Proceedings of the 1983
Winter Simulation Conference
S. Roberts, J. Banks, B. Schmeiser (eds.)

RANKING AND SELECTION IN SIMULATION

David Goldsman
School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853

In this article, we discuss the branch of statistics known as ranking and selection. We introduce some common ranking and selection terminology and procedures. Additional references for more complicated procedures are given. Applications to simulation are discussed.

1. INTRODUCTION

Over the last thirty years, a large body of literature has been concerned with certain statistical problems of ranking and selection. Ranking and selection methods form a collection of procedures which attempt to answer the following questions (among others):

- Which of k competing populations (or policies or drugs, etc.) is the 'best'?
- Of the k competing populations, what are the t ($1 \leq t \leq k$) 'best' populations with (or without) regard to order?
- Can we find a (small) subset of the k populations which contains the 'best' population?
- Can we find a subset which contains the t 'best' populations?
- Which of the k populations are 'better' than a certain 'control' population?

'Best' is used with regard to that population parameter which the experimenter deems to be the most important; e.g., the population mean or variance. The problems of determining the population with the largest mean, the smallest mean, or the smallest variance could be of interest. Following are several practical problems in which ranking and selection procedures might be useful:

- Find which of k computer systems has the greatest availablity.
- Find the face of a biased die which will turn up most frequently.
- Find the anti-cancer drug which yields the highest five-year survival rate.
- Identify which drugs (if any) are better than a certain drug currently on the market.
- Determine the most precise bathroom scale.
- Find the weed killer that yields the agricultural plot with the fewest weeds.

When studying problems such as those described above, ranking and selection is generally more efficient in its use of observations than other methods of 'classical' statistics. Hence, money and time can be saved by using ranking and selection in the appropriate situations. We shall also see that these procedures are both intuitively appealing and simple to implement. Further, it is easy to use the procedures in the simulation environment. Thus, the simulation practitioner should consider these procedures when they would appear to be applicable.

The purpose of this paper is to familiarize the reader with some of the common ranking and selection terminology and procedures. Section 2 is concerned with the indifference zone approach in ranking and selection while section 3 is concerned with the subset approach. In section 4, we give references for additional procedures. A framework for using these procedures in simulations is given in section 5.

2. THE INDIFFERENCE ZONE APPROACH

To motivate this approach, due mainly to

CH1953-9/83/0000-0387 $01.00 © 1983 IEEE

Bechhofer (1954), let us consider a simple example. Suppose that we have k=5 populations, denoted as $\Pi_1, \Pi_2, \ldots, \Pi_5$. Suppose that, unknown to us, the underlying population means, $\mu_1, \mu_2, \ldots, \mu_5$ have values 1.00, 1.32, 10.99, 0.99, 1.00, respectively. We denote the ordered unknown μ_i's by $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$. Of course, we do not know the values of the $\mu_{[i]}$'s nor do we know how the $\mu_{[i]}$'s are paired with the Π_i's. We now suppose that our goal is to select as best that Π_i associated with $\mu_{[5]}$; i.e., we wish to select the population with the largest population mean. Ideally, we would correctly select Π_3. An incorrect choice might prove costly since Π_3 has by far the largest population mean. Thus, we would wish to select the correct population (Π_3 here) with high probability.

Generalizing to the case of k populations, $\Pi_1, \ldots, \Pi_k$, we wish to choose that Π_i which is associated with $\mu_{[k]}$. If the Π_i associated with $\mu_{[k]}$ is, in fact, chosen, we say that a correct selection (CS) has been made. Suppose that we further require that the probability of a correct selection, P(CS), be $\geq$ P*, whenever $\mu_{[k]} - \mu_{[k-1]} \geq \delta^*$, where we have pre-specified $\{\delta^*, P^*\}$. To avoid the trivial procedure of simply choosing a Π_i at random (in which case P(CS)=1/k), we require that $(1/k) < P^* < 1$. Also, we must have $\delta^* > 0$. We call $\{\mu \mid \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}$ the preference zone. The complement of the preference zone is called the indifference zone for the obvious reasons. The origin of the term 'indifference zone approach' should now be clear.

To illustrate an indifference zone procedure, we consider the normal means problem. Suppose there are k independent normal populations, $\Pi_1, \ldots, \Pi_k$, with unknown means, $\mu_1, \ldots, \mu_k$, and known, common variance, σ^2. Our goal is to select that one population which has the largest mean:

1. Specify $\{\delta^*, P^*\}$, with $\delta^* > 0$, $(1/k) < P^* < 1$.
2. Calculate $N = \lceil (c_{k,P^*} \cdot \sigma/\delta^*)^2 \rceil$, where c_{k,P^*} (a constant depending only on k and P*) is tabled in, for example, Bechhofer (1954) and $\lceil . \rceil$ is the 'ceiling' function.
3. Take N independent observations from each Π_i, i=1,...,k.
4. Let X_{ij} be the j-th observation from Π_i, i=1,...,k; j=1,...,N. Calculate the sample means:
$$\overline{X}_i = (1/N) \sum_{j=1}^{N} X_{ij}, \quad i=1,\ldots,k.$$
5. Choose as best that Π_i corresponding to the largest sample mean. (Certainly, this is an intuitively appealing selection rule.)

Note that this is a single-stage procedure; the common number of independent observations, N, from each Π_i is determined a priori by the experimenter's choice of $\{\delta^*, P^*\}$.

Table A consists of c_{k,P^*}-values taken from Bechhofer (1954) for certain choices of k and P*.

P* \ k	2	3	5	10
0.75	0.9539	1.4338	1.8463	2.2637
0.90	1.8124	2.2302	2.5997	2.9829
0.95	2.3262	2.7101	3.0552	3.4182
0.99	3.2900	3.6173	3.9196	4.2456

Table A: c_{k,P^*} for selected k and P*

As an example, let us assume that k=3 and σ^2=2. Table B consists of values of $N = \lceil (c_{k,P^*} \sigma/\delta^*)^2 \rceil$ for selected choices of $\{\delta^*, P^*\}$.

P* \ δ*	0.1	0.5	2.0	4.0
0.75	412	17	2	1
0.90	995	40	3	1
0.95	1469	59	4	1
0.99	2617	105	7	2

Table B: N for k=3, σ^2=2, and selected $\{\delta^*, P^*\}$

Finally, fix σ^2=2 and δ^*=0.1. Table C consists of values of N for various choices of k and P*.

P* \ k	2	3	5	10
0.75	182	412	682	1025
0.90	657	995	1352	1780
0.95	1083	1469	1867	2337
0.99	2165	2617	3073	3606

Table C: N for σ^2=2, δ^*=0.1, and selected k and P*

From the above tables, it is easy to see that N increases as k, σ^2, or P* increase and as δ^* decreases.

3. THE SUBSET APPROACH

We now discuss another powerful ranking and selection methodology; viz., the subset approach originally formulated by Gupta (1956). Consider an example similar to one in section 2: Unknown to us, the values of the underlying means of k=5 populations, $\Pi_1, \ldots, \Pi_5$ are μ_1=1.00, μ_2=1.32, μ_3=10.99, μ_4=0.99, and μ_5=10.96. These values are unknown to us. We are interested in determining which Π_i is associated with $\mu_{[5]}$, the largest mean.

Had we used the indifference zone approach, the 'correct' choice would have been Π_3. In the subset approach, 'good' choices would be only Π_3, only Π_5, or both of Π_3 and Π_5. In fact, the subset approach simply dictates that we select a non-empty subset of $\{\Pi_1, \ldots, \Pi_5\}$ which contains the 'best' population. The selected subset might contain only one population if, e.g., $\mu_{[k]} >> \mu_{[k-1]}$; on the other hand, the subset might contain all populations if $\mu_{[k]} - \mu_{[1]}$ is very small. In any event, the number of populations in the selected subset is a random variable.

Here, a correct selection (CS) is said to have been made if and only if that Π_i associated with $\mu_{[k]}$ is included in the selected subset. We wish to have $P(CS) \geq P^*$, where P^* $((1/k) < P^* < 1)$ is pre-specified by the user. Also, it is desired to choose as small a subset as possible (while maintaining $P(CS) \geq P^*$). This prevents us from always trivially selecting all of the Π_i's, although, as noted above, such a selection is possible.

Note that an indifference zone is not specified in this approach. However, there is a trade-off between the indifference zone and subset formulations: We need only specify P^* in the subset approach, but the number of Π_i's in the selected subset is unknown a priori.

Consider the same normal means problem as in the previous section. That is, $\Pi_i \sim N(\mu_i, \sigma^2)$, where σ^2 is known and common to all k populations. We wish to select the Π_i associated with $\mu_{[k]}$:

1. Specify P^* and n.
2. Take n independent observations from each Π_i, i=1,...,k.
3. Calculate $\overline{X}_i = (1/n) \sum_{j=1}^{n} X_{ij}$, i=1,...,k.
4. Include Π_i, i=1,...,k, if and only if:

   $\overline{X}_i \geq \max_{1 \leq j \leq k} \overline{X}_j - (c_{k,P^*} \sigma/\sqrt{n})$, where c_{k,P^*} is again from Bechhofer (1954).

Thus, we see that although the indifference zone and subset approaches are different, they are nevertheless related. This relationship is not necessarily as close for more complicated ranking and selection problems.

We point out that Santner (1973,1975) and Gupta and Santner (1973) formulated the so-called restricted subset approach. The indifference zone and subset approaches can be viewed as special cases of the restricted subset approach. For more details, the reader should see the above references or Gupta and Panchapakesan (1979).

All of these procedures are intuitively appealing and easy to use. Indeed, the only work in implementing these procedures (aside from data collection) lies in looking up certain constants in tables. We do not advocate one of the three methodologies over the others. The experimenter should use whichever approach would appear to be appropriate for the problem at hand.

4. OTHER PROCEDURES

4.1 Normal Means

As discussed previously, Bechhofer (1954) and Gupta (1956) did pioneering work on the normal means problem. Tamhane and Bechhofer (1977,1979) give a two-stage procedure for finding the normal population with the largest mean when the variances of the Π_i's are all equal to a known common value. This procedure is superior to that given in Bechhofer(1954) in that the total number of observations required in this two-stage procedure is always less than that required by the single-stage procedure. This savings is a result of the fact that, after the first stage, the two-stage procedure eliminates from consideration populations indicated as being 'inferior'. We only sample from the non-eliminated populations in the second stage. Sequential (or multi-stage) procedures are given in Paulson (1964) and Bechhofer, Kiefer, and Sobel (1968).

Bechhofer, Dunnett, and Sobel (1954) give a two-stage procedure for finding the normal population with the largest mean when the variances are common and unknown. In the first stage, we sample from the Π_i's in order to obtain an estimate of the variance. In the second stage of this procedure, the estimate of the variance is used to determine how many additional observations must be taken from each of the populations in order to guarantee the indifference zone probability requirement. A two-stage elimination procedure for this case of common unknown variance is given by Gupta and Kim (1982). Gupta and Kim's procedure requires fewer observations (on the average) than does that of Bechhofer, Dunnett, and Sobel.

Dudewicz and Dalal (1975) and Rinott (1978) give two-stage procedures for the normal means problem in which the values of the variances of the Π_i's are completely unknown.

4.2 Normal Variances

Suppose that the researcher is interested in selecting that normal population which has the smallest population variance. Bechhofer and Sobel (1954) and Gupta and Sobel (1962) address this problem.

4.3 Multivariate Normal

Frequently, the normal populations with which the statistician is working are not independent. Thus, it is worthwhile to consider the problem of finding the multivariate normal component with the largest mean. This problem is of particular import to the simulator since he may use variance reduction techniques which induce (positive) correlation amongst all k simulated systems. It turns out that such correlation induction, coupled with the use of the correct multivariate normal ranking and selection procedure, can

decrease the number of observations needed in an experiment. Some multivariate normal procedures are given by Gupta, Nagel, and Panchapakesan (1973) and Gnanadesikan (1966).

4.4 Bernoulli Parameter

Suppose that we have k drugs, $\Pi_1, \ldots, \Pi_k$, and that the underlying cure rates are $p_1, \ldots, p_k$, respectively. We are interested in determining which Π_i is associated with the largest of the p_i's. We can treat this problem as that of finding the Bernoulli population with the largest probability of 'success'. Fundamental papers concerning this topic are Sobel and Huyett (1957), which uses a single-stage indifference zone approach, and Gupta, Huyett, and Sobel (1957), which uses a single-stage subset procedure. Recently, Bechhofer and Kulkarni (1981,1982), Bechhofer and Frisardi (1983), and Kulkarni and Jennison (1983) have given an extremely efficient sequential procedure for this problem.

4.5 Multinomial Parameter

Here, we attempt to find the multinomial cell with the largest underlying probability. The procedures which deal with this problem generally call for the user to take observations from the multinomial system in question until one cell has 'significantly' more observations than the other cells. That cell is then declared to be the cell with the largest underlying probability.

Bechhofer, Elmaghraby, and Morse (1959), Cacoullos and Sobel (1966), Bechhofer, Kiefer, and Sobel (1968), Alam (1971), and Ramey and Alam (1979) give such procedures, the latter being the most parsimonious with observations. Some errors in Ramey and Alam's tables are corrected by Bechhofer and Goldsman (1983). Gupta and Nagel (1967) study the problem via the subset approach.

Now, consider k arbitrary populations $\Pi_1, \ldots, \Pi_k$. We wish to select that population which has the highest probability of being the 'best', where 'best' is defined by the user. Suppose Π_i has probability p_i of being the 'best', i=1,...,k, where $\sum p_i = 1$. We will think of Π_i as corresponding to the i-th multinomial cell of a k-nomial distribution, i=1,...,k. Suppose we take an observation from each of the Π_i's. (This is called taking observations vector-at-a-time.) Clearly, of these k observations in the vector, one will be the 'best' according to some criterion of goodness. (Use randomization, if necessary.) Award one 'success' to the multinomial cell corresponding to the Π_i from which the 'best' observation came. Award the other cells zero 'successes'. We can thus think of the problem of sampling from k arbitrary populations as one of sampling from one k-nomial distribution.

Therefore, the multinomial problem is actually equivalent to finding that one of k arbitrary populations which has the highest probability of being the 'best' (see also Bechhofer and Sobel (1958)). Hence, any multinomial procedure is, in a sense, a nonparametric procedure. This has tremendous ramifications for the practitioner since, in a simulation study of k systems, nothing at all may be known about the underlying distributions of the systems.

4.6 General References

Unfortunately, only a limited number of texts have been written on the subject of ranking and selection. The first, Bechhofer, Kiefer, and Sobel (1968), is a mathematically sophisticated monograph that contains many general sequential procedures.

Gibbons, Olkin, and Sobel (1977) is a relatively easy book to read. It is meant only as an introductory text. The discussion of the subset approach is very limited. The reader should see the review of this book given in Bechhofer (1980).

An all-purpose book is that of Gupta and Panchapakesan (1979). Both the indifference zone and subset formulations are thoroughly treated. The text contains an excellent bibliography.

Dudewicz and Koo (1982) contains a comprehensive bibliography of the ranking and selection literature.

The thesis of Raskopf (1982) provides a very clear summary of several of the normal means procedures. Each procedure cited in this thesis is explicitly stated. Numerical examples are also given.

5. RANKING AND SELECTION IN SIMULATION

We now describe how ranking and selection techniques can be used by simulation practitioners. Of course, in real-life simulation problems, the data are rarely normally distributed; however, many of the aforementioned ranking and selection procedures are robust against the normality assumption. A more serious problem is that simulated outputs are frequently serially correlated. That is, adjacent output data from a simulation might not be independent. It might appear that this difficulty immediately makes the use of ranking and selection procedures an impossibility for the simulator. Such is not the case.

Consider a long run of output data. Under very mild conditions (see the description of 'phi-mixing' in Schruben (1982b) for further details), the data at the beginning of the run are virtually independent of the data at the end of the run. In fact, if a run is made that is sufficiently long, the data can be divided into nearly independent 'batches'. For example, in a run of 10000 observations, the experimenter may divide the output into ten batches, each with a batch size of 1000 observations. With the exception of the observations close to the ends of adjacent batches, the batches can be thought of as virtually independent of each other (assuming again that the run is sufficiently long).

Now, take the sample average of all of the observations in a particular batch - this is called the batched mean of that batch. If the

data are stationary (i.e., the data have the same underlying distribution), it can be shown that for sufficiently large batches, the batched means are approximately independent and normally distributed (although we do not necessarily know their mean and variance). We can thus treat each of the batched means as a single independent normal observation; the 'usual' ranking and selection procedures can now be applied to these batched means. The performances of k different simulated systems can be compared by using as normal independent observations the batched means of the k corresponding simulation runs.

Note that the same arguments can be applied to what is called the 'method of replications'. For additional material concerning the methods of batched means and replications, see Law and Kelton (1982).

An example

We now give a simple example: Consider k simulated systems, $\Pi_1, \ldots, \Pi_k$, with underlying means $\mu_1, \ldots, \mu_k$, none of which is known a priori. The goal is again to select that Π_i corresponding to $\mu_{[k]}$. Suppose that we can choose a batch size m for each of the k systems such that the batched means are approximately independent and normal. We will deal with three progressively more general cases concerning the variances of the normal batched means; namely, the variances are:

(a) the same for all k populations and known,

(b) the same for all k populations and unknown, and

(c) completely unknown.

Note that (c) would most likely be the case applicable to an actual simulation comparison study of k systems.

Finally, suppose for convenience that there is no initialization bias present in the output of any of the k systems. I.e., the output from a particular system is stationary, albeit serially correlated. (See Schruben (1982a) for a survey dealing with the problem of initialization bias.)

Case (a): Variances of batched means are common and known ($= \sigma^2$, say).

We use an indifference zone procedure directly analogous to that given in section 2:

1. Specify $\{\delta^*, P^*, m\}$, where m is the batch size.
2. Calculate $N = \lceil (c_{k,P^*}\, \sigma/\delta^*)^2 \rceil$, as before.
3. Take mN observations from each Π_i, i=1,...,k. Let X_{ij} denote the j-th observation from Π_i, i=1,...,k; j=1,...,mN.
4. Calculate $\bar{X}_i = (1/mN) \sum_{j=1}^{mN} X_{ij}$, i=1,...,k.
5. Select as the population with the largest mean that Π_i corresponding to the largest $\bar{X}_i$.

Again, we remark that the assumptions for the above illustrative case may be somewhat unrealistic.

We now relax the assumptions slightly;

Case (b): Variances of batched means are common but unknown.

We use a procedure analogous to that of Gupta and Kim (1982). This two-stage elimination procedure combines the indifference zone and subset approaches. In the first stage, 'inferior' populations are eliminated; the remaining populations receive additional sampling in the second stage, where a winner is finally chosen.

First Stage:

1.1 Specify $\{\delta^*, P^*, m, n_0\}$, where m is the batch size and n_0 is the number of batches which we will sample in the first stage.

1.2 Again, let X_{ij} be the j-th observation from Π_i, i=1,...,k; j=1,2,... Calculate the (first stage) batched means:

$\bar{X}_{it} = (1/m) \sum_{j=m(t-1)}^{mt} X_{ij}$, where

$\bar{X}_{it}$ is the t-th (first stage) batched mean from Π_i, i=1,...,k; t=1,...,n_0.

1.3 Calculate the first stage sample means:

$\bar{X}_i^{(1)} = (1/mn_0) \sum_{j=1}^{mn_0} X_{ij} = (1/n_0) \sum_{t=1}^{n_0} \bar{X}_{it}$, i=1,...,k.

1.4 Calculate the first stage 'pooled' variance:

$S^2 = (k(n_0-1))^{-1} \sum_{i=1}^{k} \sum_{t=1}^{n_0} (\bar{X}_{it} - \bar{X}_i^{(1)})^2$.

1.5 Find the subset I of {1,...,k} such that

$I = \{i \mid \bar{X}_i^{(1)} \geq \max_{1 \leq j \leq k} \bar{X}_j^{(1)} - ((dS/\sqrt{n_0}) - \delta^*)^+\}$,

where d is to be found in Gupta and Kim (1982) and where $y^+ = \max(0,y)$ for all y.

1.6 If there is only one element, say r, in I, then stop and declare that Π_r has the largest mean. Otherwise, continue to the second stage.

Second Stage:

2.1 Calculate $N = \max\{n_0, \lceil (hS/\delta^*)^2 \rceil\}$, where h is to be found in Gupta and Kim (1982).

2.2 For each Π_i such that $i \in I$, take $m(N-n_0)$ additional observations.

2.3 For each Π_i such that $i \in I$, compute the

overall sample mean:

$$\overline{X}_i = (1/mN) \sum_{j=1}^{mN} X_{ij}.$$

2.4 Declare that Π_i associated with the maximum $\overline{X}_j$ as having the largest mean.

Note that aside from taking data and making some easy calculations, the above procedure requires only that we look up two constants.

Finally, we state yet a more general procedure:

Case (c): Variances of the batched means are completely unknown.

We now adopt a two-stage indifference zone procedure from Dudewicz and Dalal (1975). In the first stage, we estimate the variances of the batched means from each Π_i. These estimates are then used to allocate 'efficiently' observations in the second stage.

First Stage:

1.1 Specify $\{\delta^*, P^*, m, n_0\}$.

1.2 As before, calculate the (first stage) batched means, $\overline{X}_{it}$, $i=1,\ldots,k$; $t=1,\ldots,n_0$.

1.3 As before, calculate the (first stage) sample means, $\overline{X}_i^{(1)}$, $i=1,\ldots,k$.

1.4 Calculate the first stage sample variances:

$$S_i^2 = (n_0-1)^{-1} \sum_{t=1}^{n_0} (\overline{X}_{it}-\overline{X}_i^{(1)})^2, \quad i=1,\ldots,k.$$

Second Stage:

2.1 Calculate $n_i = \max\{n_0+1, \lceil (S_i H/\delta^*)^2 \rceil\}$, $i=1,\ldots,k$, where H is to be found in Dudewicz and Dalal (1975).

2.2 Take $m(n_i - n_0)$ more observations from each Π_i.

2.3 Calculate the additional batched means from each Π_i:

$\overline{X}_{it}$, $i=1,\ldots,k$; $t=n_0+1,\ldots,n_i$.

2.4 For each population, calculate the 'pseudo-average':

$\overline{Y}_i = \sum_{t=1}^{n_i} a_{it}\overline{X}_{it}$, where a formula for the a_{it}'s can be found in Dudewicz and Dalal.

2.5 Declare that Π_i associated with the largest $\overline{Y}_j$ as having the largest mean.

Once more, it turns out that no tedious calculations are required in order to implement the procedure.

If one is unwilling even to make assumptions about the approximate normality of the batched means, the obvious extension of a nonparametric procedure might be useful (e.g., the multinomial procedure of Ramey and Alam (1979)). Thus, the point of this section is that all ranking and selection procedures can be directly extended for use in the simulation environment.

6. CONCLUSIONS

In this article, we have introduced the reader to basic ranking and selection terminology and procedures. The simplicity and appeal of these procedures are to be stressed. These procedures have many applications in computer simulation studies. Further, the simulator will find that they are easy to implement. It is hoped that the reader will take advantage of this important branch of statistics.

ACKNOWLEDGEMENTS

The author expresses his appreciation to Professors Robert Bechhofer and Lee Schruben at Cornell University for encouraging this line of study.

REFERENCES

Alam, K. (1971). On selecting the most probable category. Technometrics 13, 843-850.

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

Bechhofer, R. E. (1980). A book review of Gibbons, Olkin, and Sobel (1977). J. Amer. Statist. Assoc. 75, No. 371, 751-756.

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, 102-119.

Bechhofer, R. E., and Frisardi, T. (1983). A Monte Carlo study of the performance characteristics of a closed adaptive sequential procedure for selecting the best Bernoulli population. To appear in the J. of Statist. Computation and Simulation.

Bechhofer, R. E., and Goldsman, D. (1983). Further light on the Ramey-Alam sequential procedure for selecting the most probable multinomial event. In preparation.

Bechhofer, R. E., Kiefer, J., and Sobel, M. (1968). Sequential Identification and Ranking Procedures. The Univ. of Chicago Press, Chicago.

Bechhofer, R. E., and Kulkarni, R. (1981). Closed adaptive sequential procedures for selecting the best of k≥2 Bernoulli populations. Proceedings of the Third Purdue Symposium on Statistical Decision Theory and Related Topics (ed. by S. S. Gupta and

J. Berger). Academic Press, New York, I, 61-108.

Bechhofer, R. E., and Kulkarni, R. (1982). On the performance characteristics of a closed adaptive sequential procedure for selecting the best Bernoulli populations. Comm. in Statist. - Sequential Analysis, I(4), 315-354.

Bechhofer, R. E., and Sobel, M. (1954). A single-sample multiple-decision procedure for ranking variances of normal populations. Ann. Math. Statist. 24, 273-289.

Bechhofer, R. E., and Sobel, M. (1958). Nonparametric multiple-decision procedures for selecting that one of k populations which has the highest probability of yielding the largest observation (prelim. report). Abstract. Ann. Math. Statist. 29, p. 325.

Dudewicz, E. J., and Dalal, S. R. (1975). Allocation of observations in ranking and selection with unequal variances. Sankhya Ser. B 37, 28-78.

Dudewicz, E. J., and Koo, J. O. (1982). The Complete Categorized Guide to Statistical Selection and Ranking Procedures. American Sciences Press, Inc., Columbus, Ohio.

Gibbons, J. D., Olkin, I., and Sobel, M. (1977). Selecting and Ordering Populations. John Wiley, New York.

Gnanadesikan, M. (1966). Some Selection and Ranking Procedures for Multivariate Normal Populations. Ph. D. Thesis. Dept. of Statist., Purdue Univ., W. Lafayette, Ind.

Gupta, S. S. (1956). On a Decision Rule for a Problem in Ranking Means. Ph. D. Thesis (Mimeo. Ser. No. 150). Inst. of Statist., Univ. of N. Carolina, Chapel Hill.

Gupta, S. S., Huyett, M. J., and Sobel, M. (1957). Selection and ranking problems with Binomial populations. Trans. Amer. Soc. Qual. Contr., 635-644.

Gupta, S. S., and Kim, W. C. (1982). A two-stage elimination type procedure for selecting the largest of several normal means with a common unknown variance. Tech. Report #82-27, Dept. of Statist., Purdue Univ., W. Lafayette, Ind.

Gupta, S. S., and Nagel, K. (1967). On selection and ranking procedures and order statistics from the multinomial distribution. Sankhya Ser. B 29, 1-34.

Gupta, S. S., Nagel, K., and Panchapakesan, S. (1973). On the order statistics from equally correlated normal random variables. Biometrika 60, 403-413.

Gupta, S. S., and Panchapakesan, S. (1979). Multiple Decision Procedures. John Wiley, New York.

Gupta, S. S., and Santner, T. J. (1973). On selection and ranking procedures-a restricted subset selection rule. Proceedings of the 39th Session of the International Statistical Institute, Vol. 45, Book 1, 409-417.

Gupta, S. S., and Sobel, M. (1962). On selecting a subset containing the population with the smallest variance. Biometrika 49, 495-507.

Kulkarni, R., and Jennison, C. (1983). Optimal properties of the Bechhofer-Kulkarni Bernoulli selection procedure. Tech. Report #600, School of OR & IE, Cornell Univ., Ithaca, NY.

Law, A. M., and Kelton, W. D. (1982). Simulation Modeling and Analysis. McGraw-Hill, New York.

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

Ramey, J. T., Jr., and Alam, K. (1979). A sequential procedure for selecting the most probable multinomial event. Biometrika 66, 171-173.

Raskopf, B. J. (1982). A Survey of Statistical Decision Procedures for Selecting the Population with the Largest Normal Mean. M. S. Thesis, School of OR & IE, Cornell Univ., Ithaca, New York.

Rinott, Y. (1978). On two-stage procedures and related probability-inequalities. Commun. Statist.-Theor. Meth. A8, 799-811.

Santner, T. J. (1973). A Restricted Subset Selection Approach to Ranking and Selection Problems. Ph. D. Thesis (Mimeo. Ser. #318). Dept. of Statist., Purdue Univ., W. Lafayette, Ind.

Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problems. Ann. Statist. 3, 334-349.

Schruben, L. (1982a). Simulation initialization bias: the current state. Proceedings of the IMACS Conference, Aug., 1982, Montreal, Canada.

Schruben, L. (1982b). Confidence interval estimation using standardized time series. Tech. Report #518, School of OR & IE, Cornell Univ., Ithaca, New York.

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

Tamhane, A. C., and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist.-Theor. Meth. A6, 1003-1033.

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. Commun. Statist.-Theor. Meth. A8, 337-358.