Proceedings of the 1992 Winter Simulation Conference
ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson

AUTOSTAT

Van B. Norman

AutoSimulations
655 Medical Drive
Bountiful, Utah 84011, U.S.A.

1. ABSTRACT

AutoStat provides a complete statistical analysis capability in conjunction with AutoMod. Statistical analysis is an important part of using simulation to make decisions. When experimenting with a single model, AutoStat determines the "warm-up" period and computes the minimum, maximum and confidence intervals for model statistics. When comparing alternative system designs, AutoStat manages the data bases and statistical analysis for all the options. AutoStat's "Design of Experiments" feature provides a unique capability using statistical optimization theory. It reduces the number of experiments required to determine which values, from a set of options, have the most effect on system performance.

2. INTRODUCTION

The purpose of simulation is to obtain a useful answer. Thus, you must understand how much variability is in a simulation's output. As the randomness of the output increases, a greater number of runs are needed to obtain accurate results. For this reason statistics is a necessary tool. AutoStat allows you to easily create complex experiments that produce multiple representative samples of system behavior. For each response variable of interest, AutoStat computes an estimate of standard error or stochastic variability, providing a measure of the statistical accuracy of the simulation-generated response. AutoStat provides automated procedures for measuring and interpreting these outputs. Numerical and graphical output statistics convey information about system outputs, and they provide a way of comparing alternative modeling solutions.

Too often, stochastic models are run once and the output is given as 'the' answer. That answer may represent a 'typical' day or an 'unusual' day. The only way to determine if the day was typical or unusual is to run a sufficient number of replications. Many runs are usually necessary to get an accurate picture of system behavior.

3. OUTPUT ANALYSIS

If a simulation is run long enough, random values from the input distributions eventually generate virtually every possible combination of events. If there is some probability, no matter how small, of every conveyor going down or all press-brake operators calling in sick, it eventually happens in the model.

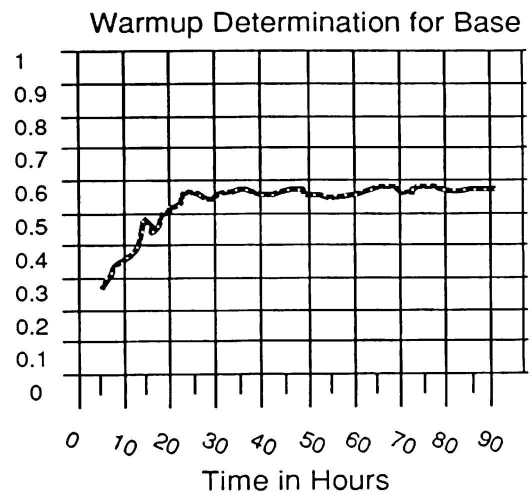
Each output from a run contains information on one instance of what can happen. The output doesn't contain information on the likelihood of it happening. The goal of a simulation is to estimate 'mean' or 'long-run average' system responses. This can be accomplished by using a point estimate, which is the 'best guess' of the actual long-run average system response. Also, an interval estimate can be used to represent the uncertainty in the point estimate. Based on the variance of the observations, a range of values is given, which contains the actual 'long-run average' system response with some level of confidence.

A point estimate is not the complete answer. An estimate is only as good as the amount of data used to generate it. A point estimate can be generated from one run as the result from that run. Additional runs are necessary to compute a standard error or confidence interval. Confidence intervals are used with stochastic (random) models.

4. WARM-UP FOR A NON-TERMINATING SIMULATION

Initial conditions affect the amount of time it takes a response to reach a steady-state. Also, the random numbers chosen to represent variability in system conditions also can affect the approach to steady-state. For non-terminating simulations, the time until a steady-state is reached is called the warm-up period. For reasons of space, only non-terminating systems will be addressed in this presentation.

There are two common strategies used to reduce initialization bias. The first is to delete the warm-up period once it is identified and use only statistics collected after that point. The second is to try to initialize the model as closely to the steady-state as possible, based on past experience or previous runs. Usually, it is better to use both strategies in combination.



It is often difficult to estimate realistic initial conditions. When examining new configurations, steady-state responses are usually unknown, and trying to find the steady-state is often the reason for the simulation! One way to determine the warm-up period is by graphing the responses and visually determining when the system is operating at steady-state. You can specify the period up to the steady-state as the warm-up period for future runs. You can then eliminate the data from the warm-up period in the analysis.

Because initialization bias can be a problem, several methods have been developed to deal with this bias (see Law and Kelton, 1991). Two statistical techniques for analyzing output data are incorporated into AutoStat: Replication/Deletion and Batch Means (see Law and Kelton, 1991). These techniques are discussed in later sections. Both techniques assume, however, that the warm-up period has been deleted.

5. ESTIMATION OF RESPONSES FOR A SINGLE SYSTEM

In order to estimate mean responses for a single system, two methods are commonly used: replication/deletion and batch means. Both methods generate confidence intervals on mean response levels. The data used to construct confidence intervals come from response levels taken over consecutive equal time intervals called snaps.

Batch means use one long replication for obtaining independent estimates of steady-state parameters. This replication is divided into several batches of equal length. If sufficiently large batch sizes are chosen, the batch means are approximately independent and normally distributed random variables. Thus, they can be used to obtain point estimates and construct confidence intervals for steady-state responses. As batches come from the same replication, consecutive batches can be correlated. Because of this, batch means have the problem of a biased estimator of the variance of the point estimate. Fortunately, if the batch size is large enough, this can be ignored and batches can be assumed to be uncorrelated. Unlike replication/deletion, the warm-up period is passed through only once.

The replication method takes the mean response from a series of independent runs as data to construct a confidence interval. In replication/deletion, statistics are collected following the warm-up period. Statistics are then collected for the remainder of the run and the results from several replications are used to construct a confidence interval. Using separate different random number seeds in each replication assures no correlation between replications, thus guaranteeing independent samples. One of the conveniences of AutoStat is that it automatically changes the random number seeds from run to run.

6. COMPARING CONFIGURATIONS

One of the greatest benefits of simulation is the ability to compare different system designs. These different designs are called configurations in AutoStat. It's more convenient and inexpensive to answer "What if..." questions using computer simulation than it is to experiment on the shop floor. Comparing configurations in AutoStat provides a framework to examine different system designs and then display and interpret the results.

7. SELECTING THE BEST SYSTEM

Selecting the best system out of three or more options may be desirable without enumerating all pairwise combinations. To do this, AutoStat selects the best system for a specified response at some level of confidence.

In order to choose the best configuration, AutoStat requires you to specify a confidence level, an indifference zone, and an initial number of runs. The confidence level represents how certain you want to be in your answer, similar to that for a confidence interval. The indifference zone is the range of responses that make no difference to you, based on practical or engineering considerations.

AutoStat utilizes a 2-stage procedure to select the best configuration. In the first stage, an initial n_0 replications of each configuration are run. AutoStat uses the results from these runs to compute the additional number of runs necessary for each configuration in order to select the best one at the desired confidence level. In the second stage, these additional runs are made and AutoStat can then select the best configuration.

8. DESIGN OF EXPERIMENTS

Design of Experiments is a statistical method whch efficiently determines the effects that model inputs have on model results. Rather than randomly picking which runs will be made, as one might otherwis be forced to do, Design of Experiments efficiently determines a set of runs which simultaneously estimate the effects that several model inputs have on the model. Design of Experiments also estimates interactions between inputs, i.e., two imputs can have no effect individually yet change in both can siumlatneously occur.

AutoStat supports two main designs: full-factorial designs, which are the most robust, and fractional factorial designs. Full factorial designs make no assumptions about factor interactions, yet can require many runs. Fractional factorial assume that high order interactions are negligible, but require only 1/2 to 1/8 the number of runs.

REFERENCES

Law, A.M. and W.D. Kelton 1991. *Simulation Modeling and Analysis, Second Edition,* McGraw-Hill, New York

Banks, Jerry and John S. Carson II 1984. *Discrete-Event System Simulation,* Prentice-Hall, Englewood Cliffs, NJ

AutoSimulations, Inc. 1992. *AutoStat User's Manual*

AUTHOR BIOGRAPHY

VAN B. NORMAN, President of AutoSimulations, Inc. received a B.S. in mathmatics at the University of Utah in 1969. He spent 6 years at Eaton-Kenway, where he implemented the first simulation animator. In 1982, using his experience in factory automation and simulation, he co-founded AutoSimulations, Inc., where he co-authored AutoMod, ASI's first graphic simulation software. He has authored papers on the application and the future of simulation in manufacturing. His interests are in world-class manufacturing operations and simulation research.