VARIANCE REDUCTION TECHNIQUES

Stephen S. Lavenberg
Peter D. Welch

ABSTRACT

This talk is concerned with statistical techniques, known as variance reduction techniques, for increasing simulation efficiency. The emphasis will be on critically assessing the practical applicability of variance reduction techniques in discrete event simulations of complex stochastic models of systems. Examples of the successful application of such techniques will be presented.

I. INTRODUCTION

Discrete event simulations of complex stochastic models of systems are often substantial consumers of computing time and hence costly. It is therefore important to carry them out as efficiently as possible. We will be concerned with statistical techniques, called variance reduction techniques, for increasing simulation efficiency. The purpose of a variance reduction technique is to produce at no extra cost a more accurate estimate than that obtained without application of the technique or to produce at less cost an estimate having the same accuracy. The means of the estimates obtained with and without application of the techniques should be equal and the accuracy of an estimate is measured by its variance.

Variance reduction techniques were originally applied in the Monte Carlo solution of deterministic problems, e.g., the numerical evaluation of multiple integrals, as well as in sampling problems in statistics. More recently they have been applied in the simulation of simple stochastic models, particularly simple queueing models. The 1974 book by Kleijnen (8) contains a chapter and an extensive bibliography on variance reduction techniques. Several techniques are described and critically appraised with respect to their applicability to system simulation. Statistical issues, such as determining confidence intervals, that arise when applying a variance reduction technique are also discussed. The goal of this talk is not to repeat this work, but rather to bring it up to date from our perspective. In particular we want to critically assess the practical applicability of variance reduction techniques in simulating complex stochastic models and to present examples of their successful application. In the remainder of this paper we briefly describe the types of simulation models we will consider, and the variance reduction techniques we will consider. We also include a bibliography of selected recent papers on variance reduction in discrete event simulations as a supplement to the bibliography in (8).

II. TYPES OF SIMULATION MODELS

The types of models we have in mind are complex discrete event stochastic simulation models of dynamic systems. Examples are a model of military engagement (12), a model of fire department operations (3), and queueing models of job shop type systems such as computer systems (10). Such models are often characterized by having one or more input streams of random variables which are transformed by the complex mechanism of the discrete event simulation into one or more output streams of random variables. The simulator may be interested in estimating steady state characteristics of an output stream. For example, in the simulation model of fire department operations referred to above one input stream consisted of the times between occurrences of serious fires and one output stream consisted of the times it took certain equipment to arrive at serious fires. The simulator was interested in estimating the steady state mean value of this output stream. In other simulation models, such as the military engagement model referred to above, there is no steady state, rather the simulation terminates when certain events occur. In the military engagement model the simulator wished to estimate several quantities including the mean time until termination and the probability that a particular event causes termination. In addition to estimating quantities for a system operating with particular policies and parameter values a simulator may wish to compare system variants. This was the case with both of the above models.

III. TECHNIQUES CONSIDERED

We will use the following notation. Let I denote the number of input streams of random variables and let

$$X(i)=(X_1(i),X_2(i),\ldots)$$

denote the i-th stream. The collection of these streams is denoted by

$$X=(X(1),\ldots,X(I)).$$

Let $\hat{q}$ be an estimator obtained from the simulation of an unknown deterministic quantity q. In general $\hat{q}$ is a complicated function of all the input streams, i.e.,

$$\hat{q}=\hat{q}(X).$$

We assume that $\hat{q}$ is an unbiased estimator of q, i.e.,

$$E[\hat{q}]=q. \qquad (1)$$

(If q is a steady state quantity then (1) usually holds approximately.) The techniques we will consider include the following.

ANTITHETIC VARIABLES

A pair of simulation runs is made where the collections of input streams used on the two runs, denoted by X and Y, have identical distributions, but the corresponding elements in one or more of the streams are negatively correlated, i.e., for some i, $X_k(i)$ and $Y_k(i)$ are negatively correlated for all k. The two runs produce estimators $\hat{q}(X)$ and $\hat{q}(Y)$ which have identical distributions, in particular

$$E[\hat{q}(X)]=E[\hat{q}(Y)],$$

$$Var[\hat{q}(X)]=Var[\hat{q}(Y)],$$

but are correlated. The goal is that $\hat{q}(X)$ and $\hat{q}(Y)$ be negatively correlated so that

$$Var[(\hat{q}(X)+\hat{q}(Y))/2] =$$

$$(Var[\hat{q}(X)]+Cov[\hat{q}(X),\hat{q}(Y)])/2$$

$$< Var[\hat{q}(X)]/2.$$

Note that since two simulation runs are made instead of one we require that the variance be reduced by at least a factor of two.

CONTROL VARIABLES

A control variable is a random variable C=C(X) produced by the simulation whose mean value E[C] is known and which is correlated with $\hat{q}$. For any constant a let

$$\hat{q}(a)=\hat{q}-a(C-E[C]). \qquad (2)$$

Then

$$E[\hat{q}(a)]=E[\hat{q}]$$

and

$$Var[\hat{q}(a)]=Var[\hat{q}]-2aCov[\hat{q},C]+a^2Var[C].$$

If

$$a<2Cov[\hat{q},C]/Var[C]$$

then

$$Var[\hat{q}(a)]<Var[\hat{q}].$$

Furthermore,

$$a^*=Cov[\hat{q},C]/Var[C]$$

minimizes $Var[\hat{q}(a)]$ and

$$Var[\hat{q}(a^*)]=Var[\hat{q}][1-\rho^2(\hat{q},C)]$$

where $\rho(\hat{q},C)$ is the correlation coefficient between $\hat{q}$ and C. In general $Cov[\hat{q},C]$ and Var[C] are not known so that a* is not known and is typically estimated as part of the simulation. The above development extends to the case of more than one control variable by using a linear combination of the control variables in (2). In this case the optimum coefficients (there is one for each control variable in the linear combination) are not known and are typically estimated.

One way of obtaining a control variable is to combine two estimators of q. Suppose

In addition to $\hat{q}$ a second unbiased estimator $\hat{q}'$ can be obtained from the simulation. For any constant a let

$$\hat{q}(a)=(1-a)\hat{q}+a\hat{q}'$$
$$= \hat{q}-a(\hat{q}-\hat{q}'). \quad (3)$$

Since

$$E[\hat{q}-\hat{q}']=0,$$

(3) is equivalent to (2) with C replaced by $\hat{q}-\hat{q}'$. A linear combination of more than two unbiased estimators, where the coefficients in the linear combination sum to one, can be used to obtain multiple control variables.

COMMON RANDOM NUMBERS

This technique is designed to be used when estimating the difference between quantities q_1 and q_2 for two system variants. Identical collections of input streams are used when simulating the two variants yielding estimators $\hat{q}_1(X)$ and $\hat{q}_2(X)$ of q_1 and q_2 respectively which are correlated. The goal is that $\hat{q}_1$ and $\hat{q}_2$ be positively correlated so that

$$Var[\hat{q}_1-\hat{q}_2]=Var[\hat{q}_1]-2Cov[\hat{q}_1,\hat{q}_2]+Var[\hat{q}_2]$$
$$<Var[\hat{q}_1]+Var[\hat{q}_2].$$

The sum on the right of the inequality is the variance that would be obtained if $\hat{q}_1$ and $\hat{q}_2$ were obtained using independent collections of input streams.

CONDITIONAL SAMPLING

Assume there is a random variable R=R(X) produced by the simulation such that the conditional mean $E[\hat{q}|R=r]$ is known for all possible values r of R. Let

$$q(r) = E[\hat{q}|R=r]$$

denote this known function of r. The random variable q(R) is an unbiased estimator of q since

$$E[q(R)] = \int E[\hat{q}|R=r]dF(r)$$
$$= E[\hat{q}],$$

where F is the distribution function of R. For any two random variables A and B it is known that

$$Var[A] = Var[E[A|B]]+E[Var[A|B]]$$
$$\geq Var[E[A|B]]. \quad (4)$$

Letting $A=\hat{q}$ and B=R in (4) we have that

$$Var[\hat{q}] \geq Var[q(R)].$$

Furthermore, equality holds if and only if $\hat{q}$ and q(R) are identical.

Antithetic variables, control variables and common random numbers were among the techniques considered by Kleijnen (8). Since that time control variables have received considerable attention (1), (6), (9), (10), (11), although common random numbers remains the most widely applied technique, e.g., (3), (12). Conditional sampling, also called conditional Monte Carlo, was only briefly mentioned in (8). However, a closely related technique was successfully applied to the simulation model of fire department operations (4).

BIBLIOGRAPHY

1. Carson, J. S., "Variance Reduction Techniques for Simulated Queuing Processes," Ph.D. Thesis, Department of Industrial Engineering, 78-8, University of Wisconsin-Madison, 1978.

2. Carson, J. S. and A. M. Law, "Conservation Equations and Variance Reduction in Queuing Simulations," 77-25, Department of Industrial Engineering, University of Wisconsin-Madison, 1977 (to appear in Oper. Res., 1979).

3. Carter, G. and E. J. Ignall, "A Simulation Model of Fire Department Operations: Design and Preliminary Results," IEEE Trans. Sys. Sci. and Cyb 6, 282-292, 1970.

4. Carter, G. and E. J. Ignall, "Virtual Measures: A Variance Reduction Technique for Simulation," Manage. Sci. 21, 607-616, 1975.

5. Fishman, G. S., "Variance Reduction Techniques for Nonstationary Simulation Models," Tech. Rep. No. 77-3, Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill, 1977.

6. Heidelberger, P., "Variance Reduction Techniques for the Simulation of Markov Processes," Ph.D. Thesis, Department of Operations Research, Stanford University, 1978.

7. Heidelberger, P. and D. L. Iglehart, "Comparing Stochastic Systems Using Regenerative Simulation with Common Random Numbers," Tech. Rep. No. 47, Department of Operations Research, Stanford University, 1978.

8. Kleijnen, J.P.C., Statistical Techniques in Simulation Part I, Marcel Dekker, Inc., New York, 1974.

9. Lavenberg, S. S., T. L. Moeller and C. H. Sauer, "Concomitant Control Variables Applied to the Regenerative Simulation of Queueing Systems," IBM Research Report RC 6413, Yorktown Heights, New York, 1977 (to appear in Oper. Res., 1979).

10. Lavenberg, S. S., T. L. Moeller, and P. D. Welch, "Control Variables Applied to the Simulation of Queueing Models of Computer Systems," in Computer Performance, Edited by K. M. Chandy and M. Reiser, North Holland Publishing Co., Amsterdam, 1977.

11. Lavenberg, S. S., T. L. Moeller and P. D. Welch, "Statistical Results on Multiple Control Variables with Application to Variance Reduction in Queueing Network Simulation," IBM Research Report, Yorktown Heights, New York, 1978.

12. McGrath, E. J. and D. C. Irving, "Application of Variance Reduction to Large Scale Simulation Problems," Comput. and Ops. Res. 1, 283-311, 1974.