Proceedings of the 1986 Winter Simulation Conference
J. Wilson, J. Henriksen, S. Roberts (eds.)

PROBLEMS IN BAYESIAN ANALYSIS OF STOCHASTIC SIMULATION

Peter W. Glynn
Dept. of Industrial Engineering
University of Wisconsin-Madison
1513 University Avenue
Madison, Wisconsin 53706

ABSTRACT:

It is argued that Bayesian methodology is an appropriate tool in certain simulation contexts. Computational problems, specific to simulation applications, are then described in some detail; possible remedies are also outlined.

1. INTRODUCTION

In this paper, we hope to show that Bayesian statistical methodology has an important role to play in the theory and practice of stochastic simulation. In particular, we will argue that Bayesian methods are often appropriate when the practitioner is attempting to quantify uncertainty in the input distributions that drive the stochastic system under study.

From a statistical viewpoint, the material to be dicussed here is standard and well known to the Bayesian community. A primary goal is to compute an appropriate posterior distribution (or, at least, its mean) for some quantity of interest. The difficulty is that because of the model complexity that is typical of stochastic simulation, the required calculations tend to be highly non-trivial, in terms of computational complexity. We intend to discuss computational remedies elsewhere; our focus here is on describing the relevant problems.

2. SIMULATION: TWO DIFFERENT APPLICATIONS

Before proceeding, it is useful to briefly discuss the nature of simulation. Our view is that:

> Monte Carlo simulation is a numerical tool for studying complex stochastic systems.

More precisely, let us suppose that the stochastic system is represented by a probability triple $(\Omega, \mathcal{F}, P)$. Let $X:\Omega \to R^d$ be a ($\mathcal{F}$-measurable) random vector such that EX exists. For a given map $g: R^d \to R$, the goal of Monte Carlo simulation is to estimate

(2.1) $\alpha = g(EX)$.

Monte Carlo methods are algorithms that numerically estimate α by drawing observations from one, or possibly more, independent copies of $(\Omega, \mathcal{F}, P)$; the estimate is then formed by appropriately combining observations.

(2.2) EXAMPLE.

Suppose that one is studying the M/M/1/∞ queue with the arrival rate $\lambda = 5$ and the service rate $\mu = 10$. (See p. 202 of ROSS (1980) for a description of the model.) Here, Ω is the space of right continous real-valued functions with left limits $D[0,\infty)$ (see ETHIER and KURTZ (1986)), $\mathcal{F}$ is the associated Borel σ-field, and P is the probability on $(\Omega,\mathcal{F})$ under which the co-ordinate process Y is the M/M/1/∞ queue described above. If one is interested in the number of customers in the system at time $t = 15$, then $X = Y(15)$ and $g(x) = x$. The standard Monte Carlo algorithm involves generating independent copies of $(\Omega,\mathcal{F}, P)$ and averaging the corresponding $Y_i(15)$'s to obtain the estimator.

If, on the other hand, one is interested in the steady-state variance, then one need only generate one observation of $(\Omega,\mathcal{F}, P)$, set $g(x_1,x_2) = x_1 - x_2^2$, and let

$$X = \left(\lim_{t\to\infty} \frac{1}{t} \int_0^t Y^2(s)ds, \lim_{t\to\infty} \frac{1}{t} \int_0^t Y(s)ds \right)$$

The main point to be emphasized here is that Monte Carlo simulation is a computational method for solving a certain class of problems. It follows that simulation is, in itself, not a statistical discipline; rather, it is most properly viewed as a sub-area of numerical analysis. From this point of view, it is misleading to regard simulation as a special type of "controlled statistical experiment".

Given that we very strongly support viewing Monte Carlo simulation as a part of numerical analysis, it is perhaps surprising to see us arguing for the application of Bayesian methods (which are inherently statistical in nature). There is a consistency here, however. In particular, Bayesian methods are reasonable within the applications environment associated with Monte Carlo simulation. In short:

> The Bayesian statistical framework does not apply to Monte Carlo simulation itself. Rather, it applies to a certain applications environment associated with Monte Carlo simulation.

To describe the connection in more detail, it is necessary to discuss the two main applications settings for Monte Carlo simulation.

APPLICATION 1: SOLVING A MATHEMATICAL PROBLEM

In Example 2.2, Monte Carlo simulation was suggested as a technique for solving certain mathematical problems associated with a specific mathematical model (as specified by $(\Omega,\mathcal{F}, P)$). The important point here is that the applications setting completely determines the probability triple

$(\Omega, \mathcal{F}, P)$. This approach to specification is typical of Monte Carlo simulation applications in:

i) mathematics: Given a specific mathematical model (such as the Schlugle model (GRASSBERGER et al (1979)) with fixed parameters), simulate in order to solve for analytically intractable quantities.

ii) physics: Given a specific physical model (as described, for example, by Schrodinger's equation, with specified initial conditions, use Monte Carlo simulation to solve.

iii) statistics: Use Monte Carlo simulation to calculate analytically intractable sampling distributions (e.g. Monte Carlo studies of power of statistical tests.)

APPLICATION 2: A DECISION SUPPORT TOOL

In most Industrial Engineering/Operations Research applications of Monte Carlo simulation, the simulation is used to assist in decision-making of some kind. For example, simulation is frequently used to test the design of a proposed system, in order to decide whether or not to proceed with development.

In such a context, the applications setting does not, in general, uniquely determine the probability triple $(\Omega, \mathcal{F}, P)$. The practitioner, using his or her applications background, formulates a model that is driven by certain input distributions and parameters. As a consequence, the model formulation takes the form $(\Omega, \mathcal{F}, P(F))$, where $F = (F_1, \ldots, F_m, p_1, \ldots, p_n)$ is a vector of input distributions F_i and probability p_i. Hence, the applications setting leads to a family of models $(\Omega, \mathcal{F}, P(F))$ indexed by F.

(2.3) EXAMPLE.

Consider the GI/G/1/∞ queue (see HEYMAN and SOBEL (1982)) in which inter-arrivals have common distribution F_1 and service times have common distribution F_2. One can view the GI/G/1/∞ queue as a family of models indexed by $F = (F_1, F_2)$.

(2.4) EXAMPLE.

A generalized semi-Markov process (GSMP) is a mathematical formalization of a discrete-event stochastic system (see GLYNN (1983)). Such systems are characterized by certain distributions $F_1, \ldots, F_m$ which govern the way clocks are re-set, and routing probabilities $p_1, \ldots, p_n$, which determine "physical state" transitions within the simulation. Thus, a GSMP can be viewed as a model indexed by $F = (F_1, \ldots, F_m, p_1, \ldots, p_n)$.

To assess the performance of the stochastic system under considerations, it is common to use a real-valued functional of the form (2.1). Since the expectation depends on P = P(F), it is clear that the quantity α can be viewed as a function of F:

$$\alpha = \alpha(F)$$

Statistical methods become relevant in trying to characterize the uncertainty in performance α of a stochastic system due to lack of knowledge of the input variable F. More specifically, when the practitioner wishes to incorporate prior knowledge of the "true" value of F into his or her uncertainty characterization, Bayesian methods are essential. To summarize:

Bayesian methods are essential, in order to incorporate a practitioner's prior on the input distributions driving the simulation.

In other words, we are not suggesting that a Bayesian philosophy is appropriate to Monte Carlo simulation itself; it is, however, an important tool in the decision context that is typical of Industrial Engineering applications of Monte Carlo simulation.

3. A BAYESIAN FRAMEWORK FOR DISCRETE-EVENT SIMULATION

As indicated in Section 2, a Bayesian framework makes sense in a stochastic simulation environment. We now wish to describe the framework in more detail.

To accomplish this goal, we shall specialize our discussion to stochastic simulations of discrete-event type. Such simulations can be characterized as GSMP's (see Example 2.4). In this case, the simulation is characterized by:

i) clock-setting distributions $F_1, \ldots, F_m$

ii) routing probabilities $p_1, \ldots, p_n$

iii) an initial distribution μ for starting the simulation

In most design problems of interest, one can assume that the initial distribution μ is determined by modeling constraints, and is not dependent on external "real-world" data. For example, if the decision problem involves studying steady-state behavior of a GSMP, any initial distribution can generally be used to initiate the simulation, so that μ can be chosen to be independent of "real-world" data. If μ can be ignored in this way, then the system under study can be viewed as a class of GSMP's parameterized by $F_1, \ldots, F_m, p_1, \ldots, p_n$.

In order to specify a prior on the index set $F = (F_1, \ldots, F_m, p_1, \ldots, p_n)$, it is convenient to assume that the F_i's are members of parametric families. In particular, suppose that $F_i = F_i(\gamma_i)$ where γ_i is an appropriate set of parameters for the family $F_i(.)$. This reduces the dimensionality of the index set from infinite dimensional to finite dimensional.

Specifically, the system can now be viewed as a class of GSMP's indexed by $\theta = (\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n) \in R^{\mathcal{P}}$ ($\mathcal{P} = n + d_1 + \ldots + d_m$, where $\gamma_i \in R^{d_i}$).

The performance of the system is measured by some real-valued functional of the form (2.1). Since the probability measure driving the stochastic system depends on θ, it follows that EX is representable as some function of θ, call it $\mu(\theta)$. Thus, $\alpha = \alpha(\theta)$ when

$$\alpha(\theta) = g(\mu(\theta))$$

It seems worth mentioning that the function g is necessary, in order to incorporate applications in which the performance functional involves central moments and/or ratios of expectations (as in the regenerative method; see IGLEHART (1978)).

Suppose that the practitioner has a prior $\varphi(d\theta)$ on the parameter θ. This induces a prior $\eta(d\alpha)$ on the performance functional α :

$$\eta(d\alpha) = \int_{\{\theta:\, g(\mu(\theta)) \in d\alpha\}} \varphi(d\theta)$$

If "real-world" data for θ is available (ie. observations for the $F_i(\alpha_i)$'s, ρ_i's), then the prior can be combined with the data to yield a posterior $\Phi(d\theta)$ on θ. This then leads to a posterior on the performance functional α :

$$(3.1) \qquad \gamma(d\alpha) = \int_{\{\theta:\, g(\mu(\theta)) \in d\alpha\}} \Phi(d\theta)$$

The posterior γ can be used to compute the posterior mean of α ; this is a Bayesian point estimate for the performance of the system. A further calculation leads to credible sets for α (see BERGER (1985)); this can be viewed as a Bayesian confidence interval for α. Basically, the posterior γ summarizes the statistical information relative to the system performance α. In the next section, we turn to describing research problems associated with the computation of γ .

4. RESEARCH PROBLEMS

As indicated in Section 3, the basic problem is:

(4.1) Given:

1) a GSMP indexed by $(F_1(\gamma_1), \ldots, F_m(\gamma_m), p_1, \ldots, p_n)$
2) a prior on $\theta = (\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n)$
3) "real-world" data on the $F_i(\alpha_i)$'s, p_i's

Compute:

1) a posterior γ on $\alpha = \alpha(\theta)$
2) the posterior mean
3) a credible set

In many settings, it is reasonable to expect that the posterior $\Phi(d\theta)$ on θ can be computed using standard Bayesian methods. For example, if the prior θ chosen from an appropriate conjugate family, then the posterior Φ can usually be analytically calculated.

Consequently, the difficulty is computing (3.1) is performing the appropriate integration. Some comments on integration problems in the classical Bayesian context can be found on pages 262-267 of BERGER (1985). The integration problem posed here is just a more complicated variant of the classical problem.

A first cut at solving this integration numerically starts by discretizing the posterior Φ by a point mass distribution $\tilde{\Phi}$ ie. $\tilde{\Phi}$ assigns mass p_k to point θ_k, $1 \le k \le \ell$. For each θ_k, one then simulates the GSMP, and estimates $\mu(\theta_k)$ by $\hat{\mu}(\theta_k)$. The posterior γ is then approximated by the measure $\tilde{\gamma}$ which assigns mass p_k to point $g(\hat{\mu}(\theta_k))$, $1 \le k \le \ell$. This solution strategy combines classical numerical integration (ie. the discretization step) with Monte Carlo simulation (calculating $\hat{\mu}(\theta_k)$). Of course, another alternative would be to sample the posterior Φ , obtaining random deviates $\theta_1, \ldots, \theta_\ell$, and approximate γ by the empirical distribution $\tilde{\tilde{\gamma}}$ which assigns mass $1/\ell$ to $g(\hat{\mu}(\theta_k))$, $1 \le k \le \ell$. In general, however, strategies involving appropriately modified classical numerical integration schemes are to be preferred to pure Monte Carlo algorithms, due to their improved variance properties.

Note that the computational effort required to compute a posterior on $\tilde{\gamma}$ is, roughly speaking, ℓ times the effort required to do an ordinary Monte Carlo simulation of the system. Thus, it is imperative to consider methods that would improve the efficiency of the numerical algorithm described above. Among the possible ideas to be investigated are:

i) methods for constructing a good approximation $\tilde{\Phi}$
ii) how to optimally split simulation effort among $\hat{\mu}(\theta_1), \ldots, (\theta_\ell)$
iii) using antithetic r.v.'s to induce correlation among $\hat{\mu}(\theta_1), \ldots, \hat{\mu}(\theta_\ell)$
iv) using derivative information on Φ to develop improved approximations of γ .

Some of these ideas are explored in GLYNN (1986). We also wish to point out that the calculation of a credible set from the approximating posterior $\tilde{\tilde{\gamma}}$ is non-trivial; this is an important topic deserving further attention.

In some applications settings, (4.1) is not a natural formulation of the Bayesian problem. In particular, the practioner may have significantly greater intuition about the distribution of α than that of θ. For example, in a manufacturing setting, the engineer may well know more about the assembly line production rate than the parameters associated with individual machine down-times. Further research is required on this important topic.

REFERENCES

Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis. Springer-Verlag, New York.

Ethier, S.N. and Kurtz, T.G. (1986). Markov Processes: Characterization and Convergence. John Wiley, New York.

Glynn, P.W. (1983). On the Role of Generalized Semi-Markov Processes in Simulation Output Analysis. Proceedings of the 1983 Winter Simulation Conference, 39-42.

Glynn, P.W. (1983). Computational Methods for Bayesian Simulation. Forthcoming technical report. University of Wisconsin.

Grassberger, P. and De La Torre, A. (1979). Reggon Field Theory (Schlogle's First Model) on a Lattice: Monte Carlo Calculations of Critical Behavior. Ann. Physics 122, 373-396.

Heyman, D.P. and Sobel, M.J. (1982). Stochastic Models in Operations Research, Volume 1. McGraw-Hill, New York.

Iylehart, D.L. (1978). The Regenerative Method for Simulation Analysis. Trends in Programming Methodology, III, Software Modeling (edited by K.M. Chandy and R.T. Yeh). Prentice-Hall, Englewood Cliffs, N.J.

Ross, S.M. (1980). Introduction to Probability Models. Academic Press, New York.

BIOGRAPHY:

Peter W. Glynn received his Ph.D. in Operations Research from Stanford University in 1982, and is currently an assistant professor at the University of Wiscosin - Madison. His research interests include queueing theory and computational probability. He is a meember of ORSA, TIMS, IMS, and the Statistical Society of Canada.

Peter W. Glynn
Dept. of Industrial Engineering
University of Wisconsin-Madison
1513 University Avenue
Madison, WI 53706
(608) 263-6790