ON SEQUENTIAL RANKING AND SELECTION PROCEDURES

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, Indiana 47907

S. Panchapakesan
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901

Short Title: Sequential Ranking and Selection

Key words and Phrases: Open, closed, and truncated sequential procedures; indifference-zone; subset selection; normal; Bernoulli; exponential family; vector-at-a-time sampling; adaptive sampling; curtailment; decision-theoretic.

AMS 1980 Subject Classification: 62F07, 62L10

ON SEQUENTIAL RANKING AND SELECTION PROCEDURES*

by

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, Indiana

S. Panchapakesan
Department of Mathematics
Southern Illinois University
Carbondale, Illinois

Technical Report # 88-32C

Department of Statistics
Purdue University

June, 1988
Revised March 1989

* This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grant DMS-8606964 at Purdue University.

ON SEQUENTIAL RANKING AND SELECTION PROCEDURES

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, Indiana 47907

S. Panchapakesan
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901

1. INTRODUCTION

Problems of statistical inference that are now commonly known as ranking and selection problems gained the attention of statistical researchers in the early 1950's. Early work in this area by Bahadur [1], Bahadur and Robbins [2], Bechhofer [4], and Gupta [21] related to single-stage procedures. Interest in sequential selection procedures arose in the early days and has steadily continued ever since. However, it was a decade before a substantial amount of original research on sequential methods for ranking and selection problems was published in the form of a monograph by Bechhofer, Kiefer and Sobel [12] which still serves well as a constant source of results and ideas.

Two–stage and multi–stage procedures can be viewed as sequential procedures with the number of stages to make the terminal decision bounded above. Such procedures arise not only in the context of efficiency compared to single–stage procedures but also out of necessity. For example, nuisance parameters lead to two or multi–stage selection procedures for the normal means selection problem when the variances are common unknown or completely unknown. More importantly, the "measure of distance" used may require multi–stage or open sequential procedures. In particular, this is the situation when the "odds ratio" is used as the measure of distance for ranking Bernoulli populations (see Bechhofer, Kiefer and Sobel [12],

Section 6.7.1).

Selection procedures have been studied under various goals such as selecting the best among k (≥ 2) populations, selecting the t best ($1 \leq t < k$), and selecting the populations better than a standard or control. In all these cases, the procedure is devised to select a subset of the k given populations which is of either a fixed size or a random size. The fixed-size subset selection in the classical formulation is known as the indifference-zone (IZ) approach and the other type is called the subset selection (SS) approach. More will be said about these in the next section.

Besides the monograph of Bechhofer, Kiefer and Sobel [12] devoted entirely to sequential procedures, a few other books on ranking and selection are: Büringer, Martin and Schriever [15], Gupta and Huang [24], and Gupta and Panchapakesan [27]. The last book [27] mentioned provides a comprehensive survey of developments in this field up to 1978 with an extensive bibliography. Dudewicz and Koo [17] have given a categorized bibliography. Recently, Gupta and Panchapakesan [28] have surveyed developments in the subset selection theory over a period of more than thirty years with emphasis on historical perspectives. A nice review of developments in the multi-stage selection theory since 1979 is given by Miescke [39].

In the present paper, we do not attempt to give a complete account of sequential methods in ranking and selection. Our purpose here is to provide a basic background, give highlights of some of the early developments and their impact on some current developments.

Section 2 gives a general background for sequential selection procedures, explaining the basic aspects of the indifference-zone and subset approaches. The

specific procedures discussed here center around selecting the normal population having the largest mean, and selecting the Bernoulli population having the largest success probability. These are discussed respectively in Sections 3 and 4 for the indifference-zone approach, and in Sections 5 and 6 for the subset selection approach. Section 7 deals with subset selection from exponential family distributions and a decision-theoretic approach to the problem.

2. SOME GENERAL ASPECTS OF SEQUENTIAL SELECTION PROCEDURES

Let $\pi_1, \ldots, \pi_k$ be k given populations. From each π_i, a sequence of independent observations $X_{i1}, X_{i2}, \ldots$ is available to the experimenter. Let X_{ij} have a density f_{θ_i} with respect to (w.r.t.) a σ-finite measure on R, which is the Lebesgue measure or a counting measure. The parameters $\theta_i, i = 1, \ldots, k$, are assumed to be unknown. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered θ_i. No prior knowledge is assumed regarding the true pairing of the ordered and unordered θ_i. Our goal is to select the population π_i which has the largest associated θ_i and is called the best population. In case of a tie, we consider one such population is tagged as the best. Let $\Omega = \{\underset{\sim}{\theta} : \underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k),\ \theta_i \in \Theta,\ i = 1, \ldots, k\}$ denote the parameter space, where Θ is taken to be some interval (finite or infinite) on the real line. Let $D(\theta_i, \theta_j) \geq 0$ be an appropriately defined distance measure between the populations π_i and π_j. For $\delta^* > 0$, define

$$\Omega(\delta^*) = \{\underset{\sim}{\theta} | D(\theta_{[k-1]}, \theta_{[k]}) \geq \delta^*\}. \tag{1}$$

In the case of location parameters θ_i, for example, a natural choice is $D(\theta_{[k-1]}, \theta_{[k]}) = \theta_{[k]} - \theta_{[k-1]}$.

Under the IZ approach of Bechhofer [4], a valid procedure R selects one of the

k populations as the best with a guarantee that

$$P_{\underset{\sim}{\theta}}(CS|R) \geq P^* \text{ whenever } \underset{\sim}{\theta} \in \Omega(\delta^*) \tag{2}$$

where $P_{\underset{\sim}{\theta}}(CS|R)$ denotes the probability of a correct selection (PCS) using the rule R under the parametric configuration $\underset{\sim}{\theta}$. Here a correct selection (CS) occurs if the population selected is indeed the best one. The minimum probability level $P^*(\frac{1}{k} < P^* < 1)$ and δ^* are specified in advance by the experimenter. The complement of $\Omega(\delta^*)$ in (1) w.r.t. Ω is called the indifference-zone since we have no PCS requirement for $\underset{\sim}{\theta}$ in this part. The part $\Omega(\delta^*)$ is known as the preference-zone.

In the subset selection (SS) approach of Gupta [21, 22], a valid procedure R selects a random-sized subset of the given populations with a guarantee that

$$P_{\underset{\sim}{\theta}}(CS|R) \geq P^* \text{ for all } \underset{\sim}{\theta} \in \Omega \tag{3}$$

where a correct selection (CS) occurs if the best population is included in the selected subset. We note that there is no indifference-zone in the SS approach.

The probability requirements (2) and (3) are usually referred to as the basic probability requirements or P^*-requirements or P^*-conditions of the respective formulations. In either of these classical approaches, one proposes a "reasonable" procedure which involves some quantities to be defined so that the P^*-requirement is met. This involves the all-important first step of finding the least favorable configuration (LFC) of $\underset{\sim}{\theta}$ (in $\Omega(\delta^*)$ or Ω, depending on the approach) for which the infimum of $P(CS|R)$ over the appropriate space takes place. The necessary quantities involved in the rule R are then determined such that this infimum is at least P^*.

One would then study its properties, evaluate its performance according to suitable criteria, and compare the performance with that of any known alternative procedures.

Although we have discussed here only the goal of selecting the one best population, the IZ and SS approaches are by no means restricted to only this goal. More generalized goals have been considered in the literature using both approaches.

A selection procedure, though not always explicitly so stated, typically consists of three parts: (1) a sampling rule, (2) a stopping rule, and (3) a terminal decision rule. Procedures are usually categorized according to the types of rules employed in the above three parts. The terminal decision identifies a procedure as a fixed size or a random size subset selection procedure. A sequential procedure is said to be closed or open according as the number of observations that can be drawn from each population is a bounded or an unbounded random variable. Any given sequential procedure (open or closed) can yield a truncated version by a modification of the stopping rule so as to achieve an earlier termination. A sequential procedure with elimination may eliminate one or more populations (which appear to be inferior) before reaching the final stage at which the terminal decision is made. Typically, further sampling from eliminated populations is discontinued although this is not the case with some procedures studied in the literature.

Sampling may be done one-at-a-time or vector-at-a-time. The former is an adaptive sampling in which the population to be sampled from next depends on the data accumulated until then. Play-the-winner sampling rule of Robbins [44] in the case of Bernoulli populations is an instance of the adaptive case. In the

vector-at-a-time sampling, a vector of observations (one from each) is taken from the non-eliminated populations.

3. SELECTION FROM NORMAL POPULATIONS: IZ APPROACH

Let $\pi_1, \ldots, \pi_k$ be k normal populations with unknown means $\theta_1, \ldots, \theta_k$, respectively, and a common variance σ^2. For defining the preference-zone in (1), we take $D(\theta_{[k-1]}, \theta_{[k]}) = \theta_{[k]} - \theta_{[k-1]}$; thus

$$\Omega(\delta^*) = \{\underset{\sim}{\theta} | \theta_{[k]} - \theta_{[k-1]} \geq \delta^* > 0\}.$$

Our goal is to select the population associated with $\theta_{[k]}$ and any valid rule should satisfy the P^*-requirement (2). We will discuss the known σ^2 case first.

3.1 Case A: Known σ^2. Let X_{ij}, $j = 1, 2, \ldots$ be a sequence of independent observations from π_i, $i = 1, \ldots, k$. Unless stated otherwise, the observations are taken vector-at-a-time. Let $Y_{im} = \sum_{j=1}^{m} X_{ij}$, $i = 1, \ldots, k$, and let $Y_{[1]m} \leq \ldots \leq Y_{[k]m}$ denote the ordered Y_{im}.

Stein's Procedure, $R_{NIZ:S}$. Stein [49], using a slightly more general model than ours, proposed an open sequential procedure with elimination which is a straightforward application of a lemma of Wald. For every population π_i sampled from at stage $m (m = 1, 2, \ldots)$, let

$$A_{im} = \sum_{j=1}^{m} [X_{ij} - \overline{X}_j - \delta^*(t_j - 1)/t_j],$$

where $\overline{X}_j$ is the average of the observations at stage j, and t_j is the number of populations sampled from at this stage. Stein's procedure (for our model) is described below:

Procedure $R_{NIZ:S}$: At stage m $(m = 1, 2, \ldots)$, eliminate all populations π_i for which $A_{im} \leq (\sigma^2/\delta^*)\ell n(1 - P^*)$ and proceed to stage $(m+1)$ to take an additional observation from each remaining population. Stop the experiment at any stage m, if there is at most one i for which $A_{im} > (\sigma^2/\delta^*)\ell n(1 - P^*)$. If there is exactly one such i at termination, then select that π_i as the best; otherwise (i.e. no such A_{im}), select π_i corresponding to the largest A_{im} at termination.

The performance characteristics of the above procedure $R_{NIZ:S}$ have not been studied.

Bechhofer-Kiefer-Sobel Procedure, $R_{NIZ:BKS}$. In their monograph, Bechhofer, Kiefer and Sobel [12] considered selection from populations belonging to an exponential family. Their procedure is an open sequential one with no elimination. This procedure is specialized by them [12, pp. 264–265] to the normal case at hand. For each m $(m = 1, 2, \ldots)$, let $W_m = \sum_{i=1}^{k-1} \exp\{-\delta^*(Y_{[k]m} - Y_{[i]m})/\sigma^2\}$.

Procedure $R_{NIZ:BKS}$: Stop sampling when $m = N$, the first positive integer for which $W_m \leq (1 - P^*)/P^*$; select the population corresponding to the largest Y_{iN}.

A virtue of the above procedure is that it can react rapidly to favorable configurations of the population means. For example, if $\theta_{[k]} - \theta_{[k-1]} >> \sigma$, then with high probability the procedure will stop at the first stage. However, if $\theta_{[k]} - \theta_{[i]}$ is small, then N (the stopping time) can be large with a considerable probability. Further, the variance of N can be large. To overcome these undesirable effects, Bechhofer and Goldsman [7] proposed a truncated version (described below) of the above procedure.

Bechhofer-Goldsman Procedure, $R_{NIZ:BG}$. This procedure modifies the stopping rule of $R_{NIZ:BKS}$ as follows: Stop sampling when, for the first time, either $W_m \leq (1-P^*)/P^*$ or $m = n_0$, whichever occurs first. Here $n_0 = n_0(k, \delta^*, P^*)$ is predetermined as the smallest positive integer which guarantees the P^*-requirement (2). The terminal decision rule is: Select the population corresponding to the largest Y_{iN}, where N is now the bounded stopping time.

Bechhofer and Goldsman [7] have tabulated the n_0 values for $k = 2(1)5$, $\delta^* = 0.2(0.2)\ 0.8$, and $P^* = 0.75, 0.90, 0.95, 0.99$. Additional n_0 values are contained in Bechhofer and Goldsman [11] for $\delta^* = 0.3(0.2)0.7$ and the same selected values of k and P^* as previously stated.

Another well-known procedure in the literature is that of Paulson [40], who was the first to consider a closed procedure with elimination, a feature to be characterized by some later authors as Paulson-type. Paulson, in fact, considered a class of procedures indexed by $\lambda \in (0, \delta^*)$, using triangular stopping regions. Let $a_\lambda = \{\sigma^2/(\delta^* - \lambda)\}\ell n\{(k-1)/(1-P^*)\}$ and let W_λ denote the largest integer less than a_λ/λ.

Paulson Procedure, $R_{NIZ:P}$: At the beginning of stage m $(m = 1, \ldots, W_\lambda)$, take one observation from each population not eliminated thus far. Now eliminate all populations π_i for which

$$Y_{im} < \max_r Y_{rm} - a_\lambda + m\lambda$$

where the maximum is over all populations π_r that remain at the beginning of stage m. If all but one population are eliminated, then stop sampling and select this one remaining population; otherwise, proceed to stage $(m+1)$. If two or more

populations remain after stage W_λ, then take an additional observation from each one of them and select the population π_i corresponding to the largest $Y_{i(W_\lambda+1)}$.

Although $R_{NIZ:P}$ guarantees the P^*-requirement, the optimum value of λ in $(0, \delta^*)$ was not settled by Paulson. However, based on his calculations, he recommended the choice of $\lambda = \delta^*/4$. Bechhofer and Goldsman [10] point out that $\lambda = \delta^*/2$ minimizes $W_\lambda + 1$, the maximum possible total number of stages to termination, for any given set of k, δ^*, and P^*.

Improvements in Paulson's Procedure. Fabian [20] improved Paulson's procedure by obtaining better lower bound on the PCS. Considering the choices of $\lambda = \delta^*/2$ and $\lambda = \delta^*/4$, Fabian's improvement is achieved by replacing $c = (k-1)/(1-P^*)$ by $c/2$ for $\lambda = \delta^*/2$ and by $1/q$ for $\lambda = \delta^*/4$, where q satisfies $(q - \frac{1}{2}q^{4/3})c = 1$. Recently, Hartmann [30] improved upon Fabian's results by replacing the reciprocal of c by $1 - (P^*)^{1/(k-1)}$.

Some Comparison Results. Bechhofer and Goldsman [10] have performed Monte Carlo studies to compare the performances of $R_{NIZ:BKS}, R_{NIZ:BG}$, $R_{NIZ:PH}$ (i.e. $R_{NIZ:P}$ with Hartmann modification), the two-stage procedure of Tamhane and Bechhofer [52, 53] and the single-stage procedure of Bechhofer [4]. We refer to the last two procedures as $R_{NIZ:TB}$ and $R_{NIZ:B}$, respectively. The performances of these procedures were studied by Bechhofer and Goldsman [10] in terms of achieved PCS, $E(N)$, and $E(T)$, where N and T are the total number of stages needed to terminate and the total number of observations taken up to termination. Their results indicate that $R_{NIZ:BG}$ does well in terms of $E(N)$ except when the θ_i are all very close to each other and P^* is high, in which case

$R_{NIZ:PH}$ with $\lambda = \delta^*/2$ is recommended. When $k > 5$, they recommend $R_{NIZ:PH}$ with $\lambda = \delta^*/2$ for the equal means (EM) configuration and $\lambda = \delta^*/4$ otherwise. For reasonably high P^* with $E(T)$ as the criterion, $R_{NIZ:PH}$ seems preferable with choices of λ as indicated above.

Kao-Lai Procedure, $R_{NIZ:KL}$. A class of truncated procedures with elimination was proposed by Kao and Lai [35] employing confidence sequences for the $(k-1)$ differences $\theta_{[k]} - \theta_i$ $(i \neq [k])$. Taking $E(T)$ as a measure of efficiency, it has been shown by Kao and Lai [35] that asymptotically $(P^* \to 1)$ their procedure is more efficient than $R_{NIZ:BKS}, R_{NIZ:P}$, and $R_{NIZ:B}$ except when $\underset{\sim}{\theta}$ is in the least favorable (slippage) configuration or in the EM-configuration; in these configurations, their procedure is at least as efficient as the others.

A Generalized Goal. Fabian [19] considered a generalized goal for ranking populations. For our problem of selecting the best population, this corresponds to δ^*-correct selection (δ^*-CS) which means selecting any π_i for which $\theta_i > \theta_{[k]} - \delta^*$. Such a π_i is called a good population. For $\underset{\sim}{\theta} \in \Omega(\delta^*)$ in the IZ approach, the best population is the only good population. Fabian [19] has shown that, for the single-stage procedure of Bechhofer [4], a stronger claim can be made, namely, that $P(\delta^*-CS|R) \geq P^*$ for all $\underset{\sim}{\theta} \in \Omega$.

Kao and Lai [35] have given a sequential procedure (by slightly modifying the elimination rule of $R_{NIZ:KL}$) which guarantees a minimum probability P^* of a δ^*-CS. As pointed out by Edwards [18], this is done at the expense of considerably slower elimination of inferior populations. Edwards [18] gave a slightly different but more general procedure, which he called an extended-Paulson sampling plan. His

procedure guarantees a minimum probability of a δ^*-CS while keeping asymptotic ($P^* \rightarrow 1$) sample size properties same as those of the IZ procedures $R_{NIZ:BKS}$, $R_{NIZ:P}$, and $R_{NIZ:B}$.

Other Developments. Hoel [31] has discussed a method of constructing Paulson-type procedures based on log-likelihood ratios, which can be applied to the normal means problem. For appropriate choices of the index defining a family of procedures, Hoel's procedure is precisely $R_{NIZ:P}$.

Recently, Bechhofer and Goldsman [8] have considered selection of normal population with the largest mean when the underlying model is a two-factor experiment with no interaction. Their procedure is a natural adaptation of $R_{NIZ:BKS}$. In a later paper [9], they studied a truncated version of this adapted procedure and carried out Monte Carlo studies of performances of these procedures and the single-stage procedure of Bechhofer [4].

3.2 Case B: Unknown σ^2. When σ^2 is unknown, there does not exist a single-stage procedure that can guarantee the P^*-requirement under the IZ formulation. This is because the necessary sample size cannot be determined without the knowledge of σ^2. Bechhofer, Dunnett and Sobel [5] proposed a two-stage procedure where the first stage samples are used to provide an estimate of σ^2; the additional second-stage sample size, if necessary, is determined based on the first stage data. Paulson [40], and Kao and Lai [35] have given procedures by modifying their earlier procedures for the case of known σ^2. These involve first taking m (≥ 2) observations from each population and then proceeding sequentially by taking one observation from each noneliminated population. Robbins, Sobel and Starr [45] proposed a procedure for

which the P^*–requirement is asymptotically ($\delta^* \to 0$) satisfied. Details of these procedures will not be discussed here. These procedures, except that of Kao and Lai [35], have been discussed in Gupta and Panchapakesan [27, Chapter 6].

4. SELECTION FROM BERNOULLI POPULATIONS: IZ APPROACH

Let $\pi_1, \ldots, \pi_k$ be k Bernoulli populations with associated success probabilities $p_1, \ldots, p_k$, respectively. Consider the preference-zone $\Omega_{\delta^*} = \{\underline{p} : \underline{p} = (p_1, \ldots, p_k), p_{[k]} - p_{[k-1]} \geq \delta^*\}$, where $p_{[1]} \leq \ldots \leq p_{[k]}$ are the ordered p_i, and $0 < \delta^* < 1$ is specified in advance. For selecting the population associated with $p_{[k]}$ (the best population), Sobel and Huyett [48] studied a single-stage procedure based on a sample of size n from each population. This procedure is $R_{BIZ:SH}$: Select the population corresponding to the largest number of observed successes, breaking ties by randomization.

For this problem, Paulson [41, 42] proposed truncated sequential procedures with elimination. There are also a number of other procedures studied by several authors; these procedures differ in their sampling and/or stopping rules. A detailed discussion of some of these procedures is given in Gupta and Panchapakesan [27, Chapter 4]. An excellent bibliography of these procedures is contained in Bechhofer and Kulkarni [13].

In our above discussion, the only distance function considered is $p_{[i]} - p_{[j]} (i > j)$. However, another equally important function is the odds ratio $p_{[i]}(1 - p_{[j]})/p_{[j]}(1 - p_{[i]})$, $i < j$, which is very important in medical applications. This has been considered in Bechhofer–Kiefer–Sobel [12]. Also of interest is $p_{[i]}/p_{[j]}$, $i > j$, which has been used by Taheri and Young [51].

Recently, Bechhofer and Kulkarni [13] proposed a closed sequential procedure. This procedure is, however, not an indifference–zone procedure. The LFC of the p–values (which is central to designing the experiment using the IZ approach) is of no concern here. The focus is on the PCS for a given n for the particular goal considered and on achieving it with minimum cost. Their sampling rule involves taking at each stage one observation from a population to be determined by the accumulated data up to that stage; in other words, it is a one-at-a-time adaptive sampling. Also, a maximum n is set for the number of observations that can be drawn from any population.

Let n_{im} and Z_{im} denote, respectively, the total number of observations taken from π_i and the number of successes among them through stage m, $i = 1, \ldots, k$ and $m = 0, 1, \ldots, kn$. Stage 0 (i.e. no observation is yet taken) is introduced for convenience in describing the procedure $R_{B:BK}$ of Bechhofer and Kulkarni [13], which is as follows:

a. At stage m $(0 \leq m \leq kn - 1)$, take the next observation from the population which has the smallest number of failures among all π_i for which $n_{im} < n$. In case of a tie among such π_i's, take the next observation from the one which has the largest number of successes. In case of a further tie, select one of this further tied set at random and draw the next observation from that population.

b. Stop sampling at the first stage m at which there exists at least one π_i satisfying

$$Z_{im} \geq Z_{jm} + n - n_{jm} \text{ for all } j \neq i. \tag{4}$$

c. Select as the best one at random from those π_i's which satisfy (4) at termination.

Bechhofer and Kulkarni [13] have shown that the PCS for $R_{B:BK}$ equals that of $R_{BIZ:SH}$ uniformly in p for $k \geq 2$. Several optimal properties of $R_{B:BK}$ have been established by Kulkarni and Jennison [37], and further stronger properties are contained in Jennison [33]. Exact numerical results are given by Bechhofer and Kulkarni [14] for performance characteristics such as the distributions of $N_{(i)}$, the number of observations taken at truncation from π_i associated with $p_{[i]}$, and of the total number $N = \sum_{i=1}^{k} N_{(i)}$ at truncation. Because of the nature of time-consuming recursive formulae, their numerical results are limited to the cases of $(k, n) = (2, 20)$ and $(3, 7)$ for the distributions of $N_{(i)}$ and N, and are limited to $(k = 2,\ n \leq 100)$ and $(k = 3,\ n \leq 40)$ for $E\{N_{(i)}\}$ and $E\{N\}$. The scope of these studies is extended to $k = 4$ and 5 by Bechhofer and Frisardi [6] employing Monte Carlo simulation. Exact analytical results for various performance characteristics of the Bechhofer–Kulkarni procedure when $k = 2$ have been given by Percus and Percus [43].

The sampling rule of the Bechhofer-Kulkarni procedure is not a play-the-winner rule (see Bechhofer and Kulkarni [14]); it is referred to as the least failures rule by Kelly [36] who proposed it for a Bernoulli multi-armed bandit problem.

The idea behind the stopping rule of the Bechhofer-Kulkarni procedure is that the sampling can be curtailed as soon as there exists one or more populations which have at least as many successes as the maximum possible number of successes from any of the other populations even if all the n observations were taken from them. This criterion as given in (4) is referred to as strong curtailment by Jennison [32] who also considered weak curtailment given by (4) with $\geq$ replaced by $>$. With either curtailment, the Bechhofer-Kulkarni procedure achieves the same PCS as

does the Sobel-Huyett single-stage procedure uniformly in $p = (p_1, \ldots, p_k)$. As noted by Jennison [32], strong curtailment is preferable to weak curtailment since the former yields a sample size no larger than that yielded by the latter.

Jennison and Kulkarni [34] have considered similar procedures for the goal of selecting the s $(1 \leq s \leq k-1)$ best of k Bernoulli populations. Recently, David and Andrews [16] have proposed procedures with strong and weak curtailments for selecting the best of k objects in a Round Robin-type paired comparison experiment. They have shown that the probabilities of selecting a particular object are the same under both curtailments for the Bradley-Terry model, but are not so, in general.

5. SELECTION FROM NORMAL POPULATIONS: SS APPROACH

Let $\pi_1, \ldots, \pi_k$ be k normal populations where π_i has mean θ_i and variance σ_i^2, $i = 1, \ldots, k$. Before discussing sequential procedures for different goals, we state the single-stage procedure of Gupta [21] when $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$ (known). His procedure is based on $\overline{X}_i$, $i = 1, \ldots, n$, the means of random samples of size n from the k populations and is given below.

Gupta's Single-stage Procedure, $R_{NSS:G}$: Select π_i if and only if

$$\overline{X}_i \geq \max_{1 \leq j \leq k} \overline{X}_j - \frac{D\sigma}{\sqrt{n}} \tag{5}$$

where the constant $D = D(k, P^*)$ is the smallest positive number for which the P^*-requirement (3) is satisfied. This constant D is given by

$$\int_{-\infty}^{\infty} \Phi^{k-1}(x + D)\, d\Phi(x) = P^* \tag{6}$$

where Φ denotes the standard normal distribution function.

Let p_i denote the probability of selecting the population associated with $\theta_{[i]}$, $i = 1, \ldots, k$. Then it is known that $p_1 \leq p_2 \leq \ldots \leq p_k$ (i.e. the procedure is monotone), where

$$p_i = \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j \neq i}}^{k} \Phi\{x + D + (\theta_{[i]} - \theta_{[j]})\sqrt{n}\,\sigma^{-1}\}\, d\Phi(x). \tag{7}$$

5.1 Barron-Gupta Procedure, $R_{NSS:BG}$. This procedure is devised for selecting a subset containing the best (i.e. one having the largest θ_i) assuming that $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$ (known) and that the successive differences of the ordered θ_i are known (this implies that the p_i in (7) are known). Their procedure employs vector-at-a-time sampling. As before, let $X_{i1}, X_{i2}, \ldots$ be a sequence of observations from π_i. At stage j, we have the observations X_{ij}, $i = 1, \ldots, k$. Define

$$Y_{ij} = \begin{cases} 1 & \text{if } Y_{ij} \geq \max\limits_{r} Y_{rj} - D\sigma \\ 0 & \text{otherwise} \end{cases} \qquad i = 1, \ldots, k,$$

where D is given by (6). In other words, $Y_{ij} = 1$ if π_i is selected by Gupta's rule in (5) based on stage j observations ($n = 1$) and $Y_{ij} = 0$, otherwise.

Now, for any stage m, define $S_{im} = \sum_{j=1}^{m} Y_{ij}$ so that S_{im} has a binomial distribution $B(m, p_i)$ with parameters m and p_i (given by (7)). This fact is used by Barron and Gupta [3] in constructing their procedure. As we will see, this procedure continues sampling from all populations until the terminal decision is made with regard to all the populations.

Barron and Gupta [3], in fact, defined a class of procedures based on a pair of sequences of real numbers, $(\{b_m\}, \{c_m\})$, satisfying for $m \geq 1$ the following conditions: (i) $b_m \leq b_{m+1}$, $c_m \leq c_{m+1}$, (ii) $b_m < c_m$, (iii) $\lim_{m \to \infty} b_m = \infty$, and

(iv) $Pr\left\{\bigcap_{m=1}^{\infty}[b_m < S_{im} < c_m]\right\} = 0$ for all $i = 1,\ldots,k$. For each such pair, the Barron-Gupta rule is as follows.

Procedure $R_{NSS:BG}$: At stage m $(m = 1,2,\ldots)$, tag each untagged population π_i for which $S_{im} \notin (b_m, c_m)$; tag it "rejected" if $S_{im} \leq b_m$ and "accepted" if $S_{im} \geq c_m$. Stop sampling when all the populations are tagged. At termination, select all those populations that were tagged "accepted."

It should be noted that once a population π_i is tagged, it remains so irrespective of later changes in S_{im}. Barron and Gupta [3] have studied in detail several properties of this procedure including its performance compared with the single-stage procedure $R_{NSS:G}$.

5.2 Swanepoel-Geertsema Procedure, $R_{NSS:SG}$. This procedure is devised for selecting a subset containing the population with the largest θ_i assuming that the σ_i^2 are unknown and possibly unequal. It is a sequential procedure with no elimination employing vector-at-a-time sampling, and is based on constructing a selection sequence. For each $n \geq 1$, let B_n be a subset of the k populations defined by n observations from each. Any sequence $\{B_n\}$ is a selection sequence if

$$Pr\{\pi_{(k)} \in B_n \text{ for all } n \geq 1\} \geq P^*$$

for all $\underset{\sim}{\theta} \in \Omega$ where $\pi_{(k)}$ denotes the best population and $0 < P^* < 1$ is given.

Swanepoel and Geertsema [50] construct a selection sequence $\{B_n\}$ where $B_1 = \{\pi_1,\ldots,\pi_k\}$ and

$$B_n = \{\pi_r : \overline{X}_r(n) \geq \max_{1\leq i\leq k} \overline{X}_i(n) - s_{nr}h(k,P^*,n)\}$$

where $\overline{X}_i(n)$ is the mean of n observations from π_i, s_{nr} is an estimator of $\max_{i \neq r} \sqrt{\frac{\sigma_i^2+\sigma_r^2}{n}}$, and h is a constant depending on k, P^*, and n. The stopping time N is defined to be the first integer $n \geq 1$ such that $|B_n| \leq m$, where $|B_n|$ denotes the size of B_n and m is an integer chosen in advance with $1 \leq m \leq k-1$. At termination, we select the subset B_N.

In the unknown true configuration of $\underset{\sim}{\theta}$, let s denote the number of θ_i's equal to $\theta_{[k]}$. If $s \leq m$, then $N < \infty$ a.s., $|B_N| \leq m$, and B_N includes the best population with minimum probability P^*.

5.3 Gupta-Liang Procedure, $R_{NSS:GL}$. Gupta and Huang [23] proposed and studied two procedures based on log-likelihood ratios which can be applied to location and scale parameter cases. One of these two procedures is with elimination. Their goal is to select all mildly t best populations (i.e. those π_i's for which $\theta_i \geq \theta_{[k-t+1]} - \delta^*$ for a specified $\delta^* > 0$, in the location case).

Recently, Gupta and Liang [25] have considered a similar setup (with some slightly modified assumptions) and proposed a sequential procedure applicable to location and scale cases but with a modified goal. For the location case with $t = 1$, the Gupta-Huang goal is to select all good populations. The Gupta-Liang goal is to select a subset which includes the best population and at the same time excludes all that are not good. An event of selecting a subset consistent with this goal is denoted by $CS(\delta^*)$ [Note that $CS(\delta^*)$ is different from δ^*–CS].

For the normal means problem with a common known variance σ^2, let $X_{i1}, X_{i2}, \ldots$ be a sequence of independent observations from π_i, $i = 1, \ldots, k$. For $m \geq 1$, define $Y_{im} = \sum_{j=1}^{m} X_{ij}$. Let S_m denote the set of contending populations at the

beginning of stage m and $|S_m|$ denotes the size of S_m. We now define the Gupta-Liang procedure.

Procedure $R_{NSS:GL}$: Choose a δ_1 in the interval $(0, \delta^*/2)$. At stage m ($m = 1, 2, \ldots$), take one observation from each population in S_m. Include in S_{m+1} only those π_i's in S_m for which

$$\frac{\delta_1}{2}(Y_{rm} - Y_{im}) - \frac{m\delta_1^2}{4} < \log \frac{k-1}{1-P^*} \text{ for all } \pi_r \in S_m, r \neq i;$$

and eliminate all other π_i's from any further consideration. Now, label as good only those π_i's in S_{m+1} that have not been labeled so far and for which

$$\frac{\delta_1 + \delta^*}{2}(Y_{im} - Y_{tm}) + \frac{m(\delta^{*2} - \delta_1^2)}{4} \geq \log \frac{k-1}{1-P^*} \text{ for all } \pi_t \in S_{m+1}, t \neq i.$$

Stop sampling if either $|S_{m+1}| = 1$ or S_{m+1} does not contain any unlabeled population, and make the terminal decision: "Select all the populations in S_{m+1}"; otherwise, go to stage $m+1$.

It should be noted that a population is not labeled unless and until it qualifies to be called good. Once so labeled, it is not examined for labeling again. It is also possible that a labeled population is eliminated at a later stage. The populations that are selected are the ones which have been found to be good at some stage and which have survived elimination. The choice of δ_1 in $(0, \delta^*/2)$ assures that the procedure terminates with probability one. The procedure guarantees that the $PCS(\delta^*)$ is at least P^*. The question of an optimal choice of δ_1 is open.

6. SELECTION FROM BERNOULLI POPULATIONS: SS APPROACH

As in Section 4, $\pi_1, \ldots, \pi_k$ are Bernoulli populations with success probabilities $p_1, \ldots, p_k$, respectively. Gupta and Sobel [29] proposed and studied a single-stage

procedure based on n independent observations from each population. Let X_i denote the number of successes from π_i, $i = 1, \ldots, k$. The Gupta-Sobel procedure $R_{BSS:GS}$ is: Select π_i if and only if $X_i \geq \max_{1 \leq j \leq k} X_j - d$, where $d = d(k, n, P^*)$ is the smallest positive integer for which the P^*-requirement is satisfied.

Sequential procedures are important in practice when the cost of sampling is high or when the observations are scarce so that it is difficult to have the sample size needed by a fixed sample size procedure in order to achieve the desired level of the PCS. In the Bernoulli model, they have the added importance of ethical considerations when the experiment concerns comparisons among drugs; one would want a drug with a small success rate θ_i to be identified soon. Since a subset selection rule also serves as a screening procedure before selecting one of the drugs as the best, it makes sense to eliminate poor drugs rather quickly so that more observations can be used for the remaining ones.

Recently, Sanchez [46] considered a class of sequential procedures which take no more than n (common sample size in $R_{BSS:GS}$) observations from each population and result in the identical terminal decision as does $R_{BSS:GS}$. All the procedures in this class share the same stopping rule S^* and terminal decision rule T^* (to be defined later). An optimal procedure in this class is defined to be the one which minimizes the expected value of N, the total number of observations taken until termination. In order to determine an optimal procedure, we should consider procedures that take observations one-at-a-time. However, this turns out to be a difficult task (see Sanchez [46]). In this context, Sanchez [46] investigated a procedure which uses a modification of the so-called least-failures sampling rule of

the Bechhofer-Kulkarni procedure $R_{BIZ:BK}$ described in Section 4. Although this procedure is not optimal, it seems to perform well enough to be of practical interest. Sanchez has considered asymptotic [46] as well as small sample [47] performance of this procedure, the latter based on simulation.

We now complete our discussion by formally describing the modified least-failures procedure of Sanchez [46]. Let n and d be the common sample size and the constant of the Gupta-Sobel procedure $R_{BSS:GS}$. Observations are taken one-at-a-time. Let x_{im}, y_{im}, and n_{im} denote the number of successes, number of failures, and the total number of observations, respectively, from π_i through stage m. Let S_S and S_E denote the subsets of selected populations and of excluded populations, respectively, into which the populations are assigned possibly at each stage according to the following rule:

$$\begin{aligned} &\text{Assign } \pi_i \text{ to } S_S \text{ if } x_{im} + \min_{j \neq i} y_{jm} \geq n - d; \\ &\text{Assign } \pi_i \text{ to } S_E \text{ if } y_{im} + \max_{j} x_{jm} \geq n + d + 1; \\ &\text{No assignment is made otherwise.} \end{aligned} \tag{8}$$

Sanchez Procedure, $R_{BSS:S}$: Least-failures sampling is employed until for some π_i, $n_{im} = n$ and $y_{im} = \max_{1 \leq j \leq k} y_{jm}$ at which time this π_i is assigned to S_S. ¿From this stage on, additional observations are taken from π_j $(j \neq i)$ until the first stage when π_j can be assigned to S_S or S_E according to (8). Sampling is stopped when no population remains to be assigned. The terminal decision is: Select all the populations in S_S.

7. SELECTION FROM EXPONENTIAL FAMILY

The Bechhofer–Kiefer–Sobel [12] book provides a very comprehensive treatment (including loss functions) of indifference–zone selection for the exponential family. In this section, we discuss some recent results of Gupta and Miescke [26] and Liang [38] for selection from k populations belonging to a one-parameter exponential family. Liang's approach is classical with the goal of $CS(\delta^*)$, same as that of the Gupta-Liang procedure $R_{NSS:GL}$ described in Section 5. Gupta and Miescke [26] adopted a decision-theoretic approach to sequential selection. Their treatment includes multi-stage selection. They have obtained results for selection of subsets of random as well as fixed sizes.

7.1 Liang Procedure, $R_{EFSS:L}$. Let $\pi_1, \ldots, \pi_k$ be k populations where π_i has density $f(x|\theta_i)$, where

$$f(x|\theta) = c(\theta)\exp(\theta x)h(x),\ x \text{ real},$$

and $\theta \in \Theta$, an interval on the real line. For specified $\delta^* > 0$, any population π_i is defined to be good if $\theta_i \geq \theta_{[k]} - \delta^*$. Liang [38] considered the goal of selecting a subset which contains the best population and excludes any that is not good (same goal as that of $R_{NSS:GL}$ in Section 5). His sequential procedure with elimination is based on certain conditional likelihood functions and it achieves the P^*-requirement for $CS(\delta^*)$. The details are omitted here.

7.2 Gupta-Miescke Decision Theoretic Approach. Consider the one-parameter exponential family $\mathcal{F}$ given by

$$\mathcal{F} = \{c(\theta)\exp(\theta x)h(x),\ x \in R\}_{\theta\in\Theta}$$

where $\Theta \subseteq R$ is an interval. We consider the class $\mathcal{P}_I$ of permutation invariant sequential procedures with or without elimination, employing vector-at-a-time

sampling. Let $X_{i1}, X_{i2}, \ldots$ be a sequence of observations available to the experimenter from π_i (with associated parameter θ_i). At stage m ($m = 1, 2, \ldots$), let n_m observations be taken from the eligible populations. Let $W_{im} = \sum_{j=1}^{N_m} X_{ij}$, where $N_m = \sum_{j=1}^{m} n_j$, denote the sufficient statistic for θ_i, based on all observations from π_i through stage m, and let $W_m = (W_{1m}, \ldots, W_{km})$, $m = 1, 2, \ldots$.

For $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k) \in \Omega = \Theta^k$, $L_m(\Omega; t_1, \ldots, t_m, t_{m+1})$ denotes the loss incurred when the procedures stops at stage m with a record $\{t_1, \ldots, t_m, t_{m+1}\}$, where t_j, $j = 1, \ldots, m$, denotes the subset of $\{\pi_1, \ldots, \pi_k\}$ that is eliminated at stage j, and t_{m+1} denotes the subset finally selected at termination. Note that $\{t_1, \ldots, t_{m+1}\}$ is a disjoint decomposition of $\{\pi_1, \ldots, \pi_k\}$. It is assumed that: (a) L_m is permutation invariant, and (b) L_m increases if a record is changed so that a better population is eliminated before an inferior one.

A natural terminal decision, at stage m, selects only those populations among the noneliminated ones which yielded the largest values of W_{im}. Gupta and Miescke [26] have shown that between any two procedures which differ only in their terminal decisions, the procedure that uses a natural rule for terminal decision has a smaller risk.

It is reasonable to speculate that, within stages where a procedure with elimination does not stop, natural subset selections are optimal as in the case of terminal decisions. However, this has been proved by Gupta and Miescke [26] only in the case of multi-stage procedures with the sizes of the subsets selected at each stage fixed, under the assumption that $\mathcal{F}$ is strongly unimodal [i.e. exponential density is logconcave]. For additional comments, see Miescke [39].

8. CONCLUDING REMARKS

As pointed out earlier, we have not attempted to provide any sort of comprehensive survey of sequential selection procedures. We have discussed only a few of the selection procedures which are dealt with in the books mentioned in Section 1. These few procedures are included to make the discussion of recent results contextually clear. There are other problems of current interest which are not included here. For example, there is some interest in multinomial selection problems with truncation and curtailed sampling. There are several papers relating to multi-stage procedures; especially, two-stage procedures. These are not included here. Also, we have not discussed sequential procedures for selecting populations better than a standard or a control.

SUMMARY

This paper describes some sequential selection procedures for selecting the normal population having the largest mean, and for selecting the Bernoulli population having the largest success probability, with emphasis on recent developments. Both the indifference-zone and subset approaches are discussed. Some results for the exponential family including a decision-theoretic approach are also described. Specifically, the review of recent accomplishments include the development of truncated procedures, detailed comparisons of the performance characteristics of procedures for normal means, curtailment, adaptive sampling, and new selection goals.

ACKNOWLEDGEMENTS

This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS-8606964, DMS-8702620 at Purdue

University. The authors are thankful to the referee for valuable comments and to Professor TaChen Liang for helpful discussions.

REFERENCES

[1] Bahadur, R.R. (1950). On a problem in the theory of k populations, Ann. Math. Statist., 21: 362–375.

[2] Bahadur, R.R. and Robbins, H. (1950). The problem of the greater mean, Ann. Math. Statist., 21: 469–487. Correction, 22 (1951): 301.

[3] Barron, A.M. and Gupta, S.S. (1972). A class of non-eliminating sequential multiple decision procedures, Operations Research-Verfahren (R. Henn, H.P. Künzi and H. Schubert, eds.), Verlag Anton Hein, Meisenheim am Glan, West Germany, pp. 11–37.

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

[5] 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 unknown variance, Biometrika, 41: 170–176.

[6] Bechhofer, R.E. and Frisardi, T. (1983). A Monte Carlo study of the performance of a closed adaptive sequential procedure for selecting the best Bernoulli population, J. Statist. Comput. Simulation, 18: 179–213.

[7] Bechhofer, R.E. and Goldsman, D.M. (1987). Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean, Comm. Statist. Simulation Comput., 16: 1067–1092.

[8] Bechhofer, R.E. and Goldsman, D.M. (1987). Sequential selection procedures for multi-factor experiments involving Koopman-Darmois populations with additivity, Statistical Decision Theory and Related Topics - IV (S.S. Gupta and J.O. Berger, eds.), Vol. 2, Springer-Verlag, New York, pp. 3–22.

[9] Bechhofer, R.E. and Goldsman, D.M. (1988). Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean (II): 2-factor experiments with no interaction, Comm. Statist. Simulation Comput., 17: 103–128.

[10] Bechhofer, R.E. and Goldsman, D.M. (1988). A comparison of the performances of procedures for selecting the normal population having the largest mean when the variances are known and equal, Technical Report No. 789, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York. To appear in a special volume in honor of Ingram Olkin on his 65th birthday.

[11] Bechhofer, R.E. and Goldsman, D.M. (1988). Truncation of the Bechhofer–Kiefer–Sobel sequential procedure for selecting the largest normal mean (III): supplementary truncation numbers and resulting performance characteristics, Technical Report 815, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

[12] Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures (with special reference to Koopman-Darmois populations), The University of Chicago Press, Chicago.

[13] Bechhofer, R.E. and Kulkarni, R.V. (1982). Closed adaptive sequential proce-

dures for selecting the best of $k \geq 2$ Bernoulli populations, Statistical Decision Theory and Related Topics - III (S.S. Gupta and J.O. Berger, eds.), Vol 1, Academic Press, New York, pp. 61–108.

[14] Bechhofer, R.E. and Kulkarni, R.V. (1982). On the performance characteristics of a closed adaptive sequential procedure for selecting the best Bernoulli population. Sequential Anal., 1: 315–354.

[15] Büringer, H., Martin, H. and Schriever, K.-H. (1980). Nonparametric Sequential Selection Procedures, Birkhauser, Boston, Massachusetts.

[16] David, H.A. and Andrews, D.M. (1987). Closed adaptive sequential paired-comparison selection procedures, J. Statist. Comput. Simulation, 27: 127–141.

[17] Dudewicz, E.J. and Koo, J.O. (1982). The Complete Categorized Guide to Statistical Selection and Ranking Procedures, Series in Mathematical and Management Sciences, Vol. 6, American Sciences Press, Inc., Columbus, Ohio.

[18] Edwards, D. (1987). Extended-Paulson sequential selection, Ann. Statist., 15: 449–455.

[19] Fabian, V. (1962). On multiple decision methods for ranking population means, Ann. Math. Statist., 33: 248–254.

[20] Fabian, V. (1974). Note on Anderson's sequential procedures with triangular boundary, Ann. Statist., 2: 170–176.

[21] 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.

[22] Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules, Technometrics, 7: 225–245.

[23] Gupta, S.S. and Huang, D.-Y. (1975). On some parametric and nonparametric sequential subset selection procedures, Statistical Inference and Related Topics (M.L. Puri, ed.), Academic Press, New York, pp. 101–128.

[24] Gupta, S.S. and Huang, D.-Y. (1981). Multiple Decision Theory: Recent Developments, Lecture Notes in Statistics, Vol. 6, Springer-Verlag, New York.

[25] Gupta, S.S. and Liang, T. (1988). On a sequential subset selection procedure, Technical Report No. 88–23, Department of Statistics, Purdue University, West Lafayette, Indiana.

[26] Gupta, S.S. and Miescke, K.J. (1984). Sequential selection procedures - A decision-theoretic approach, Ann. Statist., 12: 336–350.

[27] Gupta, S.S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations, John Wiley & Sons, New York.

[28] Gupta, S.S. and Panchapakesan, S. (1985). Subset selection procedures: review and assessment, Amer. J. Math. Management Sci., 5: 235–311.

[29] Gupta, S.S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations, Contributions to Probability and Statistics (I. Olkin, S.G. Ghurye, W. Hoeffding, W.G. Madow and H.B. Mann, eds.), Stanford University Press, Stanford, California, Chapter 20, pp. 224–248.

[30] Hartmann, M. (1988). An improvement on Paulson's sequential ranking procedure, Sequential Analysis, 7: 363–372.

[31] Hoel, D.G. (1971). A method for construction of sequential selection procedures, Ann. Math. Statist., 42: 630–642.

[32] Jennison, C. (1983). Equal probability of correct selection for Bernoulli selection procedures, Comm. Statist. Theory Methods, 12: 2887–2896.

[33] Jennison, C. (1986). On the stochastic minimization of sample size by the Bechhofer–Kulkarni Bernoulli sequential selection procedure, Technical Report 677, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, Revised 1988.

[34] Jennison, C. and Kulkarni, R.V. (1984). Optimal procedures for selecting the best s out of k Bernoulli populations, Design of Experiments: Ranking and Selection (T.J. Santner and A.C. Tamhane, eds.), Marcel Dekker, New York, pp. 113–125.

[35] Kao, S.C. and Lai, T.L. (1980). Sequential selection procedures based on confidence sequences for normal populations, Comm. Statist. A - Theory Methods, 9: 1657–1676.

[36] Kelly, F.P. (1981). Multi-armed bandits with discount factor near one: the Bernoulli case, Ann. Statist., 9: 987–1001.

[37] Kulkarni, R.V. and Jennison, C. (1986). Optimal properties of the Bechhofer-Kulkarni Bernoulli selection procedure, Ann. Statist., 14: 298–314.

[38] Liang, T. (1988). On a sequential subset selection procedure for exponential family distributions, Technical Report No. 88-20C, Department of Statistics, Purdue University, West Lafayette, Indiana.

[39] Miescke, K.J. (1984). "Recent results on multi-stage selection procedures,"

Proceedings of the Seventh Conference on Probability Theory (Braşov, 1982), Editura Academiei Republicii Socialiste România, Bucharest, pp. 259–268.

[40] Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations, Ann. Math. Statist., 35: 174–180.

[41] Paulson, E. (1967). Sequential procedures for selecting the best one of several binomial populations, Ann. Math. Statist., 38: 117–123.

[42] Paulson, E. (1969). A new sequential procedure for selecting the best one of k binomial populations, Abstract, Ann. Math. Statist., 40: 1865–1866.

[43] Percus, O.E. and Percus, J.K. (1984). On the Bechhofer–Kulkarni stopping rule for sequential clinical trials, SIAM J. Appl. Math., 44: 1164–1175.

[44] Robbins, H. (1956). A sequential decision problem with a finite memory, Proc. Nat. Acad. Sci. U.S.A., 42: 920–923.

[45] Robbins, H., Sobel, M. and Starr, N. (1968). A sequential procedure for selecting the best of k populations, Ann. Math. Statist., 39: 88–92.

[46] Sanchez, S.M. (1987). A modified least-failures sampling procedure for Bernoulli subset selection, Comm. Statist. Theory Methods, 16: 3609–3629.

[47] Sanchez, S.M. (1987). Small-sample performance of a modified least-failures sampling procedure for Bernoulli subset selection, Comm. Statist. Simulation Comput., 16: 1051–1065.

[48] Sobel, M. and Huyett, M.J. (1957). Selecting the best one of several binomial populations, Bell System Tech. J., 36: 537–576.

[49] Stein, C. (1948). The selection of the largest of a number of means, Abstract,

Ann. Math. Statist., 19: 429.

[50] Swanepoel, J.W.H. and Geertsema, J.C. (1973). Selection sequences for selecting the best of k normal populations, S. Afr. Statist. J., 7: 11–21.

[51] Taheri, H. and Young, D.H. (1974). A comparison of sequential sampling procedures for selecting the better of two binomial populations, Biometrika, 61: 585–592.

[52] Tamhane, A.C. and Bechhofer, R.E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean, Comm. Statist. A – Theory Methods, 6: 1003–1033.

[53] Tamhane, A.C. and Bechhofer, R.E. (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): An improved PCS lower bound and associated tables, Comm. Statist. A - Theory Methods, 8: 337–358.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release, distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report #88-32C	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Purdue University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Department of Statistics West Lafayette, IN 47907	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-88-K-0170 and NSF Grant MDS-8606964
8c. ADDRESS (City, State, and ZIP Code) Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
ON SEQUENTIAL RANKING AND SELECTION PROCEDURES (Unclassified)

12. PERSONAL AUTHOR(S)
Shanti S. Gupta and S. Panchapakesan

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	July, 1988	25

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Open, closed, and truncated sequential procedures; indifference-zone; subset selection; normal; Bernoulli; exponential family; vector-at-a-time sampling; adaptive sampling; decision-theoretic

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper describes some sequential selection procedures for selecting the normal population having the largest mean, and for selecting the Bernoulli population having the largest success probability, with emphasis on recent developments. Both the indifference-zone and subset approaches are discussed. Some results for the exponential family including a decision-theoretic approach are also described.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Shanti S. Gupta	22b. TELEPHONE (Include Area Code) (317) 494-6031 / 22c. OFFICE SYMBOL