THE USE OF EXPERIMENTAL DESIGN TECHNIQUES IN SIMULATION

A. L. Frank
Advanced Information Development, Inc.
Los Angeles, California

Summary

This paper describes the use of statistical experimental design techniques in simulation modeling.

Background Statement

Many statisticians are actively participating in the research of simulation experimental designing. Many simulation analysts are surveying their work and attempting to tell their fellow analysts what these theoreticians have to offer. Unfortunately, few, if any, illustrative papers have been presented demonstrating the application of statistics to an actual simulation model. More often than not, the analyst is more concerned with "debugging" than analysis. When he does finally pass through the mystical state of model validation, he performs further black magic by experimenting in a manner that would disgust any student of statistics. Unfortunately, this author has been as guilty of the above crimes as other practicioners. The purpose of this paper is to discuss an honest attempt at utilizing statistical design of experiments in an actual simulation modeling effort.

Simulation Environment

Given the follwoing situation:
A machine shop with three machine groups, i.e., three groups of machines which are identical within groups. Each machine group may have one, two, or three machines.
A set of service times S_i $i = 1,2,3$
A set of job orders J_i $i = 1,2,3,4,...$ which have interarrival times I_i $i = 1,2,3,4,.....$and routing (a,b,c). [It is assumed for purpose of illustration that all jobs have the same routing.]

Obviously in a simulation environment one would generate a predetermined job set and continue to try different strategies against this set. However, approaching this from the classical statistical design of experiments standpoint, we wish to design a simulation experiment which thoroughly analyzes the relationships of the factors involved in the experiment. In the above example the follwoing analysis illustrates this approach.

Classical Design Analysis

Assuming a predetermined job mix, the factors under consideration are:

A = the number of machines in machine group 1.
B = the number of machines in machine group 2.
C = the number of machines in machine group 3.

A secondary problem becomes to determine what is a good measure of system response. In the above example, one measure might be the average throughput time for a given job mix. Of course, there might be other measures of effectiveness for a job shop environment. In the above example, assuming a predetermined job mix, there are still many other measures which are related. This paper will deal only with the average throughput time.

Using a standard 3 x 3 x 3 factorial experimental design, the following table was derived. In a complete factorial design, all possible combinations of allowable factor levels are tested.

	Level of factor C								
	1			2			3		
Level of factor A	Level of factor B			Level of factor B			Level of factor B		
	1	2	3	1	2	3	1	2	3
1	*	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*	*	*

* denotes an "observation"

The above table indicates that 27 observations were necessary. Since in most Monte Carlo simulation runs, several replications at each set of factors are run, the 27 runs could be more realistically some integer multipile of 27 (say for example, 81.) Of course, as the levels of factors and the number of factors increase, the number of observations could also rapidly increase. It becomes apparent that this approach can necessitate extremely large computer time requirements. Therefore, the cost of the experiment may outweigh the benefits derived. Before demonstrating a possible solution to the problem of a large number of observations, it would be useful to discuss what type of results can be obtained from the above experiment.

Since in the above case we are treating quantitative factors, we could proceed to calculate the main effects and the interactions. (The actual study does, in fact, do so.) These numbers can be analyzed as to levels of significance and certain inferences may be drawn.

Improved Statistical Analysis

However, as was pointed our previously, even a simple illustrative example could tax the budget of most analysts. Therefore, it would be beneficial for the analyst to use some type of experimental technique which would allow him to reduce the number of samples. It turns out that one such technique is available. This technique not only allows him to reduce the number of required observations, but to also attempt to arrive at an optimal combination of factors. The basic technique stems from work done by G. E. P. Box and K. B. Wilson. [1.] Their method and others' extensions are excellently sumarized in several design of experiments texts. [2.] The Box method is, in essence, the basis for the SimOptimization package which is commercially available. [3.]

The technique referred to is not extremely complicated from a mathematical standpoint. However, it is complicated from a procedural point of view. Therefore, this paper shall attempt only to approximate the technique. Hopefully, the interested reader will refer to the references for the details.

Essentially what is done is to view the simulation experiment as a method for providing responses on a response surface. By carefully positioning himself on this surface, the analyst should be able to approach optimal conditions rapidly. Usually what is done is to use the method of steepest ascent to approach a stationary point. This is done by using the standard gradient techniques of numerical analysis coupled with factorial designs to estimate the function describing the response surface. Once one reaches a stationary region, he may attempt to fit a polynomial expression to the local response surface. Once this is done, the desired position may be found by standard partial differentiation.

The basic problem with this method is that the analyst must be able to stop his model at periodic intervals. Then after analyzing the results of a small series of experiments, he proceeds to conduct another series of tests. The actual desire to have a continous run arriving at an optimum solution is a programming exercise since the algorithm can be developed in a general form. The programming exercise is unfortunately not a trivial one since even with the HELP block, GPSS/360 is still not at a state where easy communication is available with other languages.

Obviously the use of statistical design of experiments brings simulation out of the dark ages with respect to statistical respectability. However, it is questionable that the simulation environment will ever reach the sanctity of the chemical laboratory for where these techniques were devised.

Bibliography

[1.] See Davies, Owen L., The Design and Analysis of Industrial Experiments, Hafner Publishing Company, New York, 1963.

[2.] One such is Cochran, William G. and Cox, Gertrude M., Experimental Designs, John Wiley & Sons, Inc., New York, 1957.

[3.] See Karr, H., Luther, E., Markowitz, H., and Russell, E., "SimOptimization Research, Phase I and Phase II", Consolidated Analysis Centers, Inc., Santa Monica, 1965.