DESIGNING SIMULATION EXPERIMENTS TO COMPLETELY RANK ALTERNATIVES

Thomas A. Bishop

ABSTRACT

Many of the problems of selecting the t-best of k populations with respect to a given parameter have been successfully solved for some time. Important applications, including applications to the design and analysis of simulation experiments, have been demonstrated and the tables required for implementation of solutions are easily available (e.g. in Gibbons, Olkin and Sobel (5)).

Solutions to the companion problem of completely ranking k populations have been scarce due to their mathematical complexity. This paper discusses examples of complete ranking problems, some recent advances made towards their solution, and cites sources of tables needed for their implementation.

I. INTRODUCTION TO THE COMPLETE RANKING PROBLEM

The statistical methodology known as ranking and selection procedures is now recognized as an important tool for solving many problems involving k populations (see Gibbons, Olkin and Sobel (5)). A description of a procedure for selecting from among k populations the one with the largest mean, and applications of that procedure to analysis of simulation experiments can be found in Dudewicz (3). The related problem of completely ordering k populations with respect to some parameter, for example the mean or the variance, has similar applications. Solutions to complete ranking problems have progressed much more slowly than solutions to selection problems. Basically this is due to the mathematical complexity of the complete ranking problem. This paper gives some examples of complete ranking problems, discusses some recent solutions to these problems, and reviews the relavent literature. Sources of tables needed to implement these procedures are given throughout the remaining sections.

II. THE COMPLETE RANKING PROBLEM

An example will best serve to introduce the reader to the complete ranking problem. Suppose we have 5 alternative air routes with similar layover patterns for flights from New York to Los Angeles. We want to order these routes with respect to the average time (measured in hours) it takes to complete the flight. If we use Route 1 for a given flight, we will observe a completion time X_{11}. For a second flight using the same route, we will observe a completion time X_{12} and so on. If we observe Route 1 for n flights, then we will have observed n completion times $X_{11}, X_{12}, \ldots, X_{1n}$. These times will not all be equal, and to account for this variability we assume that these are n observations (a random sample) from a probability distribution which has a mean completion time μ_1. The sample mean

$$\overline{X}_1 = \frac{X_{11} + X_{12} + \ldots + X_{1n}}{n}$$

will estimate μ_1. Similar statements can be made concerning the other four routes, so if we observe n flights for each route, we obtain a sample mean for each, namely, $\overline{X}_1, \overline{X}_2, \overline{X}_3, \overline{X}_4, \overline{X}_5$.

Next we order the five sample means: let $\overline{X}_{[1]} \leq \overline{X}_{[2]} \leq \overline{X}_{[3]} \leq \overline{X}_{[4]} \leq \overline{X}_{[5]}$ denote the ordered $\overline{X}_i$. We know which route generated $\overline{X}_{[1]}$, $\overline{X}_{[2]}$, and so on, and the population ordering rule is to order the routes in the same manner as their sample means. For example, if $\overline{X}_1 = 4.5$, $\overline{X}_2 = 4.3$, $\overline{X}_3 = 6.1$, $\overline{X}_4 = 5.2$ and $\overline{X}_5 = 4.0$, then the ordered values are 4.0, 4.3, 4.5, 5.2, 6.1, and we order the routes in the same way. The order of the routes in terms of preference (in this case best to worst) is

Route 5 , Route 2 , Route 1 , Route 4 , Route 3 .

The question then is: What is the probability that we have ordered these routes correctly in terms of their corresponding μ_i values? Before the experiment is performed an even more important question is: What value of n (how many flights for each route) should we use so that the probability of correct ranking (PCR) is at least a specified amount? It is clear that the probability of correct ranking, for any fixed value of n, decreases as the μ_i's become closer to each other. Therefore we will control the PCR when they are separated by some specified amount (determined by the experimenter). This is accomplished in the following way. Let $\mu_{[1]} \le \dots \le \mu_{[k]}$ denote the ordered μ_i. Then we will only control the PCR when

$$\mu_{[k]} - \mu_{[k-1]} \ge \delta^*$$

$$\mu_{[k-1]} - \mu_{[k-2]} \ge \delta^*$$

$$\vdots$$

$$\mu_{[2]} - \mu_{[1]} \ge \delta^* \quad .$$

The probability of correct ranking is a minimum subject to the above restrictions when each of the differences actually equals δ^*. In that case we say that the means are in a least-favorable configuration (LFC). The question then becomes: What is the smallest value of n such that the probability of correct ranking is at least P* when all the differences $\mu_{[i]} - \mu_{[i-1]} = \delta^*$ $i = 2, \dots, k$, where δ^* and P* are chosen by the experimenter? For the previous example we might want to choose n so that the probability of correct ranking is .90 when $\delta^* = .25$ hours.

The problem of course arises in trying to determine the appropriate value of n given P* and δ^*, or in finding P* given n and δ^*. Different methods of evaluating n are discussed in the next section.

III. METHODS OF SOLVING FOR n

III. A. Ranking Normal Means

If we assume that the observations from population i are normally distributed with mean μ_i, with known variance σ^2 and that all observations are independent, then the solution for n given P* and δ^* is

$$n = \left(\frac{\lambda\sigma}{\delta^*}\right)^2$$

where λ is obtained from table P.1 of Gibbons, Olkin and Sobel (5) for the given values of P* and k, the number of populations. The entries in this table (which were computed by numerical integration) were adapted from the work of Carroll and Gupta (2). For the previous example suppose $\sigma = .50$, $\delta^* = .25$, and $P^* = .90$. Then from table P.1 $\lambda \approx 2.80$ and

$$n = \left(\frac{(2.80)(.50)}{.25}\right)^2 = 31.36$$

so 32 flights should be observed for each route. If the true configuration of the μ_i is such that $\mu_{[i]} - \mu_{[i-1]} \ge .25$ for each $i = 2, 3, 4, 5$, then the probability that we correctly order the routes is at least .90.

In this case we have assumed a common known variance σ^2. The results of Dudewicz and Dalal (4), which use a two-stage sampling procedure, extend to the complete ranking problem and allow the assumption of a common known variance to be dropped.

This solution for n depends upon the assumption of normal populations. The following solution of Bishop and Dudewicz (1) is more general but n must usually be solved for by Monte Carlo techniques. We will demonstrate this solution by considering an example of the scale parameter problem. The general details of the solution can be found in Bishop and Dudewicz (1), and a similar solution can also be obtained for the general location problem.

III. B. Ranking Scale Parameters

We now assume that observations from the i^{th} population depend upon a parameter θ_i, that the statistic T_i estimates θ_i, and the probability distribution function for T_i is $F(x|\theta_i)$, where F is known. Then if $T_{(i)}$ is the estimate of $\theta_{[i]}$, we have

$$\text{PCR} = P(T_{(1)} < T_{(2)} < \dots < T_{(k)}) \; .$$

Now in this case we control the PCR whenever

$$\theta_{[i]}/\theta_{[i-1]} \ge c \quad \text{for all } i = 2, \dots, k$$

where $\theta_{[1]} \le \theta_{[2]} \le \dots \le \theta_{[k]}$ are the ordered θ_i. Here the LFC is the configuration of the θ_i such that $\theta_{[i]}/\theta_{[i-1]} = c$ for all $i = 2, \dots, k$. Now $Z_i = T_{(i)}/\theta_{[i]}$ $i = 1, \dots, k$ all have the same probability distribution function F(x), and

$$\text{PCR} \ge P[T_{(1)}/\theta_{[1]} \le c\,\frac{T_{(2)}}{\theta_{[2]}} \le \dots \le c^{k-1}\,\frac{T_{(k)}}{\theta_{[k]}}]$$

$$= P[Z_1 < cZ_2 < \dots < c^{k-1} Z_k].$$

Since the distribution of the Z_i is known in many cases, we can generate the Z_i's on the computer and use a Monte Carlo experiment to estimate the lower bound for the PCR. We discuss the Monte Carlo experiment in Section IV.

Returning to the example in Section II, our interest may be in ranking the routes in terms of their variability of times, rather than on the mean times. This is important because precise prediction of arrival time depends upon the variability. If we assume that the observations from the i^{th} route are normally distributed with mean μ_i and variance σ_i^2, then we use $T_i = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$. Note that the usual unbiased estimate of σ_i^2 is

$s_i^2 = T_i/(n-1)$. If we let $T_{(i)}/(n-1)$ be the estimate of $\sigma^2_{[i]}$, then

$$\begin{aligned} PCR &= P[T_{(1)}/(n-1) < T_{(2)}/(n-1) < \dots < T_{(k)}/(n-1)] \\ &= P[T_{(1)} < T_{(2)} < \dots < T_{(k)}] \\ &\geq P[Z_1 < cZ_2 < \dots < c^{k-1} Z_k] \end{aligned}$$

where $Z_i = T_{(i)}/\sigma^2_{[i]}$ has a chi-squared distribution with n-1 degrees of freedom. Hence F(x) is the distribution function for a chi-squared distribution with n-1 degrees of freedom.

Chi-squared random variables can be easily generated on a computer, and this lower bound can be evaluated by a Monte Carlo experiment. For various values of n, k, P* and c these lower bounds have been generated for the above case by Schafer and Rutemiller (6). These tables appear as table P.3 in Gibbons, Olkin and Sobel (5).

For the problem discussed above suppose we want the PCR to be at least .90 when $\sigma^2_{[i]}/\sigma^2_{[i-1]} \geq 1.6$ for all i = 2, 3, 4, 5. Then from table P.3 the required value of $\nu = n-1$ (the degrees of freedom for the chi-squared distribution) is 70 so the required value of n is 71.

Note that although we applied the results of III. B. to normal populations, we could also apply them to non-normal populations. Only the distribution function F(x) for the Z_i, which was the chi-squared distribution function in this case, would change.

IV. DESCRIPTION OF THE MONTE CARLO EXPERIMENTS

The lower bound for the PCR subject to the restriction that the $\theta_{[i]}/\theta_{[i-1]} \geq c$ i = 2, ..., k is

$$P[Z_1 < cZ_2 < \dots < c^{k-1} Z_k]$$

where the Z_i have the same known probability distribution F(x). To estimate this probability we generate $Z_1, Z_2, \dots, Z_k$ and check to see if $Z_1, cZ_2, \dots, c^{k-1} Z_k$ are in ascending order. This process is repeated N times, and we estimate the probability by

$$\widehat{PCR} = \text{'number of correct orderings'}/N ,$$

If the Monte Carlo experiment is based on N=20,000 replications, the maximum standard deviation of these probability estimates is $((.5)(.5)/20{,}000)^{1/2} \approx .0035$.

BIBLIOBRAPHY

1. Bishop, T. A. and Dudewicz, E. J., "Complete Ranking of Reliability-Related Distributions," IEEE Transactions on Reliability, Vol. R-26, No. 5 (1977), pp. 362.365.

2. Carroll, J. R. and Gupta, S. S., "On the probabilities of ranking of k populations with applications," Journal of Statistical Computation and Simulation, Vol. 5 (1977), pp. 145-157.

3. Dudewicz, E. J., "Statistics in Simulation: How to design for selecting the best alternative," Proceedings of the 1976 Bicentennial Winter Simulation Conference, edited by H. J. Highland, T. J. Schriber, and R. G. Sargent (1976), pp. 67-71.

4. Dudewicz, E. J. and Dalal, S. R., "Allocation of observations in ranking and selection with unequal variances," Sankhyā, B, 37, pp. 28-78.

5. Gibbons, J. D., Olkin, I. and Sobel, M., Selecting and Ordering Populations: A New Statistical Methodology, John Wiley & Sons (1977).

6. Schafer, R. E. and Rutemiller, H. C., "Some characteristics of a ranking procedure for population parameters based on chi-square statistics," Technometrics, Vol. 17, No. 3 (1975) pp. 327-331.