Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

BAYESIAN ANALYSIS FOR SIMULATION INPUT AND OUTPUT

Stephen E. Chick

Department of Industrial and Operations Engineering
The University of Michigan
1205 Beal Avenue
Ann Arbor, Michigan 48109-2117, U.S.A.

ABSTRACT

The paper summarizes some important results at the intersection of the fields of Bayesian statistics and stochastic simulation. Two statistical analysis issues for stochastic simulation are discussed in further detail from a Bayesian perspective. First, a review of recent work in input distribution selection is presented. Then, a new Bayesian formulation for the problem of output analysis for a single system is presented. A key feature is analyzing simulation output as a random variable whose parameters are an unknown function of the simulation's inputs. The distribution of those parameters is inferred from simulation output via Bayesian response-surface methods. A brief summary of Bayesian inference and decision making is included for reference.

1 INTRODUCTION

A decade ago, Glynn (1986) argued that 'Bayesian statistical methodology has an important role to play in the theory and practice of stochastic simulation'. At present, however, the number of contributions of applying Bayesian techniques to problems in stochastic simulation has been limited. Frequentist techniques are predominantly employed for solving problems such as input distribution selection, input parameter selection, output analysis for a single simulated system, selection of the best simulated system, simulation model validation, and statistical issues related to variance reduction.

Still, there is a recent resurgence of interest in Bayesian methodologies for problems in the analysis of stochastic simulations. This interest has been fueled, in part, by the impressive recent analytical and computational research which has enabled the implementation of Bayesian statistical calculations. Further, the Bayesian approach provides an elegant framework for modeling statistical problems.

This paper reviews some of the work done at the intersection of the Bayesian statistics and simulation communities, and proposes a Bayesian framework for analyzing output of a single simulated system. A key feature of the proposal is the inference of the entire distribution of the simulation output, rather than a focus on its mean. Issues such as model misspecification and relationships to currently-used frequentist techniques are also discussed.

2 BAYESIAN APPROACH

This section presents some general ideas in Bayesian statistics and to serve as a reference for concepts used in the remainder of the paper. A central conceptual difference between frequentist approaches and Bayesian approaches is that the Bayesian framework models all uncertainty in terms of probability distributions. This includes unknown parameters. Suppose θ is an unknown quantity, $p_\Theta(\theta) = f_\Theta(\theta)d\theta$ is a prior distribution which represents uncertainty about θ based on prior knowledge, and $f_{X|\theta}(x)$ is the likelihood of observing x, assuming θ were true. Then the posterior distribution of θ, given an observation of the data x, is given by Bayes rule,

$$p_{\Theta|x}(\theta) = \frac{f_{X|\theta}(x)\, p_\Theta(\theta)}{K(x)} \tag{1}$$

where $K(x) = \int_\Theta f_{X|\theta}(x)\, \pi(\theta)$.

The value $\tilde{\theta}$ which maximizes $f_{\Theta|x}\theta$, is called a maximum a posteriori estimate (MAP) of θ, and is a Bayesian analog of the maximum likelihood estimator (MLE). A Bayesian analog for confidence interval is a set where the posterior probability of θ is maximal.

Hypothesis testing for a Bayesian is also treated using random variables. Suppose a finite set of hypothesis are proposed. Posit a prior belief that each hypothesis is true, for example $p_H(H_0)$. The data x can be used to infer the probability that a given hypothesis is true (e.g., $p_{H|x}(H_0)$).

Bayesian statistics is often used in conjunction with decision making where the objective is to maximize the expected utility of a decision (see, e.g., Berger, 1985, or de Groot, 1970). This paper refers to the mean of output values rather than expected utility, as current work in stochastic simulation analysis generally focuses on estimation of means.

3 LITERATURE REVIEW

There are two important domains of literature in which simulation and Bayesian statistics have been discussed together. The first is the application of Bayesian and decision-theoretic techniques to statistical problems in the field of simulation. The second is the use of simulation as a numerical tool for approaching problems in Bayesian inference. This paper is primarily concerned with developments in the first domain–Bayesian approaches to statistical problems involved with analyzing simulations. Much of the analysis that this entails requires techniques from the second domain.

3.1 Bayesian Techniques for Simulation

Applications of Bayesian techniques to simulation problems can be classified into two related subareas: the simulation of stochastic systems and the simulation of deterministic, but unknown systems. For the analysis of stochastic simulations the literature is not extensive.

Stochastic simulations. Andrews and Schriber (1983) appear to be the first to discuss modeling simulation output with a Bayesian formalism. They construct a point estimator and a Bayesian confidence interval for the mean of batch-run simulation output. Specific assumptions include a Gaussian prior for the mean, and a stationary Gaussian process with autocorrelated output from batch to batch. Andrews, Birdsall, Gentner, and Spivey (1986) investigate Bayesian techniques for validation of simulation output.

Glynn (1986) describes an attractive, general framework for modeling a generalized semi-Markov process (GSMP) when the parameters of the input distributions are unknown (e.g., service times are exponential, with unknown rate μ). He notes that the probability distribution of the output depends on the prior distribution of the input parameters, and comments on potential research directions.

Nelson, Schmeiser, Taaffe and Wang (1997) evaluated several techniques for combining a deterministic approximation with a stochastic simulation estimator, among them a Bayesian analysis (Gaussian distribution) for a point estimator.

Wang and Schmeiser (1997) describe the use of Monte Carlo simulation for a Bayesian analysis. Additionally, they formulate a related optimization problem to select a prior distribution satisfying certain desirable properties. In this respect, they perform a Bayesian robustness analysis for analyzing Monte Carlo simulation output.

Chick (1996) addresses the problem of selecting an appropriate input distribution (e.g., exponential, gamma) using Bayesian hypothesis testing, as well as the problem of parameter uncertainty for a given input distribution. He also developed extensions to Latin hypercube sampling to provide variance reduction even when the input distribution (and therefore the number of parameters) is unknown.

Andradóttir and Bier (1997) discuss possible roles of Bayesian analysis in model validation, and output analysis with both normal and truncated normal distributions. They present results on importance sampling when the parameters of input distributions are unknown. A joint distribution for the input and output parameters is discussed, but a number of analytical and practical difficulties were described.

Deterministic simulations. A different angle is taken for the problem of applying Bayesian techniques to deterministic computer simulations. The problem is that a computer algorithm calculates a function of certain inputs and provides a deterministic output. Each evaluation of the algorithm is assumed to be computationally expensive. Bayesian techniques are used to infer the parameters of a postulated functional form for the algorithm's output. The design of computer experiments to learn the shape of the function is a central focus. Koehler and Owen (1995) provide a review of techniques in this well-explored area.

One formulation is the Kriging model, which assumes that the unknown real-valued function is

$$\Xi(x) = \sum_{j=1}^{N} \phi_j h_j(x) + Z(x),$$

where the $h_j(\cdot)$ are known functions, N is known, the ϕ_j are unknown with a given prior, and $Z(x)$ is a stationary Gaussian random process with

$$\mathrm{Cov}(Z(x_i), Z(x_j)) = \sigma^2 R(x_j - x_i)$$

Evaluations of Ξ at the points $x_0, \ldots, x_n$ and Bayes rule are used to infer the values of the ϕ_j (and therefore Ξ) at other values of the input x.

Currin, Mitchell, Morris, and Ylvisakir (1991) provide results for selecting points x_i which provide maximal inferential power for a special case of the Kriging model. Morris, Mitchell, and Ylvisakir (1993) extend the framework to accomodate simulation output which returns derivatives of Ξ as well as Ξ for each function evaluation.

Osio and Amon (1997) expand the problem to plan computer experiments where several functions of differing levels of accuracy could be programmed, the more simplistic models being quicker but less accurate. They evaluate orthogonal arrays of computer experiments, and select the array which gives the maximum expected information gain. They iterate with more refined computer models as necessary.

Chaloner and Verdinelli (1995) provide a thorough overview of Bayesian experimental design in general.

3.2 Simulation for Bayesian Analysis

Bayesian statistics presents a number of challenges for numerical analysis, notably for predictive inference when the denominator in Equation (1) is needed. A recent survey paper (Evans and Swartz, 1995) indicates that significant progress has been made using five general techniques: asymptotic methods, importance sampling, adaptive importance sampling, multiple quadrature, and Markov chain methods. Simulation techniques are therefore playing an important role in making a Bayesian analysis computationally tractable (Chen and Schmeiser, 1993).

Gilks, Richardson, and Spiegelhalter (1996) provide a comprehensive review of many of the theoretical, philosophical and practical issues related to Markov chain Monte Carlo (MCMC) techniques. The MCMC WWW site (`http://www.stats.bris.ac.uk/MCMC/`) contains a significant collection of references. Spiegelhalter et al. (1996) provide a package called BUGS to perform Bayesian inference via MCMC methods (`mailto:bugs@mrc-bsu.cam.ac.uk`). The software can be used to generate samples from posterior distributions that arise in Bayesian analysis.

4 INPUT DISTRIBUTION SELECTION

Bayesian formulations for statistical distribution selection have been applied to a number of fields (Draper, 1995; Madigan and York, 1995; Kass and Raftery, 1995; Raftery, 1995; Volinsky et al., 1996). A central idea is that all uncertainty, including distribution uncertainty, is to be represented by probability statements. The formulation presented by Chick (1996) can be summarized as follows.

Suppose that a statistical distribution and parameter for a sequence of random quantities is needed for input to a discrete-event simulation of a dynamic system, and some historical data $\vec{y}_n = (y_1, \dots, y_n)$ is available. Suppose that the data are believed to be conditionally independent, given the statistical distribution and parameter.

Set $\pi(m)$ to be the prior belief that distribution m is the correct distribution, $m = 1, \dots, M$. (A common, but not necessary, choice for $\pi(m)$ is $1/M$.) Set $p_{\Lambda_m|m}(\lambda_m)$ to be the prior belief that λ_m is the true parameter, given the assumption that m is the correct distribution.

Under these conditions, it is possible to determine:

$$p_{\vec{Y}_n|m}(\vec{y}_n) = \int_{\Lambda_m} p(\vec{y}_n \mid m, \lambda_m) p_{\Lambda_m|m}(\lambda_m) d\lambda_m$$

$$p_{m|\vec{y}_n}(m) = \frac{p_{\vec{Y}_n|m}(\vec{y}_n)\pi(m)}{\sum_{k=1}^{M} p_{\vec{Y}_n|k}(\vec{y}_n)\pi(k)} \quad (2)$$

$$p_{\Lambda_m|m,\vec{y}_n}(\lambda_m) = \frac{p_{\vec{Y}_n|m,\lambda_m}(\vec{y}_n) p_{\Lambda_m|m}(\lambda_m)}{p_{m|\vec{y}_n}(m)}$$

Equation (2) describes the belief that a specified distribution is correct, given the data, prior beliefs about the correct distribution, and the assumption that the correct distribution is to be found in the original set of M proposed distributions. Several approximation methods exist for the above integral. Of particular interest is the Laplace approximation,

$$p_{m|\vec{y}_n}(m) \approx \frac{\left|\tilde{\Sigma}_m\right|^{1/2} p_{\vec{Y}_n|m,\tilde{\lambda}_m}(\vec{y}_n) p_{\Lambda_m|m}\left(\tilde{\lambda}_m\right)\pi(m)}{(2\pi)^{-d_m/2} \sum_{i=1}^{M} p_{i|\vec{y}_n}(i)}$$

where the approximation error is $O(n^{-1})$, d_m is the dimension of λ_m, $\tilde{\Sigma}_m^{-1} = -\mathbf{D}^2 \log p\left(\tilde{\lambda}_m \mid m, \vec{y}_n\right)$ is minus the Hessian of the log-posterior evaluated at the MAP estimate $\tilde{\lambda}_m$ given m, and n is the number of data points. Chick (1996) provides an example where the approximation worked well.

The $\tilde{\Sigma}_m$ term is analogous to the information matrix term used for providing confidence intervals for the MLE of frequentist approaches (Leemis, 1995).

A review of work investigating the problem of automatically selecting an appropriate prior is given by Kass and Wasserman (1996). A discussion of the robustness of this selection process with respect to poor choices of prior distributions is given by Berger (1994). The problem of not including the correct distribution in the original set of distributions is discussed in Section 7. The probabilistic interpretation of distribution correctness is not without its detractors. See for example Edwards (1994).

5 OUTPUT ANALYSIS

Consider the problem of analyzing the output from a single simulated system (see Chick, 1997, for an extension to multiple systems). Focus here is on output which is independent from simulation replication to replication (batch mean output is not considered). Curiously, simulation output data has traditionally been handled quite differently than data for simulation input distributions, in that much focus has been on estimating means of output, rather than the assessing the entire distribution using parameter estimation and goodness-of-fit techniques. (Although Law and Kelton, 1991; Glynn, 1996, among others discuss quantiles of simulation output.)

The entire output distribution can be important to a decision maker (Law and Kelton, 1991; Banks et al., 1996). From the distribution, any quantity of interest can be deduced, including the mean, variance, other moments, quantiles, and other functionals of the probability distribution.

In general, the form of the distribution function and/or the values of the statistical parameters of the output will be unknown to the simulation analyst. On the other hand, there may be some intuition or formal analysis indicating the form of the output distribution. Analytical methods may indicate that the Gaussian or geometric or some other distribution may be appropriate, without giving precise information regarding the parameters. For example, the number in system of an M/M/1 in steady-state is geometric, or some other quantity of interest may have a functional central limit theorem indicating it will be close to Gaussian. It is known, however, that the parameters of the output distributions are some unknown function of the input parameters.

This thought process is depicted in Figure 1. For clarity of exposition, we summarize under the assumption that $M = 1$, and drop the subscript m. A superscript r is used to emphasize quantities specific to replication r, and a subscript i indicates the i-th coordinate of a vector.

1. A statistical parameter λ^r is selected by the simulation analyst (by sampling or other choice) for replication r.

2. Random variates $x_1^r, x_2^r, \ldots$ (service times, routing decisions, ...) are generated from distributions depending on λ^r.

3. A random output o^r is generated.

4. The analyst assumes a parametric distribution $f_{O^r|\theta^r}(o^r)$ to describe the output, where the unknown θ^r may depend on λ^r.

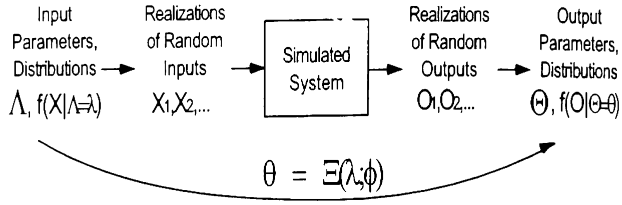

Figure 1: A Framework for Analyzing Stochastic Simulation Experiments (r is for replication number)

5. An unknown deterministic function Ξ maps input to output parameters, $\theta^r = \Xi(\lambda^r; \phi)$, where $\phi = (\phi_1, \ldots \phi_N)$ are unknown parameters with prior distribution $p_\Phi(\phi)$.

$\Lambda, \Theta, \Xi, \Phi$, and O may be scalar or vector valued.

Example 1. *Consider the M/M/1 queue, with input parameter* $\lambda = (\lambda_1, \lambda_2)$, *setting the arrival rate to be* λ_1 *and the service rate to be* λ_2. *Suppose that the number* Q *in the system at time* t *is desired. Although the distribution of* Q *can be determined analytically, suppose we 'guess that the output* $O_i = Q_i$ *comes from a series of independent simulation replications is geometric*(θ)*, where* θ *is postulated to be of the form* $\theta^r = \Xi(\lambda^r; \phi) = \phi_1 + \phi_2\lambda_1^r/\lambda_2^r + \phi_3\lambda_2^r/\lambda_1^r$.

One decision theoretic formulation of the relationships between simulation inputs and outputs is summarized in Figure 2. The figure is drawn as a Bayesian network, where probabilistic dependencies are represented with arcs (Howard and Matheson, 1984). The figure represents that for replication r, simulated variates X_i^r depend on the input parameter Λ and uniform variates U_j^r from the simulator; the output O^r can be described in terms of parameters of the output distribution Θ and the X_i^r; and the output parameters depend on Λ and Φ. In probability terms, $p_{\Lambda,\Phi,\Theta,U,X,O} = p_\Lambda p_U p_\Phi p_{O|\Theta,X} p_{\Theta|\Lambda,\Phi} p_{X|\Lambda,U}$ (some subscripts are omitted for clarity).

The relationship between $X_1^r, \ldots$ and O^r is too complex to study analytically. This complexity is one reason the system is simulated in the first place.

Figure 3 presents a simplified model obtained by conceptually 'integrating out' U and X. The complexity of random number generators and simulated variates is captured only implicitly. This simplified model, which represents $p_{\Lambda,\Phi,\Theta,O} = p_\Lambda p_\Phi p_{O|\Theta} p_{\Theta|\Lambda,\Phi}$, is the basis for the output analysis discussed in this paper.

From Figure 3, pairs of simulation inputs and outputs can be used to infer the parameters ϕ of the mapping Ξ. Suppose that R replications are run with output $\vec{o} = (o_1, \ldots, o_R)$, and denote by $\vec{\lambda} = (\lambda^1, \ldots, \lambda^R)$

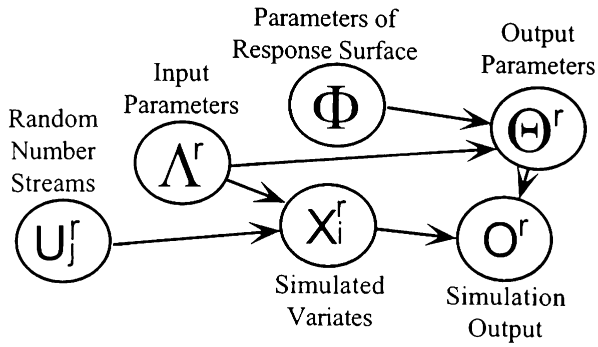

Figure 2: Approach One: Comprehensive Belief Network

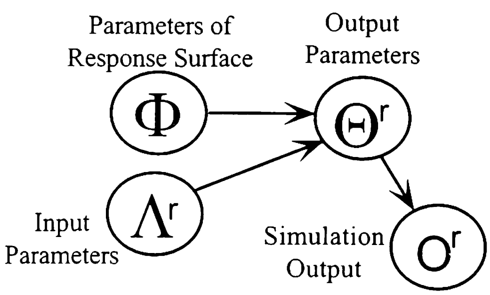

Figure 3: Approach Two: Simplified Belief Network

the input parameters used during those replications. The posterior distribution for ϕ is then

$$p_{\Phi|\mathcal{D}}(\phi) \propto p_{\Phi}(\phi) \prod_{r=1}^{R} f_{O^r|\phi,\vec{\lambda}}(o^r) \qquad (3)$$

where $\mathcal{D} = (\vec{o}, \vec{\lambda})$ represents the information learned from previous simulation replications, and $f_{O^r|\theta^r}(o^r)$ is the postulated distribution of the output o^r, given the input λ^r, mapping parameters ϕ, and $\theta^r = \Xi(\lambda^r; \phi)$.

The posterior distribution on ϕ in Equation (3) induces a posterior distribution on the output of future replications. Further, for each fixed λ, the unknown expected value of the output $\bar{O}(\lambda) = E[O \mid \lambda, \phi]$ is a random variable whose distribution can be determined with Equation (3).

$$\begin{aligned} p_{O|\mathcal{D},\lambda}(o) &= \int_{\phi} f_{O|\theta=\Xi(\lambda,\phi)}(o) p_{\Phi|\mathcal{D}}(\phi) \\ p_{\bar{O}(\lambda)|\mathcal{D},\lambda}(\bar{o}) &= \int_{\phi|E[o|\Xi(\lambda;\phi)]=\bar{o}} p_{\Phi|\mathcal{D}}(\phi) \end{aligned}$$

The distribution of $\bar{O}(\lambda)$ is important as it characterizes the system uncertainty regarding the response surface, and is distinct from systemic uncertainty about input distributions ($p_{\Lambda}(\lambda)$) and stochastic uncertainty of the system ($p_{O|\mathcal{D},\lambda}(o)$).

This can be though of as a Bayesian version of metamodeling (see, e.g., Law and Kelton, 1991). The current approach differs, however, as it infers a probability distribution for the output as a function of inputs, in addition to estimating the mean of the output. A distribution of parameters results, rather than a point estimate of means. The mean of the output can then be calculated is an integration problem, rather than a simulation problem, once the output distribution and parameter θ are known.

The approach is similar to the Kriging model in that there is assumed to be a deterministic, but unknown function, which maps input parameters to output parameters. The approach is different from the Kriging model in that the output of the simulation is stochastic, rather than deterministic, and $p_{O|\mathcal{D},\lambda}(o)$ represents uncertainty in the output rather than a Gaussian process Z.

The current proposal can also be used for determining which of several proposed functions Ξ_i is best supported by the data. See Ledersnaider (1994) for a discussion of Bayesian response surface estimation, and Chick (1997) for a discussion of Bayesian distribution selection.

6 JOINT INPUT-OUTPUT MODEL

A joint probability distribution for input and output parameters (Λ, Θ) can then be written

$$\begin{aligned} p_{\Lambda,\Theta|\mathcal{D}}(\lambda, \theta) &= p_{\Lambda|\mathcal{D}}(\lambda) p_{\Theta|\lambda,\mathcal{D}}(\theta) \\ &= p_{\Lambda} \int_{\phi} p_{\Theta|\phi,\lambda,\mathcal{D}}(\theta) p_{\Phi|\lambda,\mathcal{D}}(\phi) \end{aligned}$$

where $p_{\Lambda,\Theta} = p_{\Lambda} p_{\Theta|\lambda}$ is the prior distribution.

The posterior for λ is written in terms of a non-updated prior because under the assumptions, no information about the input parameter λ can be learned by observing the simulation output for various values of λ. The simulations give information about how λ and θ are related, not about the value of λ most appropriate for the real system.

An interesting research question is how one might learn about λ, given certain 'real-world' data about output and response variables, along with simulation input and output data.

7 MISSPECIFIED DISTRIBUTIONS

For both input and output distributions, an assumption was that the set of proposed distributions included the correct distribution. What if this is not the case? How far off-base can the results be? This section partially addresses these questions.

Suppose that $O_1, \ldots$ are i.i.d. random variables with probability distribution $p'_O(o)$, but that rather than selecting $p'_O(o)$, the simulation analyst chooses a distribution $f_{O|\theta}(o)$ from the exponential family, for some unknown θ, where

$$f_{O|\theta}(o) = F(o)G(\theta)e^{\sum_{s=1}^{S} u_s(o)\phi_s(\theta)} \qquad (4)$$

For example, the normal distribution falls into this category because

$$\begin{aligned} f(o \mid \mu, \sigma^2) &= \frac{e^{-(o-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{e^{-\mu^2/2\sigma^2}}{\sigma} e^{-x^2/2\sigma^2 + x\mu/\sigma^2} \end{aligned}$$

so that $S = 2$, $F(o) = 1/\sqrt{2\pi}$, $\theta = (\mu/\sigma^2, 1/\sigma^2)$, $G(\theta) = \exp\left[-\mu^2/2\sigma^2\right]/\sigma$ $u_1(o) = o$, $u_2(o) = o^2$, $\phi_1(\theta) = \mu/\sigma^2$, and $\phi_2(\theta) = 1/2\sigma^2$.

The following theorem indicates that the MLE and Maximum a posteriori (MAP) estimates of the parameter converge asymptotically to values closely related to the 'true' expectations of functions of the random variables under general conditions.

Theorem 1. *Let $p'_O(o)$ and $f_{O|\theta}(o)$ be as above, and suppose that $E[|u_s(o)|]$ exists for $s = 1, \ldots, S$, where the expectations are with respect to $p'_O(o)$. Further suppose that θ^* is the unique maximizer of*

$$G(\theta) \exp\left[\sum_{s=1}^{S} E[u_s(o)]\phi_s(\theta)\right]. \qquad (5)$$

Then the MLE $\hat{\theta}$ obtained from the misspecified distribution converges almost surely to the value of θ^*. Further, the MAP converges to θ^* as well for conjugate priors with non-zero probability in a neighborhood around θ^*.

Proof. Set $\pi_0(\theta) = K_1 G(\theta)^a e^{\sum_{s=1}^{S} b_s \phi_s(\theta)}$ to be a conjugate prior for θ. Then if $\vec{x}_n = (x_1, \ldots, x_n)$, the posterior is

$$p(\theta \mid \vec{x}_n) = K_2 G(\theta)^{a+n} e^{\sum_{s=1}^{S}(b_s + t_s(\vec{x}_n))\phi_s(\theta)}, \qquad (6)$$

where the sufficient statistics $t_s(\vec{x}_n) = \sum_{j=1}^{n} u_s(x_j)$. By the SLLN, $\lim_{n\to\infty} t_s(\vec{x}_n)/n \to E[u_s(o)]$ almost surely. Then

$$\frac{d\frac{\log p(\theta|\vec{x}_n)}{n}}{d\theta} \overset{n\to\infty}{\longrightarrow} \frac{d \log G(\theta)}{d\theta} + \sum_{s=1}^{S} E[u_j(X)]\frac{d\psi_j(\theta)}{d\theta} \qquad (7)$$

almost surely. Under the assumptions of the theorem, the maximum for both Equation (5) and Equation (6) are zeros of Equation (7), and the MAP converges appropriately. Set $a = b_s = 0$ for convergence of the MLE. □

In particular, the result is true for the Gaussian distribution. Assuming the output is Gaussian results in an output which asymptotically matches the mean and variance of the 'true' distribution.

More precisely, assume the output o_r is Gaussian with unknown mean μ and precision τ. Use the gamma-normal conjugate prior, with τ distributed Gamma (α_0, β_0), and μ given τ distributed conditionally Gaussian, $\mathbf{N}(\mu_0, n_0\tau)$.

Then the posterior distribution $p(\mu \mid \mathcal{D})$ of the expected value of the output μ has Student distribution

$$p(\mu \mid \mathcal{D}) \sim \mathbf{St}\left(\mu_R, (n_0 + R)(\alpha_0 + \frac{R}{2})/\beta_R, 2\alpha_0 + R\right)$$

where $\mathbf{St}(\mu, \beta, \alpha)$ is the univariate Student distribution with α degrees of freedom, mean μ, variance $\alpha/(\beta(\alpha - 2))$, and

$$\begin{aligned} \mu_R &= \frac{n_0\mu_0 + R\bar{o}}{n_0 + R} \\ \beta_R &= \beta_0 + \frac{s}{2} + \frac{n_0 R(\mu_0 - \bar{o})^2}{2(n_0 + R)} \\ \bar{o} &= \sum_{r=1}^{R_k} o_r/R \\ s &= \sum_{r=1}^{R}(o_r - \bar{o})^2. \end{aligned}$$

The Student distribution with similar parameters $(\mathbf{St}(\bar{o}, R(R-1)/s, R-1))$ is used to construct the standard frequentist confidence interval for the mean. The mean and variance for both the Bayesian and frequentist Student distributions asymptotically look like $\bar{o}$ and R^2/s, respectively. Although the interpretation of the Bayesian posterior for the unknown mean and the frequentist confidence interval have different interpretations, the above Gaussian approximation leads to similar results from a practical standpoint.

8 CONCLUSIONS

This paper reviews the literature for Bayesian approaches to statistical problems in stochastic simulation. In addition, some references were presented to related literature in the areas of Bayesian methodology in deterministic simulation, and the application of simulation techniques to integration problems which arise in Bayesian analysis.

This paper also presented a new formulation for modeling output from stochastic simulations. The work is motivated by two ideas. One, the entire distribution of simulation output is often of interest (rather than just moments) and two, the parameters of the output distribution are some unknown function of the input parameters. The approach advocated here is distinct from, but related to, both the metamodeling approach used in stochastic simulation analysis, and the Kriging model used in deterministic simulation analysis.

The framework makes explicit the differences between stochastic uncertainty (randomness inherent in a system) and two types of systemic uncertainty (the unknown distributions and parameters of the system; and the unknown response of outputs with respect to changes in inputs).

Given particular assumptions, the mean output can be shown to be a random variable with Student distribution.

A number of research directions remain open to a Bayesian analysis. These include: analysis of independent output for multiple systems, analysis of batch mean output, development of tools for eliciting reasonable prior distributions, analysis of approximation issues related to 'small samples' versus 'large samples', sensitivity and robustness analysis, theoretical and implementation work on experimental design problems for simulation replication analysis, and evaluation of the importance of collecting more data about inputs or running more replications to better characterize the output.

ACKNOWLEDGMENTS

Thanks are due to Shane Henderson and Stephen Pollock, who made valuable comments during the writing of this paper.

REFERENCES

Andradóttir, S., and V. M. Bier. 1997. Applying Bayesian ideas in simulation. Department of Industrial Engineering, University of Wisconsin-Madison, Technical Report 97-1.

Andrews, R. W., W. C. Birdsall, F. J. Gentner, and W. A. Spivey. 1986. Validation of microeconomic simulation: A comparison of sampling theory and Bayesian methods. In *Proceedings of the Winter Simulation Conference*, ed. J. R. Wilson, J. O. Henriksen, and S. D. Roberts, 380–383. Institute of Electrical and Electronics Engineers, Inc.

Andrews, R. W., and T. J. Schriber. 1983. A Bayesian batch means methodology for analysis of simulation output. In *Proceedings of the Winter Simulation Conference*, ed. S. Roberts, J. Banks, and B. Schmeiser, 37–38. Institute of Electrical and Electronics Engineers, Inc.

Banks, J., J. S. Carson, II, and B. L. Nelson. 1996. *Discrete-event system simulation.* 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Berger, J. O. 1985. *Statistical decision theory and Bayesian analysis.* 2nd ed. New York: Springer-Verlag.

Berger, J. O. 1994. An overview of robust Bayesian analysis. *TEST* 3:5–124.

Chaloner, K., and I. Verdinelli. 1995. Bayesian experimental design: A review. *Statistical Science* 10(3):273–304.

Chen, M.-H., and B. Schmeiser. 1993. Performance of the Gibbs, hit-and-run, and Metropolis samplers. *Journal of Computational and Graphical Statistics* 2:251–272.

Chick, S. E. 1996. Input distribution selection and variance reduction for discrete-event dynamic simulation: A Bayesian perspective. Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, TR96-14.

Chick, S. E. 1997. Selecting the best system: A decision-theoretic approach. In *Proceedings of the Winter Simulation Conference*, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson. Institute of Electrical and Electronics Engineers, Inc.

Currin, C., T. Mitchell, M. Morris, and D. Ylvisakir. 1991. Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association* 86:953–963.

de Groot, M. H. 1970. *Optimal statistical decisions.* New York: McGraw-Hill, Inc.

Draper, D. 1995. Assessment and propogation of model uncertainty (with discussion). *Journal of the Royal Statistical Society, Series B* 57(1):45–97.

Edwards, A. W. F. 1984. *Likelihood.* Cambridge: Cambridge University Press.

Evans, M. and T. Swartz. 1995. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science* 10(3):254–272.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. *Markov chain monte carlo in practice.* London: Chapman and Hall.

Glynn, P. 1986. Problems in Bayesian analysis of stochastic simulation. In *Proceedings of the Winter Simulation Conference*, ed. J. R. Wilson, J. O. Henriksen, and S. D. Roberts, 376–383. Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W. 1996. Importance sampling for Monte Carlo estimation of quantiles. Department of Operations Research, Stanford University, Technical Report.

Howard, R. A., and J. E. Matheson. 1984. Influence diagrams. In *Readings in the Principles and Applications of Decision Analysis*, ed. R. A. Howard, and J. E. Matheson, Strategic Decision Group, Menlo Park, CA.

Kass, R. E., and A. E. Raftery. 1995. Bayes factors. *Journal of the American Statistical Association* 90(430):773–795.

Kass, R. E., and L. Wasserman. 1996. The selection of prior distributions by formal rules. *Journal of the American Statistical Association* 91(435):1343–1370.

Koehler, J. R., and A. B. Owen. 1995. Computer experiments. Statistics Department, Stanford University, Technical Report.

Law, A. M., and W. D. Kelton. 1991. *Simulation modeling & analysis*. 2nd ed. New York: McGraw-Hill, Inc.

Ledersnaider, D. L. 1994. *A multi-model, Bayesian, resampling, sequential experimental design for response surface estimation*. Ph.D. thesis, University of Michigan, Ann Arbor. Deptartment of Industrial and Operations Engineering, unpublished.

Leemis, L. M. 1995. Input modeling for discrete-event simulation. In *Proceedings of the Winter Simulation Conference*, ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, 16–23. Institute of Electrical and Electronics Engineers, Inc.

Madigan, D., and J. York. 1995. Bayesian graphical models for discrete data. *International Statistical Review* 63(2):215–232.

Morris, M., T. Mitchell, and D. Ylvisakir. 1993. Bayesian design and analysis of computer experiments: Use of derivative in surface prediction. *Technometrics* 35(3):243-255.

Nelson, B. L., B. W. Schmeiser, M. R. Taaffe, and J. Wang. 1997. Approximation-assisted point estimation. Forthcoming, *Operations Research Letters*.

Osio, I. G. and C. H. Amon. 1997. An engineering design methodology with multistage Bayesian surrogates and optimal sampling. Forthcoming, *Journal of Research in Engineering Design*.

Raftery, A. E. 1995. Bayesian model selection in social research (with discussion by Andrew Gelman & Donald B. Rubin, and Robert M. Hauser, and a rejoinder). In *Sociological Methodology 1995*, ed. P. V. Marsden, Cambridge, Massachusetts: Blackwells.

Spiegelhalter, D. J., N. G. Best, W. R. Gilks, and H. Inskip. 1996. Hepatitis B: a case study in MCMC methods. In *Markov Chain Monte Carlo in Practice*, ed. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, London: Chapman and Hall.

Volinsky, C. T., D. Madigan, A. E. Raftery, and R. A. Kronmal. 1996. Bayesian model averaging in proportional hazard models: Assessing stroke risk. Department of Statistics, University of Washington, Seattle, WA, Technical Report no. 302.

Wang, J., and B. W. Schmeiser. 1997. Monte Carlo estimation of Bayesian robustness. Forthcoming, *IIE Transactions*.

AUTHOR BIOGRAPHY

STEPHEN E. CHICK is an assistant professor of Industrial and Operations Engineering at the University of Michigan, Ann Arbor. In addition to simulation, his research interests include engineering probability, Bayesian statistics in system design, reliability, decision analysis, and computational methods in statistics. His work experience includes several years of material handling system design for the automotive industry using simulation analysis.