The use of multiple ranking procedures to analyze simulations of management systems

Kleijnen, J.P.C.; Naylor, T.H.; Seaks, T.G.

Published in:
Management Science

Publication date:
1972

MANAGEMENT SCIENCE
Vol. 18, No. 6, February, 1972
Printed in U.S.A.

THE USE OF MULTIPLE RANKING PROCEDURES TO ANALYZE SIMULATIONS OF MANAGEMENT SYSTEMS: A TUTORIAL*†

JACK P. C. KLEIJNEN,‡ THOMAS H. NAYLOR§ AND TERRY G. SEAKS§

This paper describes the use of multiple ranking procedures to analyze data generated from computer simulation experiments with models of management systems. After outlining the rationale for the use of multiple ranking procedures with computer simulation experiments and defining some basic terminology, we examine several specific multiple ranking procedures. Careful attention is given to the assumptions underlying the different multiple ranking procedures and the extent to which these assumptions are satisfied by the data generated by simulation experiments. An example model is included to illustrate the applicability of multiple ranking procedures to simulation experiments in management science.

Introduction

Multiple ranking procedures appear to offer a useful approach to the problem of analyzing data from computer simulation experiments with models of management systems. In this paper we begin by describing the multiple ranking problem and its applicability to simulation. We focus on the assumptions underlying specific procedures and the extent to which these assumptions may or may not be violated when used with computer simulation experiments. Then we employ a multiple ranking procedure to analyze a simulation model of a queueing system and give a detailed example of how such a procedure is actually applied. Lastly, we discuss several modifications of existing multiple ranking procedures. The objective of these modifications is to circumvent some of the difficulties encountered when one or more of the assumptions underlying existing multiple ranking procedures is violated.

Multiple Ranking Procedures Defined

Frequently, the objective of computer simulation experiments with models of management systems is to find the "best" policy or decision rule. Conventional tests of homogeneity, such as the F-test that k population means are equal, or Bartlett's test that k population variances are equal, do not provide the decision-maker with the type of information that is necessary to make decisions aimed at selecting the best plan or policy—e.g., the policy which leads to the largest profit for a business firm.

Multiple ranking procedures represent a direct approach to the solution of the problem of ranking alternatives based on the empirical results obtained from some type of experiment whether it be a real world experiment or a computer simulation experiment [29]. The best rank of a set of managerial policies or decision rules is simply the ranking of the sample means associated with the given policies or decision rules. Because of random error, however, sample rankings may yield incorrect results.

* Received May 1970; revised June 1971.

† This research was supported by National Science Foundation Grant GS-2981, the Netherlands Organisation for the Advancement of Pure Research Z.W.O., and a grant from the I.B.M. Corporation. The authors are indebted to Professor Thomas H. Wonnacott of the University of Western Ontario for a number of helpful comments.

‡ Tilburg School of Economics.

§ Duke University.

With what probability can we say that the rank of a particular sample mean represents the true rank of the corresponding population mean? It is this question which multiple ranking procedures attempt to answer. Multiple ranking procedures may also be used to rank population variances, if the data have been transformed by some appropriate transformation such as a logarithmic transformation. Naylor, Wertz, and Wonnacott [29], [31] have used logarithmic transformations to rank the variances of national income generated by a simulation experiment with an econometric model. The recent book by Bechhofer, Kiefer, and Sobel [6, p. 273] describes a direct approach for ranking variances.

Historically, multiple ranking procedures can be traced back to the work of Bahadur [1] in 1950. He was one of the first statisticians to recognize the multiple ranking problem and to contribute to the theory of k sample problems.

Among the early contributions to the literature in this field were the papers by Mosteller [27] and Mosteller and Tukey [28] on *slippage procedures*. Mosteller's [27] procedure is a very simple nonparametric technique which derives its name from the fact that it attempts to tell when a sample mean is really the largest one and when it merely appears this way because it has "slipped." On the basis of our empirical results with this procedure [29], [30], we concur with Mosteller's criticism of the procedure as lacking a desirable level of power. The power of Mosteller's slippage procedure has also been discussed by Bofinger [9], Karlin and Truax [22], Kudo [24], and Paulson [32], [33].

In 1953 Bechhofer [2] proposed a path breaking procedure for selecting on the basis of sample means a single population from among k populations and guaranteeing the probability P^* that the selected population is the largest provided some other condition on the parameters is satisfied. Since the terminology used by Bechhofer [2] is basic to many multiple ranking procedures, it will be helpful if we define some of the terms which will be used throughout the remainder of this paper.

We want to choose as the "best" population the population with the highest population mean. We want to be assured with probability P^* that we have indeed selected the correct population. If the best population differs only slightly from the next best population, then (because of random error) extremely large samples are required to detect such a small difference. However, our concern over not detecting the best population when the difference between the best and the next best population is small generally will not be as great as it would be if this difference were large. Therefore, we want to take that number of observations from each population so that the probability of correct selection (CS) is at least P^*, given that the difference δ between the best and the next best population is at least δ^*. In other words, δ^* is the smallest value worth detecting, and P^* is the smallest acceptable value of the probability of achieving the desired detection. This may be stated formally as

$$\Pr(CS \mid \delta \geqq \delta^*) \geqq P^*, \tag{1}$$

where δ^* and P^* are specified by the decision-maker.

If all but one of the populations have a mean equal to the mean of the "next best" population, then the means are said to be in "least favorable configuration." If the means are not in this configuration, then the criterion given by (1) is more easily satisfied. Since we do not know the true configuration, the criterion must be satisfied for the least favorable configuration [2, p. 23].

With some of the multiple ranking procedures it is possible to find not only the best plan but also the following (and others unlisted):

1. The "best two" plans *without* regard to order.

2. The "best two" plans *with* regard to order.
3. The "best three" plans *without* regard to order.
4. The "best," "second best," "third best," etc. plan.

All of the multiple ranking procedures assume *independent observations* and most of them assume normally distributed observations. In describing the existing multiple ranking procedures it is convenient to divide them into two groups—(1) procedures assuming known variances and (2) procedures assuming unknown variances.

Procedures Assuming Known Variances

Bechhofer [2]

In 1954 Bechhofer developed a procedure for determining the number of observations required for selecting a single population and guaranteeing with probability P^* that the selected population is "best" provided the indifference quantity δ^* is specified. Bechhofer's procedure assumes *normality* and *known variances* σ_i^2. It is easy to show that the number of observations to be taken from the ith population is given by

$$N_i = (\sigma_i d/\delta^*)^2, \tag{2}$$

where d is tabulated in [2] and increases with k and P^*. Bechhofer has also demonstrated the use of his procedure with *factorial designs* (experiments with more than one factor), without interactions. He shows that a factorial design is more efficient than separate one-factor experiments [2, pp. 25–28]. If there are interactions, then it is nonsensical to search for the best population with regard to one factor, and the best population with regard to a second factor, etc. These interactions force us to treat the experiment as a single-factor experiment [3, p. 426].

Somerville [45]

A procedure has been proposed by Somerville which assumes that the k populations of interest to the decision-maker are *identical except for their means*, the *sample means are normally distributed*, and the *common variance* σ^2 *is known*. Instead of specifying P^* and δ^*, Somerville formulates *loss functions* which take into consideration the amount of use to be made of the result, the cost of making an incorrect decision, and the cost of sampling [45, p. 402]. A sample size which minimizes the maximum expected loss (i.e., a minimax sample size) is then calculated. Somerville indicates that using an estimate of σ will not seriously affect the results. Somerville's procedure has been generalized by Shirafuji [44] for the r-way classification.

Dunnett [13]

Dunnett's multiple ranking procedure assumes that the "population means $\mu_1, \mu_2, \cdots, \mu_k$ have a joint *a priori* normal distribution with means $U_1, U_2, \cdots, U_k$ known up to an additive constant, and with equal variances and equal covariances whose values are known" [13, p. 1]. Dunnett also treats the special case of prior independent normal distributions. Results are given by Dunnett for determining the number of observations n for the multiple ranking problems of Bechhofer [2] and Somerville [45].

The practical application of Dunnett's procedure is limited since certain multivariate normal integrals required by the procedure have not been tabulated for $k > 2$. However, tables are available for linear loss functions of the type proposed by Somerville [45]. The usefulness of Dunnett's procedure is subject to question by the following statement by Dunnett, "In practice . . . [the] values [of U_i] may be so poorly

determined that the experimenter may prefer to ignore them and base his decision entirely on the experimental data," [13, p. 4].

Two Sequential Procedures [6], [36]

Ramberg [39] has reported on a Monte Carlo study aimed at evaluating the performance characteristics of two sequential procedures for selecting the population with the largest mean from k *normal populations* when the population *variances* are *known* and *equal*. The two sequential procedures are Bechhofer, Kiefer, and Sobel's [6, pp. 259, 265] procedure and Paulson's [36] procedure. Both of these procedures were compared with Bechhofer's [2] fixed sample size procedure on the basis of the expected number of *stages* and the expected *total number of observations* required to terminate experimentation for the sequential procedures. Ramberg defines the number of *stages* as the total number of vectors of observations taken by a procedure until experimentation ceases and the *total number of observations* as the sum of all observations taken from all populations until experimentation ceases [39, pp. 7–8]. A stage consists of one observation from each of the populations still being sampled. On the basis of these two variables Ramberg concludes that,

> Since both of the sequential procedures guarantee a higher probability of a correct selection than P^*, the two sequential procedures are adjusted so that they exactly guarantee P^*. The performance characteristics of these adjusted procedures are then compared, and the Paulson procedure improves (sometimes markedly) relative to the Bechhofer-Kiefer-Sobel procedure [39].

Bechhofer, Kiefer, and Sobel [6, p. 264]

Bechhofer, Kiefer, and Sobel have recently developed a *sequential* procedure for normal populations with *unequal* but *known* variances σ_i^2. In this procedure, at each stage of experimentation r_i observations are taken from population π_i $(i = 1, \cdots, k)$, where the r_i's are the smallest integers satisfying

$$\frac{\sigma_1^2}{r_1} = \frac{\sigma_2^2}{r_2} = \cdots = \frac{\sigma_k^2}{r_k} = c^2, \tag{3}$$

where c is an arbitrary constant. Then the probability of a correct selection is guaranteed, if we redefine δ. The Bechhofer-Kiefer-Sobel procedure guarantees

$$\Pr\,(CS \mid \delta \geqq \delta^*) \geqq P^*, \tag{4}$$

where δ is now defined as

$$\delta = \frac{\mu_{[k]} - \mu_{[k-1]}}{c} \tag{5}$$

and $\mu_{[k]}$ and $\mu_{[k-1]}$ are the best, i.e., the largest, and the second largest population mean respectively, and c follows from (3).

Note that if all variances σ_i^2 are equal to, say, σ^2, then $r_i = 1$ and $c^2 = \sigma^2$ and the procedure is simplified considerably. It was this simple procedure that was compared with Paulson's procedure by Ramberg [39].

Bechhofer, Kiefer, and Sobel list many other multiple ranking procedures, e.g., for ranking variances of normal populations with unknown and unequal means, for ranking exponential, gamma, Weibull, Bernoulli, Poisson and negative-binomial distributions [6, pp. 267–272]. In simulation experiments the distribution is usually unknown to the analyst. We prefer to use a procedure based on normal distributions, since the normal distribution is a good approximation to many other distributions. Further-

more, procedures based on normality are often robust against deviations from normality. The reader is referred to Bechhofer, Kiefer, and Sobel [6] when one of the above distributions is a better approximation for a given simulation.

Procedures Assuming Unknown Variances

Bechhofer and Blumenthal [3], [4]

A *sequential* procedure for the case of normal populations with *common unknown variances* has been developed by Bechhofer [3]. Bechhofer [3, p. 416] has shown how to combine his procedure with various experimental designs which reduce the variance, and, therefore, decrease the sample size required. The procedure can be generalized for multi-factor experiments [3, p. 426]. In a later paper Bechhofer and Blumenthal [4] developed an equivalent but algebraically simpler procedure. Monte Carlo experiments [4] have demonstrated that even though the variance is unknown the Bechhofer-Blumenthal procedure is much more efficient than Bechhofer's single-sample procedure which assumes a known variance [2]. In the original paper [3] Bechhofer claims that the procedure yields the desired probability P^*, even if the means are in the least favorable configuration and that the procedure "can capitalize on any favorable configuration of the population means," [3, p. 414]. The latter point simply means that if the configuration of means is more favorable than the least favorable configuration, then the average sample size for the experiment to terminate will decrease [3, p. 415]. Unfortunately, Bechhofer has recently found an error in the *basic* procedure [3, p. 413] such that there is no guarantee that either the Bechhofer [3, p. 413] procedure or the Bechhofer-Blumenthal [4] procedure can capitalize on a favorable configuration of means.[1] However, Professor Bechhofer has indicated that the approximation procedure for calculating the "stopping statistic" $\tilde{Z}_m$ in §2.6 of [3, p. 414] does capitalize on a favorable configuration of population means, even though the *basic* procedure [3, p. 413] and the Bechhofer-Blumenthal [4] procedure do not.

Sasser and Naylor [42], [29] have applied this procedure with a multi-item inventory simulation experiment. Significant computer time reductions were achieved using this procedure when compared with a corresponding fixed sample size procedure [42].

Robbins, Sobel, and Starr [40]

Recently, Robbins, Sobel, and Starr derived a *sequential* procedure for normal populations with *a common unknown variance.* Their calculations are even simpler than those of Bechhofer and Blumenthal [4]. The Robbins, Sobel, and Starr method has the disadvantage of not being able to capitalize on a more favorable configuration of means; its expected sample size is independent of the configuration of the means. For the least favorable configuration, we expect that Bechhofer and Blumenthal's [4] procedure is more efficient than the procedure of Robbins, Sobel, and Starr. For Monte Carlo estimates [4, pp. 61–64] for certain combinations of δ^* and P^* indicate that the Bechhofer-Blumenthal procedure is much more efficient than the single-sample procedure for known σ^2. However, the expected sample size for the Robbins-Sobel-Starr procedure is approximately equal to the sample size of the single-sample

[1] See R. E. Bechhofer, "Correction Note—An Undesirable Feature of a Sequential Multiple-Decision Procedure for Selecting the Best One of Several Normal Populations with a Common Unknown Variance," *Biometrics*, Vol. XXVI (June 1970), pp. 347–349.

procedure [40, p. 91]. It is interesting to note that the stopping rule used by Robbins is the same as the single-sample procedure [2] for a common, known variance except that now an estimate of the variance is used. This estimate is revised each time that one observation is taken from all k populations.

Paulson [36]

Paulson has proposed a sequential procedure for "selecting the *normal* population with the greatest mean when (a) the k populations have a *common known variance* or (b) the k populations have a *common but unknown variance*, so that in each case the probability of making the correct selection exceeds a specified value when the greatest mean exceeds all other means by at least a specified amount," [36, p. 174]. Unlike the Bechhofer-Blumenthal [4] procedure the Paulson procedure can capitalize on a more favorable configuration. For, during the experiment populations are eliminated from further consideration (sampling) if their sample means differ from the sample mean of the population possessing the largest sample mean by some specified amount. Only limited information is available on which to base a comparison of the Paulson procedure with the Bechhofer-Blumenthal procedure. It is impossible to say which procedure is more efficient.

This procedure has been used by Sasser and Naylor [42], [29] to analyze data generated by a simulation experiment with an inventory model.

Gupta [16]

The multiple ranking problem treated by Gupta differs somewhat from the problem defined by Bechhofer [2]. Gupta's procedure, which may be useful in the screening phase of a simulation experiment, selects a nonempty, small, best *subset* from k populations such that the probability is at least equal to a specified value P^* that the "best" population is selected in the subset. This probability is guaranteed regardless of what the unknown means may be. There is no "indifference zone" and no "least favorable" configuration [16, p. 226]. With Gupta's procedure the sample size is given and the size of the subset is random. A *common unknown variance* is assumed.

Chambers and Jarratt [10]

Chambers and Jarratt show how double sampling may be used to select the best one of k nonnormal populations if the *variances* σ_i^2 of these populations *depend on the unknown means* μ_i, the form of this relationship being known and the same for each population, i.e.,

$$\sigma_i^2 = f(\mu_i), \qquad i = 1, \cdots, k, \tag{6}$$

where f is known and the same for all i. Such a relationship holds, for example, for binomial and Poisson populations. Chambers and Jarratt assume *asymptotically normally distributed sample means* [10, pp. 49–50]. Their procedure cannot capitalize on a more favorable configuration of the means. For the least favorable configuration Table 4 of [10, p. 59] shows that the average sample size is slightly larger than with the single-sample procedure for known variances. The authors conclude their paper with the following statement, "it is, perhaps, unlikely that (a multi-stage estimation) would be seriously in error," [10, p. 64].

Paulson [34]

A rather specialized approach is the *sequential* procedure developed by Paulson [33] which assumes a *common unknown variance* and eliminates inferior populations

during the experiment. The unique feature of this procedure is that it assumes a *standard* or *control* population. The need for such a procedure arises when one must choose between some existing policy (or decision rule) and k experimental policies. With the previous methods the correct selection of the control is assured only if the mean of the control exceeds the other population means by δ^*. Therefore, if the control mean exceeds the other population means only by an amount ϵ ($0 \leqq \epsilon \leqq \delta^*$), then an incorrect selection is indeed possible. However, if one is willing to accept a difference of ϵ, then he is not likely to be particularly interested in whether he makes an incorrect selection by incorrectly selecting population i ($i = 0, 1, \cdots, k$) instead of experimental population j ($j \neq i$, $j = 1, \cdots, k$) or by incorrectly selecting the experimental category j instead of the control population. The reason for developing special procedures for cases where not all populations are of an experimental nature is, therefore, somewhat less than obvious. Although this conclusion is reasonable in the case of business and economic problems, it may not hold in the case of other applications such as a medical problem involving a placebo and k experimental drugs.

Bechhofer, Dunnett, and Sobel [5]

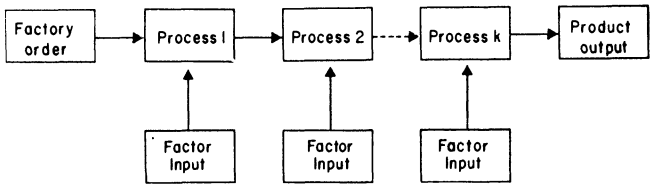
This paper describes a two-sample multiple decision procedure for ranking the means of *normal* populations with *unknown variances* but *known ratios*,

$$\sigma_i^2 = a_i\sigma^2, \qquad i = 1, \cdots, k, \tag{7}$$

where σ^2 is unknown but the a_i's are known. The procedure will be described in complete detail in the following section, when we use it to illustrate the applicability of multiple ranking procedures to a simulation experiment with a queueing model.

An Example Model

Consider a firm which produces a single product from a k-stage production process. All final products must pass through all k stages in sequential order. Each stage has its own production function which is independent of the others. The rate of output for each stage is a random variable whose probability distribution is completely determined by the factor inputs for a given stage. Thus the allocation of productive inputs is a policy variable that alters the probability distribution of the production rate for a given stage and ultimately affects total profit—but not with complete certainty. The rate of demand for the firm's product is also a random variable. Management is assumed to be interested in selecting the best (most profitable) production plan among five possible plans. The model underlying this production problem is described in complete detail in a paper by Chu and Naylor [11].



We ran simulation experiments using five different production plans for a simulated ninety-day period and then replicated each plan fifty times. We simply change the starting values in the random number generators to obtain the replications. The

following results were obtained:

Plan (j)	Sample Mean of Profit ($\bar{X}_j$)	Sample Standard Deviation of Profit (s)
I	2976.40	175.83
II	2992.30	202.20
III	2675.20	250.51
IV	3265.30	221.81
V	3131.90	277.04

We assume that for a given plan (population), j, X_{ij} is a normally and independently distributed random variable with $E(X_{ij}) = \pi_j$ and Var $(X_{ij}) = \sigma^2$ for $j = 1, \cdots, 5$. Thus the a_i's of (7) are assumed to be one, and we take our population as having a common unknown variance. The population profits, π_j, are unknown, but have some rank:

$$\pi_1 \leqq \pi_2 \leqq \cdots \leqq \pi_5 \tag{8}$$

and they have differences

$$\delta_{ij} = \pi_i - \pi_j, \qquad i, j = 1, \cdots, 5. \tag{9}$$

Suppose, as happens in many management situations, that we wish to select the plan with the largest expected population value. To apply the Bechhofer, Dunnett, and Sobel procedure, we first much choose P^* and δ^*. As in classical hypothesis testing where α and β must be specified, there is no definitive way to select these parameters. Much will depend on the situation confronting the manager, and good judgment is always needed. In many business and economic applications a choice of 0.95 or 0.90 is often appropriate, and we choose the latter for our illustration. For δ^* we select 100 based on an examination of the sample means where an inspection of five plans shows a range of roughly 500. We infer that 100 is the smallest difference that is worth detecting. In a real management simulation choice of δ^* would depend on the nature of the data and the cost of sampling weighed against the benefits of detecting a smaller difference.

Then to rank the best plan by the Bechhofer, Dunnett, and Sobel procedure, we apply the following five steps:

1. Take a first sample of N_1 observations from each of the m populations.

2. Calculate the mean square error, MS_e, which is an unbiased estimate of σ^2, having $v = m(n - 1)$ degrees of freedom for $n = N_1$.

3. Take a second sample of $N_2 - N_1$ observations from each of the m populations, $N_2 = \max \{N_1, [2MS_e(h/\delta^*)^2]\}$, where the brackets [ ] denote the smallest integer equal to or greater than the rational number contained within the brackets, and h is obtained from Table 3 of Dunnett and Sobel [14] for the given values of v and P^*. If $2\ MS_e(h/\delta^*)^2 \leqq N_1$, then no second sample is necessary and, therefore, $N_2 = N_1$.

4. For each population calculate the overall sample mean $\bar{X}_j$ where

$$\bar{X}_j = 1/N_2 \sum_{i=1}^{N_2} X_{ij}, \qquad j = 1, 2, \cdots, m. \tag{10}$$

5. Denote the ranked values of $\bar{X}_j$ by

$$\bar{X}_{[1]} < \bar{X}_{[2]} < \cdots < \bar{X}_{[m]}. \tag{11}$$

Rank the populations according to the ranking of the observed $\bar{X}_j$ and select the

population that gives rise to $\bar{X}_{[m]}$ as the population having the largest population mean.

In our simulation with 50 replications we have $n = N_1 = 50$. The mean square error is computed to be 51,901 and has $v = m(n - 1) = 5(49) = 245$ degrees of freedom. Consulting Table 3 of [14] for $P^* = 0.90$ and $v = 245$, we find $h = 1.58$. Next we determine

$$\begin{aligned} N_2 = \max\{N_1, [2\,MS_e(h/\delta^*)^2]\} &= \max\{50, [2(51{,}901)(1.58/100)^2]\} \\ &= \max\{50, 26\} \\ &= 50. \end{aligned} \tag{12}$$

Further, we note that $2\,MS_e(h/\delta^*)^2 = 26 < 50$, so that no second sample is required and $N_2 = N_1 = n = 50$. Sample means for $n = 50$ were shown in the previous table. On the basis of the ranking of the sample means, we would select operating plan IV as the plan with the highest expected profit. If, in fact, the best operating plan has an expected profit that is \$100.00 larger than the next best, then we have at least a probability of 90 per cent of correctly choosing it despite the random statistical fluctuations of sampling.

Now let us consider violations of some of the assumptions of existing multiple ranking procedures and what can be done in such a situation.

Some Problems with Simulation Experiments

All of the aforementioned multiple ranking procedures were based on the assumption of independence of the statistical error terms. This assumption will be satisfied with computer simulation experiments, if we treat one period of simulated time as one observation. Replicating this period with a new sequence of pseudorandom numbers yields another independent observation. In this way we can generate a sample of independent observations for a given period, provided we have chosen a suitable pseudorandom number generator [29]. Alternatively, if we choose not to replicate the simulation experiment, we can divide the sample record length into intervals that are longer than the interval of major autocorrelation and work with the observations on these supposedly independent intervals. This method suffers from the fact that the choice of sample record length usually does not have sufficient prior or posterior justification in most cases to make this approach much more than arbitrary.

As we have seen previously, the treatment of population variances differs considerably among multiple ranking procedures. However, no direct procedure is available for the case which is most likely to occur with computer simulation experiments—*unequal and unknown variances*. We shall consider several ways of circumventing this problem using Bechhofer's [2] problem formulation. In the absence of theoretical or empirical (Monte Carlo) evidence that these approaches will guarantee P^* given δ^*, they should be used with caution.

It should be noted that if the experimenter is willing (1) to assume a common unknown variance σ^2 and (2) to specify his indifference quantity δ^* as a percentage of either σ or σ^2, then he can use one of the multiple ranking procedures devised for the case of *common* but *known variances*, e.g., Bechhofer [2] or Bechhofer, Kiefer, and Sobel [6]. However, this approach is *impossible* if the variances are unequal or if the experimenter wishes to specify δ^* as, say, a fixed number of dollars.

Since most of the existing multiple ranking procedures assume normality, we shall also briefly mention the effects of violations of this assumption as well as alternative techniques for circumventing the problem.

We have previously discussed three sequential procedures for the case of *common unknown variances*. These include the procedures of Bechhofer and Blumenthal [3], [4]; Robbins, Sobel, and Starr [40]; and Paulson [36]. We can easily dispense with the Robbins [40] procedure on the grounds that (1) it cannot capitalize on a more favorable configuration of population means and (2) with the least favorable configuration, it would appear to be less efficient than the Bechhofer-Blumenthal [4] procedure. We shall, therefore, consider only the Bechhofer-Blumenthal [3], [4] procedure and the Paulson [36] procedure.

Bechhofer conjectures that his procedure is

> . . . relatively insensitive to lack of normality of the X_{ij} [the jth observation of the ith population]. For the procedure depends on the X_{ij} only through the
>
> $$Y_{im} = [\textstyle\sum_{j=1}^{m} X_{ij}]$$
>
> and S^2_{dm} [the unbiased estimate of σ^2 based on d_m degrees of freedom after m stages of experimentation]. And the distribution of the Y_{im} approaches normality as m grows large while S^2_{dm} is an unbiased estimate of the common σ^2 regardless of the distribution of the X_{ij} . However, the procedure can be expected to be sensitive to violations of the assumption of a constant common variance. And, in fact, it will break down completely if the violation is a severe one. With respect to the above violations, *the probability of a correct selection* for [the procedure] reacts in the same way as does the *power* of an Analysis of Variance test of homogeneity. And any set of conditions which will permit the experimenter to make a firm statement about the latter will also permit him to make a firm statement about the former [3, p. 426].

With simulation experiments the assumptions of normality and common variance are likely to be violated. Scheffé has considered the effects of violations in these assumptions on the power of an analysis of variance test of homogeneity. If the number of degrees of freedom is very large, Scheffé [41] has concluded that, "the effect of violation of the normality assumption is slight on inferences about means but dangerous on inferences about variances." Scheffé [41, p. 345] also argues that, "inequality of variances in the cells of a layout has little effect on inferences about means if the cell numbers are equal, serious effects with unequal cell numbers." In the case of a small number of degrees of freedom, Scheffé suggests that nonnormality still has relatively little effect on the power of the F-test, especially with equal cell numbers [41, p. 350]. With regard to the effect of the inequality of variance in the case of a small number of observations, Scheffé concludes that "equal group sizes should be used unless one is sure for certain groups that the variance is larger, in which case there is no harm in allocating more of the measurements to these groups" [41, p. 358].

Since Bechhofer [3] and Bechhofer and Blumenthal [4] have stated that the probability of a correct selection (with their sequential procedure) behaves in the same manner as the power of the F-test, we may use the Bechhofer-Blumenthal procedure for a common variance, if we have reason to believe the variances are (almost) equal. Such a belief may be based on a pilot phase where we obtain a number of observations from each population and estimate the variance of each population. (These pilot observations can be used again if we apply the Bechhofer-Blumenthal procedure.) If the pilot phase indicates inequality of variances, then the Bechhofer-Blumenthal procedure breaks down. A heuristic procedure for this case is described in [23].

Scheffé [41, pp. 364–365] has suggested the possibility of the use of transformations of the original output data to reduce the degree of inequality in variances—a problem that occurs often with simulation experiments. If the ratio σ_i/μ_i is a constant, then the logarithmic transformation is an appropriate transformation. Naylor, Wertz, and

Wonnacott [29], [31] have applied this transformation to national income data generated by a simulation experiment with an econometric model. The purpose of the simulation experiment was to evaluate the effects of five alternative monetary and fiscal policies on the stability of national income. The Bechhofer, Dunnett, and Sobel [5] procedure was used to determine the number of replications required for this experiment so as to rank the five economic policies on the basis of the variability of national income.

Paulson's [36] approach represents an alternative to the Bechhofer-Blumenthal [3], [4] procedure for simulation experiments where we can assume a common unknown variance. As we mentioned previously, there is a need for additional Monte Carlo experiments to evaluate the relative efficiency of the Paulson [36] procedure in contrast with the Bechhofer-Blumenthal [3], [4] procedure.

In the case of *unequal unknown variances*, life gets very difficult, indeed. Several heuristic procedures for this case are proposed in [23]. However, further theoretical and empirical (Monte Carlo) research is required in order to properly evaluate these heuristic procedures.

Conclusions

Having described several alternative multiple ranking procedures, it is obvious that at the present time an "ideal" procedure simply does not exist for computer simulation experiments. An ideal procedure for simulation experiments would take into consideration our lack of knowledge of the underlying theoretical distribution associated with a given complex computer model, i.e., the fact that the mean, variance, and form of the distribution are usually unknown to the simulator. Another desirable attribute for an idealized procedure would be that it be sequential and possess the ability to eliminate inferior populations. Finally, a more realistic formulation of the ranking problem might involve some type of Bayesian analysis in which *a priori* information and a loss function of the type proposed by Dunnett [13] are introduced. Although the multiple ranking problem appears to be amenable to Bayesian treatment, the development of such a procedure would be a nontrivial exercise.

In the absence of an ideal procedure, then heuristic techniques may offer a practical approach which is suitable for many simulation experiments. And certainly for simulations where the assumptions are met, multiple ranking procedures have a useful place in the tool kit of management scientists engaged in simulation.

References

1. Bahadur, R. R., "On the Problem in the Theory of k Populations," *Annals of Mathematical Statistics*, Vol. 21 (1950), pp. 362–375.
2. Bechhofer, R. E., "A Single Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances," *Annals of Mathematical Statistics*, Vol. 25 (1954), pp. 16–39.
3. ——, "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*, Vol. 14 (1958), pp. 408–429.
4. —— and Blumenthal, Saul, "A Sequential Multiple-Decision Procedure for Selecting the Best One of Several Normal Populations with a Common Unknown Variance. II, Monte Carlo Sampling Results and New Computing Formulae," *Biometrics*, Vol. 18 (1962), pp. 52–67.
5. ——, Dunnett, C. W. and Sobel, M., "A Two-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with a Common Unknown Variance," *Biometrika*, Vol. 41 (1954), pp. 170–176.

6. ——, KIEFER, JACK AND SOBEL, MILTON, *Sequential Identification and Ranking Procedures*, University of Chicago Press, Chicago, 1968.
7. —— AND SOBEL, M., "A Sequential Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances," *Annals of Mathematical Statistics*, Vol. 24 (1953), pp. 136–137.
8. —— AND SOBEL, MILTON, "A Single-Sample Multiple Decision Procedure for Ranking Variances of Normal Populations," *Annals of Mathematical Statistics*, Vol. 25 (1954), pp. 273–289.
9. BOFINGER, V. J., "The k-Sample Slippage Problem," *Australian Journal of Statistics*, Vol. 7 (1965), pp. 20–31.
10. CHAMBERS, M. L. AND JARRATT, P., "Use of Double Sampling for Selecting Best Population," *Biometrika*, Vol. 41 (1964), pp. 49–64.
11. CHU, KONG AND NAYLOR, THOMAS H., "A Dynamic Model of the Firm," *Management Science*, Vol. 11 (1965), pp. 736–750.
12. DUNNETT, C. W., "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association*, Vol. 50 (1955), pp. 1096–1121.
13. ——, "On Selecting the Largest of k Normal Population Means," *Journal of the Royal Statistical Society Series B* (1960), pp. 1–40.
14. —— AND SOBEL, M., "A Bivariate Generalization of Student's t-Distribution, with Tables for Certain Special Cases," *Biometrika*, Vol. 41 (1954), pp. 153–169.
15. —— AND ——, "Approximations to the Probability Integral and Certain Percentage Points of a Multivariate Analogue of Student's t-Distribution," *Biometrika*, Vol. 42 (1955), pp. 258–260.
16. GUPTA, S. S., "On Some Multiple Decision (Selection and Ranking) Rules," *Technometrics*, Vol. 7 (1965), pp. 225–246.
17. —— AND SOBEL, M., "On a Statistic Which Arises in Selection and Ranking Problems," *Annals of Mathematical Statistics*, Vol. 28 (1957), pp. 957–967.
18. —— AND ——, "On Selecting a Subset Which Contains All Populations Better Than a Standard," *Annals of Mathematical Statistics*, Vol. 29 (1958), pp. 235–244.
19. —— AND ——, "On Selecting a Subset Containing the Population with the Smallest Variance," *Biometrika*, Vol. 49 (1962), pp. 495–507.
20. —— AND ——, "On the Smallest of Several Correlated F Statistics," *Biometrika*, Vol. 49 (1962a), pp. 509–523.
21. HUYETT, M. J. AND SOBEL, M., "Selecting the Best One of Several Binomial Populations," *Bell System Technical Journal*, Vol. 36 (1957), pp. 537–576.
22. KARLIN, S. AND TRUAX, D., "Slippage Problems," *Annals of Mathematical Statistics*, Vol. 31 (1960), pp. 296–394.
23. KLEIJNEN, JACK P. C. AND NAYLOR, THOMAS H., "Multiple Ranking Procedures," Social System Simulation Program Working Paper Number 24, Duke University, September 1968.
24. KUDO, AKIO, "On the Invariant Multiple Decision Procedures," *Bulletin of the Mathematical Statistics Research Association of Statistical Science*, Vol. 6 (1956), pp. 57–68.
25. LEHMANN, E. L., "A Theory of Some Multiple Decision Problems," *Annals of Mathematical Statistics*, Vol. 28 (1957), pp. 1–25.
26. ——, "Some Model I Problems of Selection," *Annals of Mathematical Statistics*, Vol. 32 (1961), pp. 990–1012.
27. MOSTELLER, FREDERICK, "A k-Sample Slippage Test for an Extreme Population," *Annals of Mathematical Statistics*, Vol. 19 (1948), pp. 58–65.
28. —— AND TUKEY, J. W., "Significance Levels for a k-Sample Slippage Test," *Annals of Mathematical Statistics*, Vol. 21 (1950), pp. 120–123.
29. NAYLOR, THOMAS H., *Computer Simulation Experiments with Models of Economic Systems*, John Wiley and Sons, New York, 1971.
30. ——, WERTZ, KENNETH AND WONNACOTT, THOMAS, "Methods for Analyzing Data from Computer Simulation Experiments," *Communications of the ACM*, Vol. 10 (1967), pp. 703–710.
31. ——, —— AND ——, "Some Methods for Evaluating the Effects of Economic Policies Using Simulation Experiments," *Review of the International Statistical Institute*, Vol. 36 (1968), pp. 184–200.
32. PAULSON, EDWARD, "On the Comparison of Several Experimental Categories with a Control," *Annals of Mathematical Statistics*, Vol. 23 (1952), pp. 239–246.
33. ——, "An Optimum Solution to the k-Sample Slippage Problem for the Normal Distribution," *Annals of Mathematical Statistics*, Vol. 23 (1952), pp. 610–616.

34. ——, "A Sequential Procedure for Comparing Several Experimental Categories with a Standard or Control," *Annals of Mathematical Statistics*, Vol. 33 (1962), pp. 438–443.
35. ——, "A Sequential Procedure for Choosing One of k Hypotheses Concerning the Unknown Mean of a Normal Distribution," *Annals of Mathematical Statistics*, Vol. 34 (1963), pp. 549–554.
36. ——, "A Sequential Procedure for Selecting the Population with the Largest Mean for k Normal Populations," *Annals of Mathematical Statistics*, Vol. 35 (1964), pp. 174–180.
37. ——, "Sequential Estimation and Closed Sequential Decision Procedures," *Annals of Mathematical Statistics*, Vol. 35 (1964a), pp. 1048–1058.
38. PFANZAGL, J., "Ein kombiniertes test und klassifikations-problem," *Metrika*, Vol. 2 (1959), pp. 11–45.
39. RAMBERG, J. S., "A Comparison of the Performance Characteristics of Two Sequential Procedures for Ranking Means of Normal Populations," Technical Report Number 4, Department of Industrial Engineering, Cornell University, May 1966.
40. ROBBINS, H., SOBEL, M. AND STARR, N., "A Sequential Procedure for Selecting the Largest of k Means," *Annals of Mathematical Statistics*, Vol. 39 (1968), pp. 88–92.
41. SCHEFFÉ, HENRY, *The Analysis of Variance*, John Wiley and Sons, New York, 1959.
42. SASSER, W. EARL ET AL., "The Application of Sequential Sampling to Simulation: An Example Inventory Model," *Communications of the ACM*, Vol. 13 (May 1970), pp. 287–296.
43. SEAL, K. C., "On a Class of Decision Procedures for Ranking Means of Normal Populations," *Annals of Mathematical Statistics*, Vol. 26 (1955), pp. 387–398.
44. SHIRAFUJI, M., "Note on the Determination of the Replication Numbers for the Slippage Problem in r-Way Layout," *Bulletin of Mathematical Statistics*, Vol. 7 (1956), pp. 46–51.
45. SOMERVILLE, P. N., "Some Problems of Optimum Sampling," *Biometrika*, Vol. 41 (1954), pp. 420–429.

[Footnotes]

[1] **Correction Note: An Undesirable Feature of a Sequential Multiple-Decision Procedure for Selecting the Best One of Several Normal Populations with a Common Unknown Variance**

Robert E. Bechhofer

Biometrics, Vol. 26, No. 2. (Jun., 1970), pp. 347-349.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28197006%2926%3A2%3C347%3ACNAUFO%3E2.0.CO%3B2-C

References

[1] **On a Problem in the Theory of k Populations**

Raghu Raj Bahadur

The Annals of Mathematical Statistics, Vol. 21, No. 3. (Sep., 1950), pp. 362-375.

Stable URL:

http://links.jstor.org/sici?sici=0003-4851%28195009%2921%3A3%3C362%3AOAPITT%3E2.0.CO%3B2-C

[2] **A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with known Variances**

Robert E. Bechhofer

The Annals of Mathematical Statistics, Vol. 25, No. 1. (Mar., 1954), pp. 16-39.

Stable URL:

http://links.jstor.org/sici?sici=0003-4851%28195403%2925%3A1%3C16%3AASMDPF%3E2.0.CO%3B2-X

LINKED CITATIONS

[3] **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**

Robert E. Bechhofer

Biometrics, Vol. 14, No. 3. (Sep., 1958), pp. 408-429.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28195809%2914%3A3%3C408%3AASMPFS%3E2.0.CO%3B2-A

[4] **A Sequential Multiple-Decision Procedure for Selecting the Best One of Several Normal Populations with a Common Unknown Variance, II: Monte Carlo Sampling Results and New Computing Formulae**

Robert E. Bechhofer; Saul Blumenthal

Biometrics, Vol. 18, No. 1. (Mar., 1962), pp. 52-67.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28196203%2918%3A1%3C52%3AASMPFS%3E2.0.CO%3B2-H

[5] **A Two-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with a Common Unknown Variance**

Robert E. Bechhofer; Charles W. Dunnett; Milton Sobel

Biometrika, Vol. 41, No. 1/2. (Jun., 1954), pp. 170-176.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28195406%2941%3A1%2F2%3C170%3AATMDPF%3E2.0.CO%3B2-Q

[8] **A Single-Sample Multiple Decision Procedure for Ranking Variances of Normal Populations**

Robert E. Bechhofer; Milton Sobel

The Annals of Mathematical Statistics, Vol. 25, No. 2. (Jun., 1954), pp. 273-289.

Stable URL:

http://links.jstor.org/sici?sici=0003-4851%28195406%2925%3A2%3C273%3AASMDPF%3E2.0.CO%3B2-R

[10] **Use of Double Sampling for Selecting Best Population**

M. L. Chambers; P. Jarratt

Biometrika, Vol. 51, No. 1/2. (Jun., 1964), pp. 49-64.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28196406%2951%3A1%2F2%3C49%3AUODSFS%3E2.0.CO%3B2-M

LINKED CITATIONS

[11] **A Dynamic Model of the Firm**
Kong Chu; Thomas H. Naylor
Management Science, Vol. 11, No. 7, Series A, Sciences. (May, 1965), pp. 736-750.
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28196505%2911%3A7%3C736%3AADMOTF%3E2.0.CO%3B2-A

[12] **A Multiple Comparison Procedure for Comparing Several Treatments with a Control**
Charles W. Dunnett
Journal of the American Statistical Association, Vol. 50, No. 272. (Dec., 1955), pp. 1096-1121.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28195512%2950%3A272%3C1096%3AAMCPFC%3E2.0.CO%3B2-K

[14] **A Bivariate Generalization of Student's t-Distribution, with Tables for Certain Special Cases**
Charles W. Dunnett; Milton Sobel
Biometrika, Vol. 41, No. 1/2. (Jun., 1954), pp. 153-169.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195406%2941%3A1%2F2%3C153%3AABGOSW%3E2.0.CO%3B2-2

[15] **Approximations to the Probability Integral and Certain Percentage Points of a Multivariate Analogue of Student's t-Distribution**
C. W. Dunnett; M. Sobel
Biometrika, Vol. 42, No. 1/2. (Jun., 1955), pp. 258-260.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195506%2942%3A1%2F2%3C258%3AATTPIA%3E2.0.CO%3B2-Z

[16] **On Some Multiple Decision (Selection and Ranking) Rules**
Shanti S. Gupta
Technometrics, Vol. 7, No. 2. (May, 1965), pp. 225-245.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196505%297%3A2%3C225%3AOSMD%28A%3E2.0.CO%3B2-F

[17] **On a Statistic which Arises in Selection and Ranking Problems**
Shanti S. Gupta; Milton Sobel
The Annals of Mathematical Statistics, Vol. 28, No. 4. (Dec., 1957), pp. 957-967.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195712%2928%3A4%3C957%3AOASWAI%3E2.0.CO%3B2-2

LINKED CITATIONS

[18] **On Selecting a Subset Which Contains All Populations Better Than a Standard**
Shanti S. Gupta; Milton Sobel
The Annals of Mathematical Statistics, Vol. 29, No. 1. (Mar., 1958), pp. 235-244.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195803%2929%3A1%3C235%3AOSASWC%3E2.0.CO%3B2-N

[19] **On Selecting a Subset Containing the Population with the Smallest Variance**
Shanti S. Gupta; Milton Sobel
Biometrika, Vol. 49, No. 3/4. (Dec., 1962), pp. 495-507.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196212%2949%3A3%2F4%3C495%3AOSASCT%3E2.0.CO%3B2-H

[20] **On the Smallest of Several Correlated F Statistics**
Shanti S. Gupta; Milton Sobel
Biometrika, Vol. 49, No. 3/4. (Dec., 1962), pp. 509-523.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196212%2949%3A3%2F4%3C509%3AOTSOSC%3E2.0.CO%3B2-S

[22] **Slippage Problems**
Samuel Karlin; Donald Truax
The Annals of Mathematical Statistics, Vol. 31, No. 2. (Jun., 1960), pp. 296-324.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196006%2931%3A2%3C296%3ASP%3E2.0.CO%3B2-4

[25] **A Theory of some Multiple Decision Problems, I**
E. L. Lehmann
The Annals of Mathematical Statistics, Vol. 28, No. 1. (Mar., 1957), pp. 1-25.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195703%2928%3A1%3C1%3AATOSMD%3E2.0.CO%3B2-9

[26] **Some Model I Problems of Selection**
E. L. Lehmann
The Annals of Mathematical Statistics, Vol. 32, No. 4. (Dec., 1961), pp. 990-1012.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196112%2932%3A4%3C990%3ASMIPOS%3E2.0.CO%3B2-1

LINKED CITATIONS

[27] **A k-Sample Slippage Test for an Extreme Population**
Frederick Mosteller
The Annals of Mathematical Statistics, Vol. 19, No. 1. (Mar., 1948), pp. 58-65.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28194803%2919%3A1%3C58%3AASTFAE%3E2.0.CO%3B2-C

[28] **Significance Levels for a k-Sample Slippage Test**
Frederick Mosteller; John W. Tukey
The Annals of Mathematical Statistics, Vol. 21, No. 1. (Mar., 1950), pp. 120-123.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195003%2921%3A1%3C120%3ASLFAKS%3E2.0.CO%3B2-K

[32] **On the Comparison of Several Experimental Categories with a Control**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 23, No. 2. (Jun., 1952), pp. 239-246.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195206%2923%3A2%3C239%3AOTCOSE%3E2.0.CO%3B2-1

[33] **An Optimum Solution to the k-Sample Slippage Problem for the Normal Distribution**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 23, No. 4. (Dec., 1952), pp. 610-616.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195212%2923%3A4%3C610%3AAOSTTS%3E2.0.CO%3B2-D

[34] **A Sequential Procedure for Comparing Several Experimental Categories with a Standard or Control**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 33, No. 2. (Jun., 1962), pp. 438-443.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196206%2933%3A2%3C438%3AASPFCS%3E2.0.CO%3B2-1

[35] **A Sequential Decision Procedure for Choosing One of k Hypotheses Concerning the Unknown Mean of a Normal Distribution**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 34, No. 2. (Jun., 1963), pp. 549-554.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196306%2934%3A2%3C549%3AASDPFC%3E2.0.CO%3B2-J

LINKED CITATIONS

[36] **A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 35, No. 1. (Mar., 1964), pp. 174-180.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196403%2935%3A1%3C174%3AASPFST%3E2.0.CO%3B2-G

[37] **Sequential Estimation and Closed Sequential Decision Procedures**
Edward Paulson
The Annals of Mathematical Statistics, Vol. 35, No. 3. (Sep., 1964), pp. 1048-1058.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196409%2935%3A3%3C1048%3ASEACSD%3E2.0.CO%3B2-8

[40] **A Sequential Procedure for Selecting the Largest of k Means**
H. Robbins; M. Sobel; N. Starr
The Annals of Mathematical Statistics, Vol. 39, No. 1. (Feb., 1968), pp. 88-92.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196802%2939%3A1%3C88%3AASPFST%3E2.0.CO%3B2-B

[43] **On a Class of Decision Procedures for Ranking Means of Normal Populations**
K. C. Seal
The Annals of Mathematical Statistics, Vol. 26, No. 3. (Sep., 1955), pp. 387-398.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195509%2926%3A3%3C387%3AOACODP%3E2.0.CO%3B2-4

[45] **Some Problems of Optimum Sampling**
Paul N. Somerville
Biometrika, Vol. 41, No. 3/4. (Dec., 1954), pp. 420-429.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195412%2941%3A3%2F4%3C420%3ASPOOS%3E2.0.CO%3B2-T