 must be taken from each population, a total of 1488.6 observations needed to satisfy (1.7.3) with $\epsilon = .001$. This is better than a 30% savings in using $\mathcal{L}_1(c^*)$.

Table 5 gives the same data as Table 4 but for $P^* = .90$. In this case the savings are generally less when using $\mathcal{L}_1(c^*)$ over $R(n)$, and δ generally must be much smaller. In general in both tables for a fixed k as δ increases the ratio $\frac{M}{N}$ increases. For a fixed δ , $\frac{M}{N}$ is smallest for $k = 2$, while the maximum increases from $k = 7$ to $k = 50$ as δ increases. Thus it appears that $\mathcal{L}_1(c^*)$ is a better procedure when the mean that has slipped to the right has not slipped far. Close examination of both tables reveals that the ratio $\frac{M}{N}$ does not increase

TABLE 4

Sample Size Comparisons for the Sequential and

Fixed Sample=Sized Rules for the Slippage

Configuration for the Normal Population: $P^* = .75$

$k \backslash \delta$	0.05	0.10	0.20	0.30	0.40	0.50	0.60	1.00	2.00
2	5422.7 11240.0 .482	1315.7 2810.0 .468	336.6 702.5 .479	151.8 312.2 .486	87.5 175.6 .498	57.3 112.4 .510	40.6 78.1 .520	16.1 28.1 .573	15.1 19.7 .766
3	7553.9 13600.0 .555	1922.4 3460.0 .556	491.4 850.0 .578	219.1 377.7 .580	124.3 212.5 .585	82.0 136.0 .603	57.7 94.5 .611	22.1 34.0 .650	6.7 8.5 .788
4	9890.1 15120.0 .654	2418.0 3780.0 .640	586.1 945.0 .620	259.4 420.0 .618	148.0 236.2 .627	95.6 151.2 .632	67.5 102.9 .656	26.1 37.8 .690	7.9 9.5 .832
5	10752.1 15640.0 .687	2485.3 3910.0 .636	637.2 977.5 .652	284.1 434.4 .654	161.6 244.4 .661	104.9 156.4 .672	74.8 108.7 .688	29.0 39.1 .742	8.7 9.8 .888
6	10752.1 15880.0 .677	2754.5 3970.0 .694	679.7 992.5 .685	305.9 441.1 .696	171.5 248.1 .691	111.9 158.8 .705	84.5 110.4 .765	31.2 39.7 .786	9.4 9.9 .949
7	12119.5 16400.0 .739	2835.3 4100.0 .692	695.0 1025.0 .678	317.9 455.5 .698	180.1 256.2 .693	117.0 164.0 .713	84.8 114.0 .744	32.8 41.0 .800	10.0 10.2 .980
8	12695.4 16920.0 .750	2803.7 4230.0 .663	708.1 1057.5 .670	325.3 470.0 .692	183.9 264.4 .796	120.4 169.2 .712	86.4 117.6 .735	34.2 42.3 .809	10.6 10.6 1.000
9	12695.4 17440.0 .728	2803.7 4360.0 .643	749.6 10900.0 .688	331.1 484.4 .685	188.6 272.5 .692	126.7 174.4 .726	88.9 121.2 .733	35.2 43.6 .807	10.9 10.9 1.000
10	13037.6 17680.0 .737	2960.2 4420.0 .670	731.4 1105.0 .662	346.3 491.1 .705	192.9 276.3 .698	127.2 176.8 .719	91.7 122.9 .746	36.4 44.2 .824	11.4 11.1 1.027
25	13037.6 20440.0 .638	3192.4 5110.0 .625	817.1 1277.5 .640	369.7 567.7 .651	212.7 319.4 .666	143.5 204.4 .702	103.5 142.1 .726	43.7 51.1 .855	14.8 12.8 1.156
50	13695.1 21600.0 .634	3096.1 5400.0 .573	786.9 1350.0 .583	365.1 599.9 .609	220.0 337.5 .665	148.1 216.0 .686	107.2 150.1 .714	46.7 51.0 .864	17.2 13.5 1.27

TABLE 5

Sample Size Comparisons for the Sequential and
Fixed Sample-Sized Rules for the Slippage

Configuration for the Normal Population: $P^* = .90$

$k \backslash \delta$	0.05	0.10	0.20	0.30	0.40	0.10	0.60	1.00	2.00
2	8786.5 15360.0 .572	2079.3 3840.0 .541	542.4 960.0 .565	257.6 426.6 .604	148.0 240.0 .608	97.0 153.6 .632	69.2 106.5 .650	28.3 38.4 .737	9.1 9.6 .948
3	12469.0 17440.0 .716	3069.0 4360.0 .704	849.1 1090.0 .779	385.1 484.4 .795	217.9 272.5 .800	145.6 174.4 .835	100.9 121.2 .833	46.4 43.6 1.064	12.2 10.9 1.12
4	14282.3 18600.0 .786	3781.5 4650.0 .813	1024.6 1162.5 .881	477.0 516.6 .923	260.5 290.6 .896	170.2 186.0 .915	122.4 129.3 .947	49.1 46.5 1.056	14.5 11.6 1.25
5	14282.3 19600.0 .729	4177.9 4800.0 .853	1085.0 1225.0 .886	494.3 544.3 .908	290.9 306.3 .950	186.7 196.0 .953	135.0 136.2 .991	51.0 59.0 1.091	16.2 12.3 1.32
6	14282.3 20160.0 .708	4230.0 5040.0 .839	1148.4 1260.0 .911	531.2 559.9 .949	300.5 315.0 .954	201.2 201.6 .998	148.8 140.1 1.062	59.0 50.4 1.171	17.8 12.6 1.413
7	18633.7 20720.0 .899	4463.5 5180.0 .862	1148.4 1295.0 .887	550.7 575.5 .957	333.2 323.8 1.029	219.8 207.2 1.061	152.7 144.0 1.060	63.3 51.8 1.222	19.1 13.0 1.469
8	18633.7 21040.0 .886	4653.4 5260.0 .885	1335.6 1315.0 1.016	566.9 584.4 .970	337.9 328.8 1.028	224.1 210.4 1.065	161.2 146.2 1.103	66.8 52.6 1.270	20.3 13.2 1.538
9	18633.7 21320.0 .874	4653.4 5330.0 .873	1335.6 1332.5 1.002	589.0 592.2 .995	341.6 333.1 1.026	232.2 213.2 1.089	167.5 148.2 1.130	70.0 53.3 1.313	21.4 13.3 1.609
10	18633.7 21680.0 .859	4653.4 5420.0 .859	1335.6 1355.0 .986	589.0 602.2 .978	341.6 338.8 1.008	246.6 216.8 1.137	168.4 150.7 1.117	72.0 54.2 1.328	72.4 13.6 1.647
25	18633.7 24040.0 .775	4653.4 6010.0 .774	1381.8 1520.5 .909	632.1 667.7 .947	383.4 375.6 1.021	265.6 240.4 1.105	195.3 167.1 1.169	86.7 60.1 1.443	30.4 15.0 2.027
50	18801.4 25600.0 .734	5257.6 6400.0 .821	1338.0 1600.0 .836	683.3 711.0 .961	403.9 400.0 1.010	278.1 256.0 1.086	200.2 177.9 1.125	104.1 64.0 1.627	37.8 16.0 2.36

monotonically for very small values of δ , this is due to the fact that p_k and p_{k-1} do not vary greatly as k increases. In fact, for $P^* = .75$ and $\delta = 0.05$, $p_k = .763$, $p_{k-1} = .748$ for $k = 7, 8$, and 9 to three significant places. Thus for small δ rounding errors play a somewhat bigger role than for larger δ . Another factor in all tables of this section is that in practice the exact value of C^* cannot be used to obtain the various results, but a close approximation of c^* is used instead. This tends to remove some of the apparent monotonicity as well.

C. The Equally-Spaced Means Problem

For the configuration in (1.7.2) equation (1.1.1) becomes,

$$(1.7.6) \quad p_i = \int_{-\infty}^{\infty} \left[\prod_{\substack{j=1 \\ j \neq i}}^k \phi(x+d-(j-i)\delta \frac{\sqrt{n}}{\sigma}) \right] \varphi(x) dx, \quad i = 1, 2, \dots, k$$

Using tables and extensions of tables in [22] (1.7.6) was evaluated for $\sigma = 1$. A numerical comparison of $\beta_1(c^*)$ and $R(n)$ was carried out using the same method as in Section 1.7B. That is M and N were evaluated so that (1.7.3) holds for $\beta_1(c^*)$ and $R(n)$ with $\epsilon = .001$. For $P^* = .75$, Table 6 gives the values of M , n and the ratio $\frac{M}{N}$ for selected values of k and δ . The upper value being M , the middle value being n , and the lower value being the ratio $\frac{M}{N}$. Table 7 contains the same information for $P^* = .90$.

It can be seen from Table 6 that the behavior of the ratio $\frac{M}{N}$ is similar to that in Table 4 for the slippage configuration. That is the smaller δ is the smaller the ratio. In fact for $k = 2$ there is a 50% or better saving in the expected number of samples using $\beta_1(c^*)$

TABLE 6

Sample Size Comparisons for the Sequential and
Fixed Sample-Sized Rules for the Equally-Spaced Means
Configuration for the Normal Population: $P^* = .75$

$k \backslash \delta$	0.05	0.10	0.20	0.30	0.40	0.50	0.60
2	5422.7	1315.7	336.0	151.8	87.5	57.3	40.6
	11240.0	2810.0	702.5	312.2	175.6	112.4	78.1
	.482	.468	.479	.486	.498	.510	.520
3	7430.9	1841.7	470.9	208.9	118.1	77.2	54.5
	13440.0	3360.0	840.0	372.3	210.0	134.4	93.4
	.553	.548	.561	.561	.562	.574	.584
4	8258.4	2194.8	551.4	280.3	141.2	90.4	64.2
	14880.0	3720.0	930.0	463.3	232.5	148.8	103.4
	.555	.590	.593	.605	.607	.608	.623
5	8824.8	2393.9	619.2	276.9	155.8	101.0	72.4
	15880.8	3970.0	992.5	441.1	248.1	158.8	110.4
	.556	.603	.624	.628	.628	.636	.656

TABLE 7

Sample Size Comparisons for the Sequential and
Fixed Sample-Sized Rules for the Equally-Spaced Means
Configuration for the Normal Population: $P^* = .90$

$k \backslash \delta$	0.05	0.10	0.20	0.30	0.40	0.50	0.60
2	8786.5	2079.3	542.4	257.6	146.0	97.0	69.2
	15360.0	3840.0	960.0	426.6	240.0	153.6	106.5
	.572	.541	.565	.604	.608	.632	.650
3	1186.4	3262.7	783.2	372.1	214.1	144.6	100.6
	17440.0	4360.0	1090.0	484.4	272.5	174.4	121.2
	.680	.748	.719	.768	.786	.829	.830
4	14435.0	3792.2	961.9	459.1	259.4	174.5	123.8
	18640.0	4660.0	1165.0	517.7	291.3	186.4	129.5
	.774	.814	.826	.887	.890	.936	.956
5	17436.4	4517.0	1207.6	540.1	310.1	197.0	147.7
	19520.0	4880.0	1220.0	542.2	305.0	195.2	135.7
	.893	.926	.990	.996	1.02	1.01	1.09

instead of $R(n)$ for $\delta \leq .50$. For a fixed δ as k increases from 2 to 5 $\frac{M}{N}$ increases. Of course Table 6 only goes to $k = 5$, and since Table 4 showed erratic behavior as k increased to 50 one cannot make a general statement about this monotonic behavior. Table 7 shows the same general behavior but, the savings using $\rho_1(c^*)$ over $R(n)$ are, in general, less.

Thus based on the numerical computations for the slippage configuration and the equally spaced means configuration one conclusion is that $\rho_1(c^*)$ is a more preferable procedure when the means are close and $R(n)$ is a more preferable procedure as any one mean gets significantly larger (or smaller) than the others.

CHAPTER II

A CLASS OF ELIMINATING SEQUENTIAL PROCEDURES
FOR SELECTING THE BEST OF K NORMAL
POPULATIONS WITH COMMON KNOWN VARIANCE

2.1 Definition of the Class of Procedures

In this section the general nature of an eliminating sequential multiple decision procedure will be outlined. As in Chapter I let $\Pi_1, \Pi_2, \dots, \Pi_k$ denote k given normal populations with means $\theta_1, \theta_2, \dots, \theta_k$ respectively and a common known variance σ^2 . Again the object is to select a small subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ so as to guarantee, with a prescribed probability P^* , the the population with the largest (or equivalently the smallest) mean is included in the selected subset. The notation of this chapter will be consistent with that of Chapter I.

The class of procedures $\mathcal{L}_1(\underline{c})$ defined in Section 1.1 for all \underline{c} satisfying (1.1.2) is a non-eliminating class, that is at each stage of any procedure in the class a sample must be taken from each at the k populations regardless of the fact that a decision to accept or reject a particular population may have occurred at an earlier stage. Thus, from the point of view of a population about which a decision has been reached the additional samples taken from it are unnecessary. These additional samples are not used to

make a decision about the population they are taken from, but rather to keep the vector of probabilities (p_1, p_2, \dots, p_k) for $R(1)$ as defined in (1.1.1) constant from stage to stage of the procedure.

In order to avoid taking these additional samples, a sequential ranking and selection procedure \mathcal{S}_2 modifying \mathcal{S}_1 will be defined for a suitable class of pairs of sequences \underline{c} . The following definition is required.

Definition (2.1.1). For the procedure \mathcal{S}_2 to be defined let N_1 be that stage of the procedure at which a decision is made, for the first time, to accept or reject one or more populations. Let $N_2 \geq N_1$ be the next stage of the procedure at which such a decision is made, and in general let $N_i \geq N_{i-1}$ be the stage of the procedure at which the i th decision to accept or reject one or more populations is made.

Define N to be the stage at which the procedure terminates.

It is clear that since there are k populations to start with $N \leq N_k$. In Section 2.2 it is shown that for all $\underline{c} \in C_1'$, defined in Section 1.3, N is finite with probability one. If for some $j < k$ $N_j = N$ then we take $N_i = N$ for all $i > j$.

Definition (2.1.2). Let k_j $j=1, 2, \dots, k-1$ denote the number of populations undecided upon immediately after stage N_j . Clearly $k > k_1 \geq k_2 \geq \dots \geq k_{k-1}$.

For the procedure \mathcal{S}_2 as for the procedure \mathcal{S}_1 it is assumed that the means are known but that the exact pairing of the i th ranked mean $\theta_{[i]}$ to the population it came from is unknown for all $i=1, 2, \dots, k$. Again as in Chapter I we need not know the means, but

rather their ranked differences $(\theta_{[i]} - \theta_{[j]})$. The method by which the test statistics $T_{1m}, T_{2m}, \dots, T_{km}$, $m \geq 1$ for the procedure \mathcal{L}_2 , corresponding to the test statistics $S_{1m}, S_{2m}, \dots, S_{km}$ $m \geq 1$ for the procedure \mathcal{L}_1 as defined in Section 1.1, are obtained is as follows.

At each stage $m=1, 2, \dots, N_1$ draw a sample of one from each of the k populations independently of all previous samples. Perform the fixed sample size procedure $R_{(1)}$, as defined in Section 1.1, and define random variables $Y_{1m}, Y_{2m}, \dots, Y_{km}$ where $Y_{im} = 0$ or 1 as a failure (Π_i is rejected under $R_{(1)}$) or a success (Π_i is accepted under $R_{(1)}$) occurs with population π_i on the m th stage. Up to this point the procedure is identical to \mathcal{L}_1 . However, at stage N_1 for the first time a decision to accept or reject one or more of the k populations is made, and there remain $k_1 \leq k-1$ populations to decide upon. If $k_1 \leq 1$ the procedure \mathcal{L}_2 defined below will terminate and no further samples need be taken. If, however, $k_1 \geq 2$ then for $m=N_1+1, \dots, N_2$ draw a sample of one, independently of all previous samples, from each of the remaining k_1 populations only. Perform $R_{(1)}$ on each of these samples, and define $Y_{i_1 m}, Y_{i_2 m}, \dots, Y_{i_{k_1} m}$ where $Y_{i_j m} = 0$ or 1 as a failure or a success is obtained for population Π_{i_j} on the m^{th} stage and $\Pi_{i_1}, \Pi_{i_2}, \dots, \Pi_{i_{k_1}}$ are the unknown remaining populations. No further samples are taken from populations about which a decision to accept or reject has been made at N_1 , here then is how \mathcal{L}_2 differs from \mathcal{L}_1 . Now at stage N_2 more populations are decided upon and k_2 remain. If

$k_2 \leq 1$ the procedure will stop. If $k_2 \geq 2$ then for $m=N_2+1, \dots, N_3$ a sample of one is taken, independently of all previous samples, from each at the remaining k_2 populations. $R_{(1)}$ is performed at each stage and $Y_{i_1 m}, Y_{i_2 m}, \dots, Y_{i_{k_2} m}$ are defined to be 0 or 1 as a

failure or a success is obtained from $\Pi_{i_1}, \dots, \Pi_{i_{k_2}}$ under $R_{(1)}$,

where $\Pi_{i_1}, \dots, \Pi_{i_{k_2}}$ are the remaining populations yet undecided upon.

The procedure continues in this manner, at each stage sampling only from those populations yet undecided upon until it terminates at the first stage N_j when $k_j \leq 1$.

It is clear from the nature of the sampling, that for each $i=1, 2, \dots, k$ $Y_{i1}, Y_{i2}, \dots, Y_{im}$ are independent random variables taking the value 0 or 1, and so for each $i=1, 2, \dots, k$ define $T_{im} = \sum_{j=1}^m Y_{ij}$.

Then for each $\underline{c} = (\{a_m\}, \{b_m\})$ where

$$(i) \quad a_1 < 0 < a_m \leq a_{m+1}, \quad b_m \leq b_{m+1}, \quad m \geq 1$$

$$(2.11) \quad (ii) \quad \lim_{m \rightarrow \infty} a_m = \lim_{m \rightarrow \infty} b_m = \infty$$

$$(iii) \quad P\left\{ \bigcap_{v=1}^{\infty} [a_v < T_{iv} < b_v] \right\} = 0, \quad i=1, 2, \dots, k$$

the following procedure is defined.

\mathcal{L}_2 : For each $i=1,2,\dots,k$ stop sampling from Π_i at the first stage $m \geq 1$ when $T_{im} \notin (a_m, b_m)$. Accept Π_i if $T_{im} > b_m$ and reject Π_i if $T_{im} \leq a_m$. If at any stage only one population remains undecided upon accept that population.

T_{im} corresponds to the test statistic S_{im} used for the procedure \mathcal{L}_1 , but while it is clear that S_{im} is a binomial random variables with parameters p_i and m , no such statement can be made about T_{im} . This is because the distribution of Y_{im} $i=1,2,\dots,k$ is unknown for $m > N_1$. Define now for each $i=1,2,\dots,k$ and $m \geq 1$ $p_{im} = P\{Y_{im} = 1\}$. Then for $m=1,\dots,N_1$ and all $i=1,2,\dots,k$ $p_{im} = p_i$ as defined in (1.1.1), since the two procedures \mathcal{L}_1 and \mathcal{L}_2 are identical up to stage N_1 . However for $m > N_1$, following (1.1.1) with $n=1$,

$$(2.1.2) \quad p_{im} = \int_{-\infty}^{\infty} \prod'_{\substack{j=1 \\ j \neq i}}^k \Phi(x+d+(\theta_{[i]} - \theta_{[j]}) \frac{1}{\sigma}) \varphi(x) dx, \quad i=1,2,\dots,k.$$

where Π' refers to the fact that there are less than $k-1$ factors $\Phi(x+d+(\theta_{[i]} - \theta_{[j]}) \frac{1}{\sigma})$ but how many and which ones remaining is unknown. This depends on which populations are accepted or rejected at N_1, N_2, \dots, N_{k-1} . Clearly since populations are accepted or rejected only at stages N_j , $j=1,2,\dots,k-1$ then for $i=1,2,\dots,k$, with $N_0 \equiv 0$,

$$(2.1.3) \quad p_{im} = p_{iN_{j+1}} \quad \text{for } m=N_j+1, \dots, N_{j+1} \quad \text{and } j=0, \dots, k-2.$$

It is also clear that as m goes from N_j to N_{j+1} some of the factors $\Phi(x+d+(\theta_{[i]} - \theta_{[j]}) \frac{1}{\sigma})$ are replaced by 1 so that for $i=1,2,\dots,k$,

$$(2.1.4) \quad p_i \leq p_{im} \leq p_{im+1}, \quad m \geq 1,$$

and in particular if $N_j < N$

$$(2.1.5) \quad p_i = p_{iN_1} < p_{iN_2} < \dots < p_{iN_j}.$$

Since at N_j , $j=1,2,\dots,k-1$ at least one population must be either accepted or rejected we have the following lower bounds for $i=1,2,\dots,k$,

$$(2.1.6) \quad \int_{-\infty}^{\infty} \prod_{\substack{\ell=j \\ \ell \neq i}}^k \Phi(x+d+(\theta_{[i]} - \theta_{[\ell]}) \frac{1}{\sigma}) \varphi(x) dx \leq p_{iN_j} \quad j=1,\dots,k-1.$$

Since the distribution of T_{im} is in general unknown to us, it can be expected that much less information can be derived about the procedure \mathcal{L}_2 . Define for each $i=1,2,\dots,k$,

$$(2.1.7) \quad r_i^! = P\{\text{accepting } \Pi_{(i)} | \mathcal{L}_2(\underline{c})\}.$$

In general then,

$$(2.1.8) \quad r'_k = P\{cs | \mathcal{L}_2(\underline{c})\} = \sum_{m=1}^{\infty} P\left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(k)v} < b_v] \cap [T_{(k)m} \geq b_m] \right\}.$$

If as in Chapter I we let M_i be the stage at which population $\Pi_{(i)}$ is accepted or rejected then (2.1.1) guarantees that $E\{M_i | \mathcal{L}_2(\underline{c})\} < \infty$ for all $i=1,2,\dots,k$, so that,

$$(2.1.9) \quad P\{\text{accept } \Pi_i | \mathcal{L}_2(\underline{c})\} + P\{\text{reject } \Pi_i | \mathcal{L}_2(\underline{c})\} = 1.$$

2.2. Some Properties of $\mathcal{L}_2(\underline{c})$

Consider two generic populations Π and Π' . Let $\{Y_m, m \geq 1\}$ and $\{Y'_m, m \geq 1\}$ be two sequences of independent random variables generated by Π and Π' respectively as shown in Section 2.1.

Suppose $p_m = P\{Y_m = 1\}$ and $p'_m = P\{Y'_m = 1\}$ and that $p_m \leq p'_m$

for $m \geq 1$. Finally set $T_m = \sum_{j=1}^m Y_j$ and $T'_m = \sum_{j=1}^m Y'_j$, and consider

the following lemma.

Lemma (2.2.1). There exists a sequence of independent random variables $\{U_m, m \geq 1\}$ such that,

(i) $P\{U_m \leq u\} = P\{Y_m \leq u\}$ for all u , $m \geq 1$, and

(ii) $P\{U_m \leq Y'_m\} = 1$ for $m \geq 1$.

Proof. Consider a sequence of independent random variables

$\{Z_m, m \geq 1\}$ where Z_m is independent of Y'_n for $m \neq n$ and
 $P\{Z_m = 1 \mid Y'_m = 1\} = \frac{p_m}{p'_m}$, $P\{Z_m = 0 \mid Y'_m = 0\} = 1$ for all $m \geq 1$.

Then as in Lemma (1.2.2) if we set $U_m = Z_m Y'_m$ for all $m \geq 1$ the lemma follows.

Theorem (2.2.1). $r'_1 \leq r'_2 \leq \dots \leq r'_k$.

Proof. It will suffice to prove $r'_1 \leq r'_2$. For the first N_1 stages of the procedure, since no populations have been accepted or rejected, $P\{Y_{(1)m} = 1\} = p_1$ and $P\{Y_{(2)m} = 1\} = p_2$ for any $m=1,2,\dots,N_1$ where p_1 and p_2 are defined by (1.1.1) with $n=1$. It was shown in Gupta [22] that $p_1 \leq p_2$. For $m > N_1$ from (2.1.2)

$$P\{Y_{(1)m} = 1\} = p_{1m} = \int_{-\infty}^{\infty} \prod_{j=2}^k \Phi(x+d+(\theta_{[1]} - \theta_{[j]})\frac{1}{\sigma}) \varphi(x) dx, \text{ and}$$

$$P\{Y_{(2)m} = 1\} = p_{2m} = \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j \neq 2}}^k \Phi(x+d+(\theta_{[2]} - \theta_{[j]})\frac{1}{\sigma}) \varphi(x) dx. \text{ Then at}$$

any stage m before which neither $\Pi_{(1)}$ or $\Pi_{(2)}$ have been accepted

or rejected since $\theta_{[1]} \leq \theta_{[2]}$ it is clear that $p_{1j} \leq p_{2j}$ $j \leq m$.

Now Lemma (2.2.1) allows us to consider the sequence $\{Y_{(1)m}\}$ as

bounded above by the sequence $\{Y_{(2)m}\}$, i.e. $P\{Y_{(1)m} \leq Y_{(2)m}\} = 1$

for all $m \geq 1$. Thus the event $\left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(1)v} < b_v] \cap [T_{(1)m} \geq b_m] \right\}$

implies that there exists an $n=1,2,\dots,m$ such that the event

$\left\{ \bigcap_{v=1}^{n-1} [a_v < T_{(2)v} < b_v] \cap [T_{(2)n} \geq b_n] \right\}$ holds. Therefore it follows

that

$$\left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(1)v} < b_v] \cap [T_m \geq b_m] \right\} \subset \bigcup_{n=1}^m \left\{ \bigcap_{v=1}^{n-1} [a_v < T_{(2)v} < b_v] \cap [T_{(2)n} \geq b_n] \right\}.$$

Then taking the union of these events for $m \geq 1$, and noting that they are all disjoint events, we have,

$$r_1' = \sum_{m=1}^{\infty} P \left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(1)v} < b_v] \cap [T_{(1)m} \geq b_m] \right\} \leq \sum_{m=1}^{\infty} P \left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(2)v} < b_v] \cap [T_{(2)m} \geq b_m] \right\} = r_2'.$$

This completes the proof.

Corollary (2.2.1). The procedure $\mathcal{L}_2(\underline{c})$ is unbiased.

Proof. It follows from Theorem (2.2.1) and (2.1.9) that the probability of rejecting $\Pi_{(k)}$ is less than or equal to the probability of rejecting any other population. Thus $\mathcal{L}_2(\underline{c})$ is unbiased.

The next theorem corresponds to Theorem (1.2.1) which showed the

monotonicity of procedure $\mathcal{L}_1(\underline{c})$ with respect to the ordered pairs $\underline{c}, \underline{c}'$ satisfying (1.1.2). Here we consider ordered pairs $\underline{c}, \underline{c}'$ which satisfy (2.1.1). The proof of the following theorem follows from Lemma (1.2.1) with the sequence $\{S_m, m \geq 1\}$ replaced by $\{T_m, m \geq 1\}$.

Theorem (2.2.2). If $\underline{c}, \underline{c}'$ satisfy (2.1.1) and $\underline{c} \leq \underline{c}'$ then,

- (i) $P\{\text{selecting } \Pi_i | \mathcal{L}_2(\underline{c})\} \geq P\{\text{selecting } \Pi_i | \mathcal{L}_2(\underline{c}')\}$ and
- (ii) $P\{\text{rejecting } \Pi_i | \mathcal{L}_2(\underline{c})\} \leq P\{\text{rejecting } \Pi_i | \mathcal{L}_2(\underline{c}')\}$.

In particular $P\{cs | \mathcal{L}_2(\underline{c})\} \geq P\{cs | \mathcal{L}_2(\underline{c}')\}$.

Proof. Apply Lemma (1.2.1) to $\{T_{im}, m \geq 1\}$ $c=1,2,\dots,k$, and note that $P\{\text{selecting } \Pi_i | \mathcal{L}_2(\underline{c})\} = p_{\underline{c}}(A)$ as defined in the lemma. This will demonstrate (i) above, and (ii) follows from (i) and (2.1.9). Finally apply (i) to $\Pi_{(k)}$ for the last result which completes the proof.

Thus the behavior of \mathcal{L}_2 over an ordered class C_2 where $\underline{c} \in C_2$ satisfies (2.1.1) is the same as \mathcal{L}_1 over an ordered class C_1 where $\underline{c} \in C_1$ satisfies (1.1.2).

We would now like to compare procedures $\mathcal{L}_1(\underline{c})$ and $\mathcal{L}_2(\underline{c})$ for any \underline{c} satisfying both (1.1.2) and (2.1.1). The existence of such \underline{c} will be demonstrated for $\underline{c} = (\{cm-d_1\}, \{cm+d_2\})$ where $c \in (0,1)$ and is rational, in the last lemma of this section. The main

theorem of this section compares $r_i = P\{\text{selecting } \Pi_{(i)} | \mathcal{A}_1(\underline{c})\}$ and $r'_i = P\{\text{selecting } \Pi_{(i)} | \mathcal{A}_2(\underline{c})\}$. Consider any generic population Π . Under $\mathcal{A}_1(\underline{c})$ Π generates a sequence of independent, identically distributed random variables taking the values 0 or 1 as described in Section 1.1; and under $\mathcal{A}_2(\underline{c})$ Π generates a sequence of independent random variables taking the values 0 or 1 as described in Section 2.1. If we let $\{s_m, m \geq 1\}$ be the sequence at partial sums of the variables generated by $\mathcal{A}_1(\underline{c})$ and $\{T_m, m \geq 1\}$ be the sequence of partial sums of the variables generated by $\mathcal{A}_2(\underline{c})$ then from (2.1.4) and lemma (2.2.1) with $p_m = p, m \geq 1$ it follows that we can consider S_m dominated by T_m by $P\{S_m \leq T_m\} = 1$ for all $m \geq 1$.

Following the notation of Section (2.1) define, $A_m(\underline{c}) =$

$$\left\{ \bigcap_{v=1}^{m-1} [a_v < S_v < b_v] \cap [s_m \geq b_m] \right\}, B_m(\underline{c}) = \left\{ \bigcap_{v=1}^{m-1} [a_v < T_v < b_v] \cap [T_m \geq b_m] \right\},$$

$$P_{\underline{c}}(A) = \sum_{m=1}^{\infty} P(A_m(\underline{c})), \text{ and } P_{\underline{c}}(B) = \sum_{m=1}^{\infty} P(B_m(\underline{c})). \text{ Consider the following}$$

lemma.

Lemma (2.2.2). $P_{\underline{c}}(A) \leq P_{\underline{c}}(B)$ for all \underline{c} satisfying (1.1.2) and (2.2.1).

Proof. From the above discussion since $P\{S_m \leq T_m\} = 1$ for all $m \geq 1$

clearly $[S_m \geq b_m] \subseteq [T_m \geq b_m]$. The event $\bigcap_{v=1}^{m-1} [a_v < S_v < b_v]$ implies

either $\bigcap_{v=1}^{m-1} [a_v < T_v < b_v]$ or there exists an $n=1, \dots, m-1$ such that

the event $\bigcap_{v=1}^{n-1} [a_v < T_v < b_v] \cap [b_n \geq b_n]$ holds. Thus $A_m(\underline{c}) \subseteq \bigcup_{n=1}^m B_n(\underline{c})$.

Therefore $P(\bigcup_{m=1}^{\infty} A_m(\underline{c})) \leq P(\bigcup_{m=1}^{\infty} B_m(\underline{c}))$ but clearly $A_m(\underline{c}) \cap A_n(\underline{c}) = \emptyset$

and $B_m(\underline{c}) \cap B_n(\underline{c}) = \emptyset$ if $m \neq n$. Thus $P_{\underline{c}}(A) \leq P_{\underline{c}}(B)$ and the proof

is complete.

Theorem (2.2.3). For all $i=1,2,\dots,k$, $r_i \leq r'_i$.

Proof. The proof follows immediately from Lemma (2.2.2) since

$$r_i = \sum_{m=1}^{\infty} P\left\{ \bigcap_{v=1}^{m-1} [a_v < s_{(i)v} < b_v] \cap [s_{(i)v} \geq b_m] \right\}, \text{ and}$$

$$r'_i = \sum_{m=1}^{\infty} P\left\{ \bigcap_{v=1}^{m-1} [a_v < T_{(i)v} < b_v] \cap [T_{(i)m} > b_m] \right\}. \text{ Thus apply Lemma (2.2.2)}$$

with $\Pi = \Pi(i)$.

Lemma (2.2.3). For $\underline{c} = (\{cm-d_1\}, \{cm+d_2\})$ where $c \in (0,1)$ is

rational and d_1, d_2 are positive, $P\{N < \infty\} = 1$.

Proof. As shown in Section 1.3 we can convert the problem of the first crossing of the sequence $\{T_{im}, m \geq 1\}$, $i=1,2,\dots,k$ out of

the interval $(cm-d_1, cm+d_2)$ to that of a one-dimensional random

walk $R'_{im} = \sum_{i=1}^m Y'_i$ on the fixed interval $(-sd_1, sd_2)$ where $c=r/s$

and $P\{Y'_{im} = s-r\} = p_{im}$, $P\{Y'_{im} = -r\} = 1-p_{im}$. Now until N_1 , the

first stage at which one or more populations are accepted or rejected

$p_{im} = p_i$, $i=1,2,\dots,k$ and the procedures \mathcal{L}'_2 and \mathcal{L}'_1 coincide.

However it was shown in Section 1.3 and Section 1.4 that for each

$i=1,2,\dots,k$ $E\{M_i | \mathcal{L}'_1(c)\} < \infty$ where M_i is the first stage Π_i is

tagged, therefore $P(N_1 < \infty) = 1$. From stage N_1 until stage N_2

for those population left we have $P_{i,j,m} = P_{i,j,N_2}$ for $j=1,2,\dots,k_1$.

Here again we have a random walk on $(-sd_1, sd_2)$ for each of the k_1 remaining populations. The starting point of each of the random walks is some point in $(-sd_1, sd_2)$. However, it is well-known (see Cox and Miller [8], Chapter I) that for such random walks the probability of leaving a finite interval in a finite time is one regardless of the starting point, therefore $P\{N_2 < \infty\} = 1$. Since there are at most $k-1$ stages at which populations are tagged then the same argument will show that $P\{N_j < \infty\} = 1$ $j=3,4,\dots,k-1$. But $N \leq N_{k-1}$, hence $P\{N < \infty\} = 1$ and the proof of the lemma is complete.

2.3. Some Monte Carlo Results for the
Performance of $\mathcal{L}_2(\underline{c})$ With $\underline{c} \in C_1^1$
for the Slippage Problem

Since no exact expressions for $P\{\text{selecting } \Pi_1 | \mathcal{L}_2(\underline{c})\}$ and $E\{m_1 | \mathcal{L}_2(\underline{c})\}$ can be obtained, a Monte-Carlo simulation was carried out for five normal populations with common variance $\sigma^2 = 1$, and certain slippage configurations of the means. A power residue method was used to generate a sequence of pseudo random numbers on $(0,1)$ (see System Bulletin No. 2, Purdue University, Computer Science Department, December 1967.) These uniform random variables were then converted to normal random variables with mean 0 and variance 1 by application of the central limit theorem. They were then adjusted for slippage of the mean. Five and ten thousand simulations of $\mathcal{L}_2(\underline{c})$ were then carried

out and the number of times a population was accepted was divided by the number of simulations, which gives an indication of the true probability of acceptance using $\mathcal{L}_2(\underline{c})$.

The results were not conclusive and in only one case was there any saving in the number of samples needed in using $\mathcal{L}_2(\underline{c})$ instead of $R(n)$. However, a trend in the behavior of the probabilities and expectations with respect to c and d indicates that there may be savings in many cases. The first problem is the choice of $\underline{c} = (\{cm-d\}, \{cm+d\})$ with $c \in (0,1)$, rational and $d > 0$. In Section 2.2 it is shown (Theorem 2.2.3) that $r_i < r'_i$, and in Section 1.5 it is shown that by using c^*, d^* then $r_k \geq 1-\epsilon$, $r_i \leq \epsilon$, $i=1,2,\dots,k-1$ and the approximate minimum of (1.5.3) is attained. Hence for c^*, d^* we have $P\{cs | \mathcal{L}_2(c^*)\} \geq 1-\epsilon$. With $\epsilon = .001$, the Monte-Carlo technique was applied, and the results appear in Table 8, where the probability of selection using $\mathcal{L}_1(c^*)$ given by

TABLE 8

Monte-Carlo Results for the Slippage Problem Using \underline{c}^* , with $P^* = .75$

Configuration of the Means

<u>Population</u>	<u>(0,0,0,0, .4)</u>		<u>(0,0,0,0,1)</u>	
$\Pi_{(5)}$	1.000	(1.000) 79.6	1.000	(1.000) 14.9
$\Pi_{(i)}, i=1,2,3,4$.765	(.146) 79.3	.626	(.089) 14.0

the value on the left, and the expected number of samples using $\mathcal{L}_1(c^*)$

the value on the right. The number in parenthesis gives the value of

$R(n)$ with n approximately equal to the expected number of samples

using $\mathcal{L}_2(\underline{c}^*)$. It is clear that the procedure $\mathcal{L}_2(\underline{c}^*)$ accepts poor populations with a large probability. One of the reasons for this can be seen in equation (2.12), where p_{im} is given as a function of the number of populations left after stage $m-1$. It appears that as k decreases p_{im} increases quickly so that for five populations p_{im} becomes large at a very fast rate. It may be that for k large the rate at which p_{im} increases would be slow so that $\mathcal{L}_2(\underline{c}^*)$ might be a good procedure for $k \geq 25$.

From the definition of $\mathcal{L}_2(\underline{c})$, if at any stage only one population remains, that population is accepted. It would then seem reasonable to try and insure that Π_k has the largest expected number of stages until absorption so as not to be forced to accept a "poor" population. From Table A3 in the appendix it can be seen that for a fixed c , the expected number of stages until absorption is maximized for $p=c$. Also since the family of pairs of sequences C_1^i is ordered, we can make the initial p_{im} , $m \leq N_1$ smaller by increasing c . Thus the choice $c=p_k$ appears to be a reasonable choice to make. For $p^* = .75$ Table 9 gives results similar to Table 8 for a Monte-Carlo simulation for the configuration $(0,0,0,0,2)$ with $c=p_k$ and several values of d .

TABLE 9

Monte-Carlo Results for the Slippage Configuration $(0,0,0,0,2)$

Using $\mathcal{L}_2(p_k)$ with $p^* = .75$.

Pop.	$d = 0.5$	$d = 1.0$	$d = 2.0$	$d = 3.0$
$\Pi(5)$.994(.996) 2.9	1.000(1.000) 4.6	1.000(1.000) 6.2	1.000(1.000) 7.8
$\Pi(i) \ i < 5$.130(.305) 1.8	.210 (.180) 3.3	.266 (.026) 4.9	.311 (.005) 6.6

The values to the right give $\mathcal{L}_2(p_k)$ ($R(n)$) selection probabilities and the value to the left gives the expected number of stages until the population is tagged using $\mathcal{L}_2(p_k)$. n is taken approximately equal to that value.

It thus appears that for small enough d ($d = .5$) there exists in this case a procedure $\mathcal{L}_2(\underline{c})$ which does a bit better than $R(n)$. This was done for relatively large δ and in Chapter I it was shown that $\mathcal{L}_1(\underline{c})$ improved on $R(n)$ for very small δ , thus $\mathcal{L}_2(\underline{c})$ may improve with smaller δ . Table 10 shows the same results as Table 9 for $\delta = 1$. It is clear that d has not been chosen small enough, to obtain favorable results for $\mathcal{L}_2(p_k)$, if such results exist. However the monotonic behavior of the various probabilities is consistent with Table 9

TABLE 10

Monte-Carlo Results for the Slippage Configuration (0,0,0,0,1)

Using $\mathcal{L}_2(p_k)$ with $p^* = .75$

Pop.	$d = 2$	$d = 3$	$d = 4$
$\Pi(5)$.993(.999) 8.1	.998(1.000) 10.4	1.000(1.000) 12.9
$\Pi(i) \ i < 5$.445(.230) 6.9	.487 (.190) 9.3	.522 (.123) 11.8

CHAPTER III

SOME GENERALIZATIONS AND RELATED PROBLEMS

3.1 Generalizations of the Sequential Procedures \mathcal{L}'_1 and \mathcal{L}'_2 .

Let $\Pi_1, \Pi_2, \dots, \Pi_k$ be k populations, each of which has an observable random variable X_i whose distribution $f(x_i; \theta_i)$ depends on a parameter θ_i . The problem stated in the introduction of this thesis is to define sequential procedures which select a small subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ containing the population whose distribution has the largest (or smallest) parameter value, with a prescribed probability P^* , $\frac{1}{k} < P^* < 1$. Chapters I and II define sequential procedures \mathcal{L}'_1 and \mathcal{L}'_2 based on a fixed sample-size procedure $R(n)$, defined in Section 1.1, for the case where X_i , $i=1, 2, \dots, k$ is a normal random variable with mean θ_i and known variance σ^2 . In this section the extension of \mathcal{L}'_1 and \mathcal{L}'_2 to a wider class of distributions will be discussed. In addition, a generalization of \mathcal{L}'_1 and \mathcal{L}'_2 will be proposed. The notation of this section will be consistent with Chapters I and II.

As in Chapter I, after taking a sample of one from each of the k populations and performing the fixed sample-size procedure $R(1)$, we denote a success from population Π_i as the event,

$$(3.1.1) \quad x_i \geq x \max - \sigma d$$

where $x \max = \max (x_1, x_2, \dots, x_k)$ and d is a suitably chosen constant. The decision to accept or reject a population Π_i using \mathcal{L}_1 is then based on the statistic S_{im} , the number of success with Π_i in m independent trials using $R(1)$. The evaluation of certain necessary probabilities and expectations concerning the procedure \mathcal{L}_1 depends on knowing $p_i, i=1,2,\dots,k$, the probability of a success with Π_i . Thus given p_1, p_2, \dots, p_k , the procedure \mathcal{L}_1 is essentially selecting and ranking k binomial populations by means of $S_{1m}, S_{2m}, \dots, S_{km}$. Therefore the extension of \mathcal{L}_1 to a larger class of distribution functions depends on the extension of fixed sample-size procedures $R(n)$ to those distributions. A discussion of such extensions can be found in Gupta [23].

Gupta defines decision rules based on functions $h_b(x)$, where for all $b \in [0, \infty]$ (or $[1, \infty]$) and every x ,

$$(a) \quad h_b(x) \geq x$$

$$(b) \quad h_0(x) = x \quad (\text{or } h_1(x) = x)$$

$$(c) \quad \lim_{b \rightarrow \infty} h_b(x) = \infty$$

$$(d) \quad h_b(x) \text{ is continuous and monotone increasing in } b.$$

Then based on observations x_i from Π_i he defines procedures R_{h_b}

as follows.

R_{n_b} : Select Π_i if and only if $h_b(x_i) \geq x_{\max}$ where b is chosen such that $\inf_{\Omega} P\{cs|R_{n_b}\} \geq P^*$ and Ω is the parameter space.

Given that $f(x;\theta)$ is stochastically increasing in θ , i.e., $F(x;\theta_2) \leq F(x;\theta_1)$ for all $\theta_2 > \theta_1$ where $F(x;\theta) = \int_{-\infty}^x f(t;\theta)dt$ under certain other conditions Gupta showed that it is possible to find a unique b such that $\inf_{\Omega} P\{cs|R_{n_b}\} = P^*$. The procedure R_{n_b} also has the following monotonic property. If $\theta_i \geq \theta_j$; then,

$$(3.1.2) \quad P\{\text{selecting the population with parameter } \theta_i\} \geq$$

$$P\{\text{selecting the population with parameter } \theta_j\}.$$

This monotonic ordering of p_1, p_2, \dots, p_k is just the property needed to prove most of the lemmas of Section 1.2, so that the procedure \mathcal{L}_1 will retain such properties as monotonicity and unbiasedness when extended to the wider class of distributions.

Two important examples occur when θ is a location or scale parameter. In these cases it is clear that $f(x;\theta)$ is stochastically increasing. It has been shown by Gupta [15], [22], [23] and others that the $\inf_{\Omega} P\{cs|R_{n_b}\}$ occurs when $\theta_1 = \theta_2 = \dots = \theta_k = \theta$ and is independent of the value of θ .

For the location parameter when $f(x;\theta) = f(x-\theta)$, it is shown in [23] that $h_b(x) = x+b$, $b \in [0, \infty)$. An example of this type of distribution is the normal distribution with known variance. Here $R_{h_b} = R(n)$ as defined in Section 1.1 with $b = \frac{\sigma d}{\sqrt{n}}$.

For the scale parameter $f(x;\theta) = \frac{1}{\theta} f(\frac{x}{\theta})$ and $h_b(x) = bx$, $b \in [1, \infty)$. An example of this type of distribution is the gamma distribution. Based on a sample of n independent observations from each population Gupta in [20] defines $R_{h_b} = b_{h_b}(n)$ as the procedure which selects Π_i if and only if $\bar{x}_i \geq b \bar{x}_{\max}$ the $0 < b \leq 1$, where b is chosen so that $\inf_{\Omega} P\{cs | R_{h_b}(n)\} = P^*$.

There are other types of univariate distributions for which fixed sample-size procedures exist to which \mathcal{L}_1 is applicable. An example given in 3.2 is the binomial distribution. Extensions to multivariate distributions are also possible, examples are given in Gupta [23] and Gupta and Panchapakesan [24] of fixed sample-size rules for which the sequential procedure \mathcal{L}_1 can be defined. From the definition of \mathcal{L}_2 , given in Section 2.1, it is clear that for any application of \mathcal{L}_1 there is a similar application for \mathcal{L}_2 .

A slight generalization of \mathcal{L}_1 and \mathcal{L}_2 will now be proposed. Often it is not feasible to take only one observation at a time from a given population. For example, the fixed sample-size rule we are applying \mathcal{L}_1 (or \mathcal{L}_2) to may require a minimum of more than one sample. For example when selecting the smallest variance of k

normal populations with unknown mean (see Gupta and Sobel [18]) the

fixed sample size rule uses the statistics $\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$,

$i=1,2,\dots,k$ which clearly requires a minimum of $n = 2$. Another case where it is necessary to take more than one observation at a time occurs for discrete distributions. Here the P^* condition can not, in general, be met exactly, so that the fixed sample-size rule constant is chosen so that $\inf_{\Omega} R\{cs|R(n)\} \geq P^*$. For large P^* the

fixed sample size rule that takes just one observation from each population may be forced to accept all the populations to meet the P^* condition. Clearly if $p_1=p_2=\dots=p_k = 1$ the procedure \mathcal{L}_1 will be useless, hence it is necessary to take 2 or more observations from each population, at each stage. An example of this will be studied for the binomial distribution in Section 3.2.

Thus to generalize \mathcal{L}_1 , at the first stage take an independent sample of size $\ell \geq 1$ from each of the k populations and perform $R(\ell)$. For each $i=1,2,\dots,k$ define $Y_{i1} = 0$ or 1 as a failure (Π_i is rejected by $R(\ell)$) or a success (Π_i is accepted by $R(\ell)$) is obtained with population Π_i . At stage 2 take an additional sample of ℓ from each population, independently of the first sample, perform $R(\ell)$ and define $Y_{i2} = 0$ or 1 as a failure or success is gotten with Π_i on the second sample. Continuing in this manner at the m th stage take a sample of ℓ from each population, independently of all previous samples, perform $R(\ell)$ and define $Y_{i\ell} = 0$ or 1 as a failure or a success is gotten from Π_i from the m th sample. Then

let $S_{im} = \sum_{j=1}^m Y_{ij}$ for $i=1,2,\dots,k$ and given suitable $\underline{c} = (\{a_m\}, \{b_m\})$ perform $\mathcal{L}_1(\underline{c})$ as defined in Section 1.1. The generalization of $\mathcal{L}_2(\underline{c})$ is similar to that of $\mathcal{L}_1(\underline{c})$ but at any stage you only sample from these populations yet undecided upon.

3.2 Some Sample Size Comparisons

for $\mathcal{L}_1(c^*)$ and $R(n)$ for Binomial Populations

In this section we apply the sequential procedure $\mathcal{L}_1(c^*)$ to a fixed sample size rule, given by Gupta and Sobel [17], for selecting a subset containing the best of several binomial random variables. So let $\Pi_1, \Pi_2, \dots, \Pi_k$ denote k binomial populations, and let γ_i be the probability of a "success" with population Π_i , $i=1,2,\dots,k$. Assume $\gamma_1, \gamma_2, \dots, \gamma_k$ are known but that the correct pairing of $\gamma_{[i]}$ to $\Pi_{(i)}$ is unknown for each i . Now define x_i to be the observed number of successes based on n observations from Π_i , $i=1,2,\dots,k$, and let $x_{\max} = \max_{1 \leq i \leq k} x_i$. Then the following fixed sample procedure is defined in [17].

$R(n)$: Select Π_i if and only if $x_i \geq x_{\max} - d$, where d is chosen as the smallest integer such that $\inf_{\Omega} P\{cs|R(n)\} \geq P^*$ and Ω is the parameter space of the γ_i 's.

Comparisons made here will be for the slippage configuration, where $\gamma_{[1]} = \gamma_{[2]} = \dots = \gamma_{[k-1]} = \gamma$, $\gamma_{[k]} = \gamma + \delta$, $\delta > 0$. Then as shown in [17],

$$(3.2.1) \quad P\{cs|R(n)\} = \sum_{x=0}^n \binom{n}{x} (\gamma+\delta)^x (1-(\gamma+\delta))^{n-x} \left\{ \sum_{j=0}^{x+d} \binom{n}{j} \gamma^j (1-\gamma)^{n-j} \right\}^{k-1}$$

and

$$(3.2.2) \quad P\{\text{selecting } \Pi_{(i)} | R(n)\} = \sum_{x=0}^n \binom{n}{x} \gamma^x (1-\gamma)^{n-x} \left\{ \sum_{j=0}^{x+d} \binom{n}{j} \gamma^j (1-\gamma)^{n-j} \right\}^{k-2} \\ \left\{ \sum_{j=0}^{x+d} \binom{n}{j} (\gamma+\delta)^j (1-(\gamma+\delta))^{n-j} \right\}.$$

For $k = 2$, and $P^* = .75$ a comparison was made between $R(n)$ defined above and $c_1^{\ell}(c^*)$ as defined Section 3.1 for $\ell=1$. The results appear in Table 11 below. For $k=2$, and $P^* = .90$ the procedure $R(1)$ chooses all populations since in this case $d=1$. Thus a comparison was made between $R(n)$ and $c_1^{\ell}(c^*)$ for $\ell=2$. That is at each stage of $c_1^{\ell}(c^*)$ two observations from each population were taken. The results appear in Table 12.

TABLE 11

Sample Size Comparisons for the Sequential and
Fixed Sample-Size Rules for the Slippage Configuration,
for the Binomial Population: $P^* = .75$

$\gamma+\delta$	$\delta = .25$		$\delta = .50$	
.75	43.8	105.2	11.9	27.0
.80	42.3	103.1	11.8	25.2
.85	39.8	95.0	11.7	23.6
.90	36.6	85.6	11.5	22.0
.95	33.3	75.2	11.5	21.0
.99	26.4	55.1	10.3	17.1

In Table 11 the value to the left is the expected number of samples using $\mathcal{L}_1(c^*)$ so that (1.7.3) with $\epsilon = .001$ is satisfied and the value to the right is the fixed-sample size n necessary so that $R(n)$ satisfies (1.7.3) with $\epsilon = .001$. It can be seen that $\mathcal{L}_1(c^*)$ saves samples in every case shown but as δ increases the savings become less. Table 12 gives a similar result for $P^* = .90$, though here $R(n)$ appears the better procedure. Both tables indicate that for small δ the sequential test would produce more savings in the number of samples needed.

TABLE 12

Sample Size Comparisons for the Sequential and
Fixed Sample-Size Rules for the Slippage Configuration,
for the Binomial Population: $P^* = .90$

$\gamma + \delta$	$\delta = .5$		$\delta = .50$	
.75	139.6	140.0	39.2	34.0
.80	142.2	135.6	40.0	32.6
.85	148.2	128.3	41.2	31.3
.90	148.1	115.3	41.4	30.5
.95	142.7	98.0	40.2	28.0
.99	140.0	94.3	38.0	25.0

3.3 Some Related Problems

A: A Sequential Multiple Decision Rule for
Selecting the Best of k Binomial Populations.

Consider as in Section 3.2 k population $\Pi_1, \Pi_2, \dots, \Pi_k$ where each Π_i , $i=1,2,\dots,k$ has an observable random variable x_i with probability of success γ_i . We will suppose that the γ_i 's are known but that the correct pairing of $\gamma_{[i]}$ to $\Pi_{[i]}$ is unknown

for all $i=1,2,\dots,k$. This was not assumed by Gupta and Sobel. The problem of this section as in previous sections is to find sequential procedures which select a subset of $\Pi_1, \Pi_2, \dots, \Pi_k$ which includes $\Pi_{(k)}$ with a prescribed probability P^* . The previous section applied the procedure $c_1^f(c^*)$ to a fixed sample-size rule found in Gupta and Sobel [17]. However since we are involved with binomial populations, and the procedure c_1^f essentially converts all selection procedures, to which it applies, to selection of binomial populations; it is therefore reasonable to consider the following procedure, which makes use of the fact that the underlying distributions are already binomial.

Let S_{im} denote the number of successes from population Π_i $i=1,2,\dots,k$ in m independent trials. Then for suitable $\underline{c} = (\{a_m\}, \{b_m\})$ perform the following procedure.

- c : Stop sampling from $\Pi_i, i=1,2,\dots,k$, at the first stage $S_{im} \notin (a_m, b_m)$. Accept Π_i if $S_{im} \geq b_m$ and reject Π_i if $S_{im} \leq a_m$.

Note that the populations are independent and so therefore are $S_{1m}, S_{2m}, \dots, S_{km}$. Thus there is no need to sample from each population at each stage. It is clear that since the γ_i 's are ordered the procedure c has all the properties of c_1^f proved in Section (1.2), and for $\underline{c} \in C_1'$ all the properties of c_1^f proved in Sections 1.3, 1.4 and 1.5. In particular for $\underline{c} = \underline{c}^*$ as defined in (1.5.7) and (1.5.10) we have for any $\epsilon > 0$,

- (1) $P\{cs|c\} \geq 1 - \epsilon$
- (2) $1 - \epsilon < ES < 1 + (k-1) \epsilon$

as well as approximately minimizing $\max_{1 < i < k} E\{M_i | \mathcal{L}\}$ over all

$c \in (\gamma_{k-1}, \gamma_k)$, where M_i is the stage at which Π_i is accepted or rejected.

B. Selection of Populations Better than a Standard

The problem defined in Section 1.1, is for $\Pi_1, \Pi_2, \dots, \Pi_k$, normal random variables with mean θ_i respectively and a common known variance $\sigma^2 = 1$. In Section 1.1 it is assumed that p_1, p_2, \dots, p_k as defined in (1.1.1) are known, and procedures to select the best population are discussed. Suppose now that p_1, p_2, \dots, p_k are not known, but that the experimenter furnishes two constants p and Δ , in addition to P^* , such that $p, \Delta > 0$, and $p + \Delta < 1$. Now suppose the problem is to select all populations $\Pi_{(i)}$ such that $p_i \geq p + \Delta$, while rejecting these populations $\Pi_{(i)}$ where $p_i \leq p$. It is assumed that the experimenter has no preference as to selection or rejection of those $\Pi_{(i)}$ such that $p < p_i < p + \Delta$.

Then for any $\underline{c} \in C_1'$, that is $\underline{c} = (\{cm-d\}, \{cm+d\})$ with $c \in (p, p+\Delta)$ Theorem 1.4.4 guarantees that there exists a d such that for the procedure $\mathcal{L}_1(\underline{c})$ defined in Section (1.1), and any $\epsilon > 0$,

$$(1) \quad P\{\text{selecting } \Pi_{(i)} | \mathcal{L}_1(\underline{c})\} \geq 1 - \epsilon \quad \text{if } p_i \geq p + \Delta$$

$$(2) \quad P\{\text{selecting } \Pi_{(i)} | \mathcal{L}_1(\underline{c})\} \leq \epsilon \quad \text{if } p_i \leq p.$$

In particular, if $\bar{c} = p + \Delta/2$, $c^* = [\ln \frac{1-p}{1-p-\Delta}] [\ln \frac{p+\Delta}{p} \frac{1-p}{1-p-\Delta}]^{-1}$ and

c is chosen such that $c \in (\bar{c}, c^*)$ then as shown in Section 1.5 the approximate minimum of $\max_{1 \leq i \leq k} E\{M_i | \mathcal{L}_1(c)\}$ is attained, over all $c \in (p, p+\Delta)$.

BIBLIOGRAPHY

1. Bahadur, R.R. (1950). On the problem in the theory of k populations, Ann. Math. Statist. (21), 362-375.
2. Bechhofer, R.E. (1954). A single-sample multiple decision procedure for making means of normal populations with known variances. Ann. Math. Statist. (25), 16-29.
3. Bechhofer, R.E. and Sobel, M. (1954). A single sample decision procedure for ranking variances of normal populations. Ann. Math. Statist. (25), 273-289.
4. Bechhofer, R.E., Dunnett, C.W., and Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common known variance. Biometrika (41), 170-176.
5. Bechhofer, R.E. (1958). A sequential multiple decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs. Biometrics (14), 408-429.
6. Bechhofer, R.E. and Blumenthal, S. (1962). A sequential multiple-decision procedure selecting the best one of several normal population with a common unknown variance II: Monte Carlo sampling results and new computing formulae. Biometrics (18), 52-67.
7. Bechhofer, R.E., Kiefer, J., and Sobel, M., Sequential Identification and Ranking Procedures, University of Chicago Press, 1967.
8. Cox, D.R., and Miller, H.D., The Theory of Stochastic Processes, John Wiley and Sons, 1965.
9. Deely, J.J. (1965). Sequential selection of the best of k populations. Department of Statistics Mimeograph Series Purdue University #91.
10. Deely, J.J. and Gupta, S.S. (1968). On the properties of subset selection procedures. Sankhya Series A 1-14.
11. Dunnett, C.W. (1960). On selecting the largest of k normal population means. J. Roy. Stat. Soc. Ser. B (22), 1-40.
12. Feller, W., An Introduction to Probability Theory and Its Application Vol. 1, John Wiley and Sons, 2nd Edition, 1960.

13. Gantmacher, F.R., The Theory of Matrices Volume I, Chelsea Publishing Company, (1959).
14. Gantmacher, F.R., The Theory of Matrices Volume II, Chelsea Publishing Company, (1959).
15. Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill, North Carolina.
16. Gupta, S.S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist. (28), 957-967.
17. Gupta, S.S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. Contributions to Probability and Statistics. Ch. XX. Stanford University Press.
18. Gupta, S.S. and Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. Biometrika (49), 495-507.
19. Gupta, S.S. and Sobel, M. (1962b). On the smallest of several correlated F statistics. Biometrika (49), 509-523.
20. Gupta, S.S. (1963a). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math. Tokyo (14), 199-216.
21. Gupta, S.S. (1963b). Selection and ranking procedures and order statistics for the binomial distribution. Classical and Contagious Discrete Distributions, Statistical Publishing Society ch. 6, (1965).
22. Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules, Technometrics (7), 225-245.
23. Gupta, S.S. (1966). On some Selection and Ranking Procedures for Multivariate Normal Populations Using Distance Functions. Multivariate Analysis, Academic Press, (1966).
24. Gupta, S.S. and Panchapakesan, S. (1968). Some Selection and Ranking Procedures for Multivariate Normal Populations. To appear in the Proceedings Second Symposium on Multivariate Analysis. Academic Press.
25. Guttman, I. and Tiao, G.C. (1964). A Bayesian approach to some best population problems, Ann. Math. Statist. (35), 825-835.
26. Lehmann, E.L. (1961). Some Model I problems of selection. Ann. Math. Statist. (32), 990-1012.
27. Mosteller, F. (1948). A k-sample slippage test for an extreme population. Ann. Math. Statist. (19), 58-65.

28. Paulson, E. (1949). A multiple decision procedure for certain problems in analysis of variance. Ann. Math. Statist. (20), 95-98.
29. Paulson, E. (1962). A sequential procedure for comparing several experimental categories with a standard or control. Ann. Math. Statist. (33), 438-443.
30. Paulson, E. (1963). A sequential decision procedure for choosing one of k hypothesis concerning the unknown mean of a normal distribution. Ann. Math. Statist. (34), 549-554.
31. Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. Ann. Math. Statist. (35), 174-180.
32. Paulson, E. (1967). Sequential procedures for selection of the best one of several binomial populations. Ann. Math. Statist. (38), 117-123.
33. Rizvi, M.H. (1963). Ranking and selection problems of normal populations using the absolute values of their means: fixed sample size case. Department of Statistics, University of Minnesota, Technical Report #31.
34. Seal, K.C. (1955). On a class of decision procedures for ranking means of normal populations. Ann. Math. Statist. (26), 387-398.
35. Sobel, M. and Huyett, M.H. (1957). Selecting the best one of several binomial populations. Bell System Tech. J. (36), 537-576.
36. Spitzer, F., Principles of Random Walk, University Series in Higher Mathematics, 1964.
37. Wald, A., Sequential Analysis, John Wiley and Sons, 1947.

APPENDIX

TABLES SHOWING THE PERFORMANCE OF $\mathcal{A}_1(\underline{c})$ FOR SOME $\underline{c} \in C_1'$

This appendix consists of a set of tables showing the performance of the sequential procedure $\mathcal{A}_1(\underline{c})$ for $\underline{c} \in C_1'$ such that $\underline{c} = (\{cm-d\}, \{cm+d\})$ for $c \in (0,1)$ and rational, and $d > 0$. In all the tables of this appendix p will be the probability of a success, associated with the fixed sample size procedure $R(1)$ defined in Section 1.1, for some population π . The following is an explanation of the tables.

(1) Table A1. Let $r(m) = P\{\text{selecting } \pi \text{ at stage } m | \mathcal{A}_1(\underline{c})\}$, and $t(m) = P\{\text{rejecting } \pi \text{ at stage } m | \mathcal{A}_1(\underline{c})\}$ for $m \geq 1$. Then for

$c = .75, .80, d = 1, 2, 3$, and $p = .65, .80, .95$ Table A1 gives

$\sum_{m=1}^n r(m)$ (upper value) and $\sum_{m=1}^n t(m)$ (lower value) for $n=2(1) 18$,

20, 25, 30, ∞ . The values are obtained from formulas derived in

Theorem (1.3.1). For example, for $c=.75, d=2$ and $p=.80$ the proba-

bility of selecting π on or before the 17th stage is .41076, and

the probability of rejecting π on or before the 17th stage is

.12980. A major feature of the procedure $\mathcal{A}_1(\underline{c})$ shown by Table A1

is that for a fixed c and p , $\sum_{m=1}^n r(m)$ converges to $\sum_{m=1}^{\infty} r(m)$ faster

for smaller d . For example, for $c=.80, p=.95$, and $d=1$ it can be

seen that $\sum_{m=1}^{\infty} r(m) - \sum_{m=1}^{30} r(m) = .00000$, while for $d=3$ $\sum_{m=1}^{\infty} r(m) -$

$\sum_{m=1}^{30} r(m) = .00934$. The same behavior can be observed for $\sum_{m=1}^n t(m)$.

Since $r(n) = \sum_{m=1}^n r(m) - \sum_{m=1}^{n-1} r(m)$, and $t(n) = \sum_{m=1}^n t(m) - \sum_{m=1}^{n-1} t(m)$ the

table also gives values for $r(n)$ and $t(n)$.

(2) Table A2. For $c=.667, .750, .800, d=3(1) 10$ and $p=.3(11) .9$ and $.95$ Table A2 compares exact (upper values) and approximate (lower values) of selecting a population using $\mathcal{L}_1(c)$. The exact and approximate values are obtained by formulas in Theorem (1.3.3) and formula (1.4.13) respectively. The table verifies the behavior shown in Section 1.4. That is as d increases the probabilities tend monotonically to 0 if $p < c$, to 1 if $p > c$, and to $1/2$ if $p=c$. The latter case can be best seen when $c=.80$. Also for a fixed c and d as p increases the probabilities increase monotonically. Thus in any column if the probability is .00000 for some p , it will be .00000 for all $p' < p$. Similarly if in any column the probability is 1 for some p , it will be 1 for all $p' > p$. In general the approximations are less than or equal to the exact values.

Table A3. For the same values of $c, d,$ and p as in Table A2, Table A3, presents a comparison between exact (upper values) and approximate (lower values) values for the expected number of stages until a population is tagged. The exact and approximate values are gotten by formulas in Theorem (1.3.4) and formula (1.4.14) respectively. This table also verifies the results of Section 1.4. For a fixed c and p the values increase monotonically with d , but for a fixed c

and d the values attain a maximum at $p=c$, increasing for $p < c$ and decreasing for $p > c$. This again can best be seen for $c = .800$. In general, the approximations are less than or equal to the exact value.

TABLE A1

The Probability of Selection and the Probability of Rejection of a
Population on or Before the n th Stage Using $\mathcal{L}_1(\underline{c})$ for $\underline{c} \in C_1$.

$$c = .75$$

n	p = .65			p = .80			p = .95		
	d = 1	d = 2	d = 3	d = 1	d = 2	d = 3	d = 1	d = 2	d = 3
2	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	.12250	.00000	.00000	.04000	.00000	.00000	.00250	.00000	.00000
3	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	.28175	.04287	.00000	.10400	.00800	.00000	.00725	.00012	.00000
4	.17851	.00000	.00000	.40960	.00000	.00000	.81451	.00000	.00000
	.43702	.12648	.01501	.18080	.02720	.00160	.01402	.00048	.00001
5	.17851	.00000	.00000	.40960	.00000	.00000	.81451	.00000	.00000
	.43702	.12648	.01501	.18080	.02720	.00160	.01402	.00048	.00001
6	.17851	.00000	.00000	.40960	.00000	.00000	.81451	.00000	.00000
	.48412	.16452	.02866	.19718	.03334	.00262	.01445	.00052	.00001
7	.17851	.00000	.00000	.40960	.00000	.00000	.81451	.00000	.00000
	.54534	.23046	.05973	.22340	.04645	.00549	.01526	.00060	.00001
8	.24714	.03186	.00000	.57737	.16777	.00000	.95417	.66342	.00000
	.60504	.31350	.10877	.25486	.06677	.01106	.01642	.00076	.00002
9	.24714	.03186	.00000	.57737	.16777	.00000	.95417	.66342	.00000
	.60504	.31350	.10877	.25486	.06677	.01106	.01642	.00076	.00002
10	.24714	.03186	.00000	.57737	.16777	.00000	.95417	.66342	.00000
	.62315	.34519	.13141	.26157	.07222	.01287	.01650	.00077	.00002
11	.24714	.03186	.00000	.57737	.16777	.00000	.95417	.66342	.00000
	.64669	.39748	.17279	.27230	.08329	.01694	.01664	.00080	.00002
12	.27352	.05637	.00569	.64609	.30521	.06872	.97812	.89094	.54036
	.66964	.46133	.22866	.28519	.09994	.02370	.01683	.00086	.00003
13	.27352	.05637	.00569	.64609	.30521	.06872	.97812	.89094	.54036
	.66964	.46133	.22866	.28519	.09994	.02370	.01683	.00086	.00003
14	.27352	.05637	.00569	.64609	.30521	.06872	.97812	.89094	.54036
	.67660	.48523	.25175	.28794	.10432	.02567	.01685	.00086	.00003
15	.27352	.05637	.00569	.64609	.30521	.06872	.97812	.89094	.54036
	.68565	.52437	.29189	.29234	.11315	.02987	.01687	.00087	.00003
16	.28367	.07403	.01225	.67424	.41076	.15316	.98223	.96409	.81833
	.69448	.57188	.34384	.29761	.12634	.03657	.01691	.00089	.00003
17	.28367	.07403	.01225	.67424	.41076	.15316	.98223	.96409	.81833
	.69448	.57188	.34384	.29761	.12634	.03657	.01691	.00089	.00003
18	.28367	.07403	.01225	.67424	.41076	.15316	.98223	.96409	.81833
	.69718	.58959	.36457	.29761	.12980	.03845	.01691	.00089	.00003
20	.28757	.08682	.01824	.68577	.49219	.23531	.98223	.98772	.93154
	.70506	.65356	.46292	.30270	.14712	.04860	.01692	.00090	.00003
25	.28907	.09613	.02531	.69049	.55535	.30998	.98305	.99539	.97462
	.70861	.73879	.57156	.30479	.16339	.05949	.01692	.00090	.00003
30	.28965	.10293	.02761	.69242	.60449	.37627	.98308	.99789	.99063
	.70926	.76470	.61599	.30572	.17777	.07059	.01692	.00090	.00003
∞	.29001	.12151	.04770	.69377	.77864	.85769	.98308	.99910	.99997
	.70999	.87849	.95230	.30623	.22136	.14231	.01692	.00090	.00003

The upper value gives the selection probability and the lower value gives the rejection probability as defined in Theorem (1.3.1).

Table A1 (cont'd.)

c = .80

p = .65

p = .80

p = .95

n	d = 1	d = 2	d = 3	d = 1	d = 2	d = 3	d = 1	d = 2	d = 3
2	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	.12250	.00000	.00000	.04000	.00000	.00000	.00250	.00000	.00000
3	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	.28175	.04287	.00000	.10400	.00800	.00000	.00725	.00012	.00000
4	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	.43702	.12648	.01501	.18080	.02720	.00160	.01402	.00048	.00001
5	.11603	.00000	.00000	.32768	.00000	.00000	.77378	.00000	.00000
	.57158	.23517	.05402	.26272	.05792	.00672	.02259	.00116	.00003
6	.11603	.00000	.00000	.32768	.00000	.00000	.77378	.00000	.00000
	.57158	.23517	.05402	.26272	.05792	.00672	.02259	.00116	.00003
7	.11603	.00000	.00000	.32768	.00000	.00000	.77378	.00000	.00000
	.60985	.27638	.07621	.27910	.06611	.00877	.02310	.00121	.00003
8	.11603	.00000	.00000	.32768	.00000	.00000	.77378	.00000	.00000
	.65960	.34335	.11948	.30532	.08250	.01368	.02407	.00134	.00004
9	.11603	.00000	.00000	.32768	.00000	.00000	.77378	.00000	.00000
	.70810	.42344	.18043	.33678	.10661	.02220	.02545	.00156	.00006
10	.15227	.01346	.00000	.46190	.10737	.00000	.93134	.59874	.00000
	.75014	.50719	.25416	.37033	.13765	.03489	.02719	.00190	.00009
11	.15227	.01346	.00000	.46190	.10737	.00000	.93134	.59874	.00000
	.75014	.50719	.25416	.37033	.13765	.03489	.02719	.00190	.00009
12	.15227	.01346	.00000	.46190	.10727	.00000	.93134	.59874	.00000
	.76209	.53551	.28258	.37704	.14503	.03833	.02730	.00193	.00009
13	.15227	.01346	.00000	.46190	.10737	.00000	.93134	.59874	.00000
	.77763	.57986	.33035	.38778	.15926	.04544	.02749	.00198	.00009
14	.15227	.01346	.00000	.46190	.10737	.00000	.93134	.59874	.00000
	.79279	.63154	.39009	.40066	.17966	.05640	.02777	.00207	.00011
15	.16360	.02187	.00156	.51687	.19534	.03518	.96343	.84257	.46329
	.80592	.68457	.45596	.41441	.20543	.07126	.02813	.00221	.00012
16	.16360	.02187	.00156	.51687	.19534	.03518	.96343	.84257	.46329
	.80592	.68457	.45596	.41441	.20543	.07126	.02813	.00221	.00012
17	.16360	.02187	.00156	.51687	.19534	.03518	.96343	.84257	.46329
	.80965	.70226	.47957	.41716	.21148	.07500	.02815	.00222	.00012
18	.16360	.02187	.00303	.51687	.19534	.03518	.96343	.84275	.46329
	.81451	.72981	.51793	.42155	.22307	.08249	.02819	.00225	.00013
20	.16713	.02687	.00303	.53934	.26379	.07842	.96996	.93691	.74631
	.82334	.79440	.61438	.43246	.26040	.10844	.02832	.00234	.00014
25	.16824	.02985	.00412	.54861	.31751	.12092	.97129	.97371	.88462
	.82879	.86158	.72947	.43486	.30447	.14379	.02836	.00239	.00015
30	.16858	.03165	.00440	.55239	.35994	.16019	.97156	.98817	.94814
	.83049	.90247	.81066	.44289	.33964	.17645	.02837	.00241	.00015
∞	.16874	.03441	.00698	.55501	.52567	.51613	.97163	.99758	.99968
	.83126	.96558	.99302	.44499	.46433	.48387	.02837	.00242	.00032

The upper value gives the selection probability and the lower value gives the rejection probability as defined in Theorem (1.3.1).

TABLE A2

Comparisons of Exact and Approximate Values of the
Probability of Selecting a Population Using $f_1(c)$ for $c \in C_1$.

$$c = .667$$

$d \backslash p$	3	4	5	6	7	8	9	10
.30	.00004	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.40	.00088	.00008	.00001	.00000	.00000	.00000	.00000	.00000
.50	.01301	.00323	.00073	.00017	.00004	.00001	.00000	.00000
.60	.14913	.08862	.05114	.02911	.01642	.00921	.00515	.00288
.70	.72577	.78097	.82802	.86698	.89829	.92292	.94198	.95717
.80	.98466	.99573	.99892	.99972	.99993	.99998	.99999	1.00000
.90	.99983	.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
.95	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

$$c = .750$$

$d \backslash p$	3	4	5	6	7	8	9	10
.30	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.40	.00003	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.50	.00067	.00006	.00001	.00000	.00000	.00000	.00000	.00000
.60	.01183	.00271	.00062	.00014	.00003	.00001	.00000	.00000
.70	.17965	.11523	.07167	.04401	.02678	.01618	.00974	.00585
.80	.85823	.91345	.94835	.96982	.98254	.98997	.99426	.99671
.90	.99797	.99982	.99997	.99999	1.00000	1.00000	1.00000	1.00000
.95	.99987	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

The upper value gives the exact probability and the lower value gives the approximate probability as defined in Theorem (1.3.3) and (1.4.13) respectively.

TABLE A2 (cont'd.)

$$c = .800$$

p \ a	3	4	5	6	7	8	9	10
.30	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.40	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.50	.00005	.00000	.00000	.00000	.00000	.00000	.00000	.00000
.60	.00137	.00015	.00002	.00000	.00000	.00000	.00000	.00000
.70	.03192	.01029	.00329	.00105	.00034	.00011	.00003	.00001
.80	.51642	.51243	.50980	.50820	.50704	.50617	.50549	.50483
.90	.99237	.99845	.99961	.99991	.99998	1.00000	1.00000	1.00000
.95	.99568	.99904	.99979	.99995	.99999	1.00000	1.00000	1.00000
	.99991	.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
	.99995	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

The upper value gives the exact probability and the lower value gives the approximate probability as defined in Theorem (1.3.3) and (1.4.13) respectively.

TABLE A3
 Comparisons of Exact and Approximate Values of the Expected
 Number of Stages to Tag a Population Using $\rho_1(\underline{c})$ for $\underline{c} \in C_1$.

$c = .667$

$\frac{d}{p}$	3	4	5	6	7	8	9	10
.30	8.60372	11.28677	14.03340	16.75234	19.48321	22.20893	24.93687	27.66386
.40	8.18118	10.90906	13.63636	16.36364	19.08091	21.81818	24.54545	27.27278
.50	11.75565	15.50056	19.26015	23.00805	26.75884	30.50860	34.25868	38.00866
.60	11.23029	14.99748	18.74970	22.49997	26.25000	30.00006	33.75000	37.50000
.70	18.29568	24.61052	30.71997	36.75121	42.76050	48.76299	54.76364	60.76367
.80	17.53255	23.85139	29.95604	35.98754	41.99657	47.99907	53.99975	59.99994
.90	33.09341	50.99954	69.00990	86.48004	103.29462	119.54454	135.37125	150.90322
.95	31.72734	49.45058	67.35486	84.77022	101.55605	117.79089	133.60958	149.13723
	39.75252	66.69666	97.85318	131.68418	166.95291	202.75547	238.49365	273.81345
	38.60186	65.41451	96.53744	130.41573	165.78469	201.71659	237.59461	273.05193
	21.79886	29.75166	37.41842	44.97430	52.49213	59.99764	67.49930	74.99980
	21.69942	29.71594	37.40642	44.97048	52.49095	59.99728	67.49920	74.99977
	12.85282	17.14250	21.42854	25.71428	30.00000	34.28571	38.57143	42.85714
	12.85133	17.14238	21.42854	25.71428	30.00000	34.28571	38.57143	42.85714
	10.58617	14.11765	17.64706	21.17647	24.70588	28.23529	31.76471	35.29412
	10.58813	14.11764	17.64706	21.17647	24.70588	28.23529	31.76471	35.29412

The upper value gives the exact expectations and the lower value gives the approximate expectations as defined in Theorem (1.3.4) and (1.4.14) respectively.

TABLE A3 (cont'd.)

c = .750

d/p	3	4	5	6	7	8	9	10
.30	7.13713	9.41928	11.58861	13.83273	16.05320	18.27096	20.49680	22.71762
	6.66667	8.88889	11.11111	13.33333	15.55556	17.77778	20.00000	22.22222
.40	9.17038	12.05081	14.88794	17.75304	20.60873	23.46547	26.32307	29.18001
	8.57094	11.42855	14.28571	17.14286	20.00000	22.85714	25.71429	28.57143
.50	12.78870	16.81179	20.80628	24.80894	28.80849	32.80848	36.80853	40.80851
	11.98400	15.99813	19.99980	23.99998	28.00000	32.00000	36.00000	40.00000
.60	20.77636	27.78575	34.55610	41.25421	47.92927	54.59822	61.26548	67.93230
	19.52819	26.52158	32.29180	39.98867	46.66363	53.33254	59.99980	66.66661
.70	41.31767	64.73384	88.91969	112.78295	135.91359	158.27076	179.96521	201.14586
	38.82998	61.82296	85.75503	109.47087	132.51822	154.82837	176.49610	197.66140
.80	42.53566	65.83342	89.50817	112.66219	135.05904	156.75930	177.91445	198.67682
	41.25401	64.72848	88.65441	112.04634	134.63487	156.47631	177.72991	198.55851
.90	19.93318	26.65503	33.33142	39.99970	46.66662	53.33333	60.00000	66.66667
	19.90937	26.65075	33.33072	39.99959	46.66660	53.33332	60.00000	66.66667
.95	14.99897	19.99995	25.00000	30.00000	35.00000	40.00000	45.00000	50.00000
	14.99835	19.99992	25.00000	30.00000	35.00000	40.00000	45.00000	50.00000

The upper value gives the exact expectations and the lower value gives the approximate expectations as defined in Theorem (1.3.4) and (1.4.14) respectively.

TABLE A3 (cont'd.)

c = .800

d \ p	3	4	5	6	7	8	9	10
.30	6.45981	8.56829	10.57875	12.52345	14.54671	16.55493	18.54475	20.54504
.40	6.00000	8.00000	10.00000	12.00000	14.00000	16.00000	18.00000	20.00000
.50	8.10829	10.67508	13.16431	15.64967	18.15840	20.65830	23.15660	25.65731
.60	7.49998	10.00000	12.50000	15.00000	17.50000	20.00000	22.50000	25.00000
.70	10.81044	14.17982	17.50440	20.83501	24.17069	27.50354	30.83674	34.17018
.80	9.99894	13.33328	16.66666	20.00000	23.33333	26.66667	30.00000	33.33333
.90	16.13184	21.18528	26.18639	31.18683	36.18738	41.18728	46.18728	51.18729
.95	14.95883	19.99389	24.99915	29.99989	34.99999	40.00000	45.00000	50.00000
	30.23735	41.36985	51.87526	62.08310	72.16394	82.19428	92.20542	102.20944
	28.11492	39.18089	49.67117	59.87387	69.95304	79.98288	89.99386	99.99782
	67.20000	112.00000	179.20000	256.00000	358.40000	448.00000	595.20000	704.00000
	56.25000	100.00000	156.25000	225.00000	336.25000	430.00000	586.25000	695.00000
	29.50931	39.85544	49.96013	59.98945	69.94728	79.99931	89.99983	99.99996
	29.35889	39.80946	49.94725	59.98600	69.99639	79.99909	89.99977	99.99994
	19.98963	26.66613	33.33329	40.00000	46.66654	53.33333	60.00000	66.66667
	19.98963	26.66579	33.33326	39.99999	46.66667	53.33333	60.00000	66.66667

The upper value gives the exact expectations and the lower value gives the approximate expectations as defined in Theorem (1.3.20) and (1.4.14) respectively.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
multiple decision rules selection and ranking sequential procedures random walk Monte-Carlo simulations						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.
- 2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.
- 2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.
3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.
4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.
5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.
6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.
- 7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.
- 7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.
- 8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.
- 8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.
- 9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.
- 9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).
10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U)

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.