A Bayesian Framework for the Incorporations of Priors and Sample Data in Simulation Experiments

D.F. Muñoz[*] and D.G. Muñoz

Department of Industrial & Operations Engineering, Instituto Tecnológico Autónomo de Mexico, Mexico

Abstract: In this article, we propose a theoretical framework to estimate performance measures in simulation experiments, incorporating both sample data from a random component and priors on input parameters of the simulation model. Our approach takes into account both the inherent uncertainty of the model as well as parameter uncertainty. We discuss the estimation of a conditional expectation under a Bayesian framework and point and variability estimators are proposed when direct sampling from the posterior distribution is not allowed. The application and properties of the proposed methodology are illustrated through an inventory model and simulation experiments using a Markovian model.

Keywords: Bayesian estimation, simulation input analysis, simulation output analysis, conditional expectations.

1. INTRODUCTION

When using a model for decision making, it is not uncommon to have uncertainty in one or several parameters of the model, so that the decision maker can provide a prior (i.e., a probability distribution) for these parameters. The existence of priors to quantify parameter uncertainty of a model is quite useful; for example, a parameter related to the demand (e.g., mean time between customer arrivals) for a new business process is very unlikely to be known accurately, so the performance of the process must be evaluated based on a prior for the parameter of the future demand. A prior for a parameter can also be constructed from forecasts based on expert judgment [1]. On the other hand, it is becoming frequent to find that besides a prior, we may have data providing additional information about some parameters; for example, information about demand patterns of a new product can be obtained through acceptance testing [2]. These reasons have motivated us to investigate how we can incorporate both forecasts for a parameter and observations from a random component when using a simulation model to estimate the performance of a system.

A random component (also called random input) of a simulation model is a sequence $U_1, U_2, \ldots$ of random quantities that are needed as input to the simulation. When the U_i's are assumed to be independent and identically distributed (i.i.d.), a random component is identified by the corresponding probability distribution. Random components reflect the inherent uncertainty of the model, frequently called stochastic uncertainty. The distributions corresponding to the random components of a simulation model, however, depend upon parameters that usually reflect the particular scenario of the system that is being simulated. The uncertainty on the value of these parameters constitute another source of

uncertainty that is called parameter uncertainty [3-5] and may be quantified using information about the random component. As we will detail in this paper, parameter uncertainty can be incorporated in a simulation model using observations of a random component and/or priors on the uncertain parameters.

The classical approach to specify the distribution of a random component in a simulation experiment consists of selecting the distribution (and its parameters) that best "fit" a set of observations of the component, which are commonly assumed to be i.i.d., i.e., for a given probability distribution the value of the parameters are set using a point estimation procedure (e.g., the method of maximum likelihood), and the "best distribution" is usually chosen according to a measure of goodness of fit (e.g., the mean square error) [6]. A classical approach, however, can not be applied when there are no observations of the random component (e.g., it is a new scenario). On the other hand, several authors [3, 5] criticized the classical approach for ignoring parameter uncertainty. Since we assume parameter uncertainty in this article, we will adopt a different approach based on Bayesian estimation.

Different estimation methods that incorporate parameter uncertainty have been proposed for input analysis in simulation experiments, some of these [7, 8] attempt to estimate an expected value in the form of $E\left[W \mid \Theta = \hat{\theta}(x_0)\right]$, where W is the output of a simulation (see equation (4) below), Θ denotes a parameter and $\hat{\theta}(x_0)$ is a function of the sample data x_0. In this article, however, we adopt a Bayesian approach, since do not assume the existence of a particular value for the parameter, instead we quantify parameter uncertainty through a prior (when no data is available) and a posterior probability distribution (when data is available). Since this article suggests the estimation of conditional expectations, our approach has more similarity with that of [3, 5, 9] where the estimation of performance measures in the

*Address correspondence to this author at the Department of Industrial & Operations Engineering, Instituto Tecnológico Autónomo de Mexico, Mexico; E-mail: davidm@itam.mx

form of $E[W|x_0]$ from the output of simulation experiments is discussed.

In this article we present a theoretical framework to estimate a conditional expectation (given the data) that may be used as a performance measure of the transient behavior of a simulation model, and extend the results initially presented in [10]. The organization of the paper is as follows. In section 2 we introduce the notation and mathematical framework that is appropriate to formulate our estimation problem, and present two examples that illustrate its potential applications. In section 3 we discuss different methodologies to solve the estimation problem. In section 4 we present experimental results from applying the proposed methodologies to a specific example, and finally in section 5 we present our conclusions and directions for future work.

2. MATHEMATICAL FRAMEWORK

We assume that the output of our simulation model depends upon two random components with density (probability in the discrete case) functions $f(x|\theta_1)$ and $g(y|\theta_2)$ respectively, where $\{f(x|\theta_1), \theta_1 \in P_1\}$ and $\{g(y|\theta_2), \theta_2 \in P_2\}$ are families of density functions that depend upon the parameters θ_1 and θ_2, with P_1 and P_2 as their corresponding parameter spaces. Under a Bayesian framework, the parameters of the density functions are random variables Θ_1, Θ_2 and $\theta_1 \in P_1$, $\theta_2 \in P_2$ denote particular values for the parameters. We denote $\Theta = (\Theta_1, \Theta_2)$, $\theta = (\theta_1, \theta_2)$, and $P = P_1 \otimes P_2$.

For the first random component we have a set of observations $x_0 = (x_1, x_2, \ldots, x_n)$ that arise from a random sample $X = (X_1, X_2, \ldots, X_n)$ of the density function $f(x|\theta_1)$. In other words (see e.g., Bernardo and Smith, 1994), the conditional (given $[\Theta_1 = \theta_1]$) density of $X_1, X_2, \ldots, X_n$ satisfies:

$$f_n(x_0|\theta_1) = f(x_1|\theta_1)f(x_2|\theta_1)\ldots f(x_n|\theta_1), \tag{1}$$

where $x_0 = (x_1, \ldots, x_n) \in \Re^n$ and $\theta = (\theta_1, \theta_2) \in P$. In our particular applications the sample data x_0 provides information only on Θ_1, so that we assume that the sample information is independent of the parameter Θ_2, in other words, the conditional (given $[\Theta = \theta]$) density of $X_1, X_2, \ldots, X_n$ satisfies:

$$f_n(x_0|\theta) = f_n(x_0|\theta_1), \tag{2}$$

where $x_0 = (x_1, x_2, \ldots, x_n) \in \Re^n$, $\theta = (\theta_1, \theta_2) \in P$, and $f_n(x_0|\theta_1)$ is defined in (1).

For the second random component there are no observations, but there is a prior density function $p_{\Theta_2}(\theta_2)$. We also assume the existence of a prior density function for the parameter Θ_1, which is not restrictive, since we can use a non-informative prior distribution when we wish to consider a prior that "favors" no possible values of Θ_1 over others [11,

12]. For simplicity, we assume the priors for Θ_1 and Θ_2 are independent, in other words, the joint density of Θ_1, Θ_2 satisfies:

$$p_\Theta(\theta) = p_{\Theta_1}(\theta_1)p_{\Theta_2}(\theta_2), \tag{3}$$

for any $\theta = (\theta_1, \theta_2) \in P$.

Let $Y = \{Y(s), s \geq 0\}$ be the stochastic process (possibly $\Re^d$-valued) that represents the output of a simulation experiment. Since we wish to estimate performance measures for transient simulation, we let T denote the run length and assume that T is a stopping time with respect to the process Y [13]. Using this notation, we define our performance measure as:

$$r(x_0) = E[W|X = x_0], \tag{4}$$

where $W = g(Y(s), 0 \leq s \leq T)$ is the random variable that reflects the system's performance. We remark that the a expected value is usually chosen as a good performance measurement since it can be interpreted as a long-run average.

We assume that the random sample X is independent of the simulation output, on which case, if $h(w|x_0, \theta)$ and $h(w|\theta)$ denote the conditional density functions of W (given $[X_0 = x_0, \Theta = \theta]$, and $[\Theta = \theta]$, respectively), then:

$$h(w|x_0, \theta) = h(w|\theta), \tag{5}$$

for $w \in \Re$, $x_0 = (x_1, x_2, \ldots, x_n) \in \Re^n$ and $\theta \in P$.

Our interest is not restricted to propose point estimation procedures, but also to assess the precision of the point estimators through an asymptotic confidence interval [14].

It is worth mentioning that in the case where no sample data is available, the performance measure (4) incorporates only information from a prior on the parameter Θ.

According to Bayes Theorem, the posterior density function for the parameter Θ_1 is:

$$p_{\Theta 1}(\theta_1|x_0) = \frac{p_{\Theta_1}(\theta_1)f_n(x_0|\theta_1)}{\int_{P_1} p_{\Theta_1}(\theta_1)f_n(x_0|\theta_1)d\theta_1}, \tag{6}$$

where $f_n(x_0|\theta)$ is defined in (1). Also from (2) and (3), the (joint) posterior density function for $\Theta = (\Theta_1, \Theta_2)$ is:

$$p_\Theta(\theta|x_0) = p_{\Theta_2}(\theta_2)p_{\Theta_1}(\theta_1|x_0), \tag{7}$$

for any $\theta = (\theta_1, \theta_2) \in P$, where $p_{\Theta_1}(\theta_1|x_0)$ is defined in (6). The next proposition is required in section 3 and shows how $r(x_0)$ explicitly depends upon the posterior density function $p_\Theta(\theta|x_0)$.

Proposition 1. If $p_\Theta(\theta|x_0)$ is the posterior distribution defined in (7), then:

$$r(x_0) = \int_P r_1(\theta) \, p_\Theta(\theta|x_0) \, d\theta \, ,$$

where $r_1(\theta) = E[W|\Theta = \theta]$ and $r(x_0)$ is defined in (4).

Proof. It follows from (5) that:

$$E[W|X = x_0, \Theta = \theta] = E[W|\Theta = \theta],$$

for $\theta \in P$, so that:

$$\begin{aligned} r(x_0) &= E[W|X = x_0] = E[E[W|X = x_0, \Theta]|X = x_0] \\ &= E[E[W|\Theta]|X = x_0] = E[r_1(\Theta)|X = x_0] \\ &= \int_P r_1(\theta) \, p_\Theta(\theta|x_0) \, d\theta. \end{aligned}$$

The variability of the estimators that we will propose in section 3 rely both on the variance of $r_1(\Theta)$ as well as the variance of W with given Θ. For this reason, we introduce the following notation:

$$\sigma_p^2 = E[r_1^2(\Theta)] - (E[r_1(\Theta)])^2, \tag{8}$$

and

$$\sigma_s^2 = E[\sigma^2(\Theta)], \tag{9}$$

where:

$$\sigma^2(\theta_0) = E[W^2|\Theta = \theta_0] - (E[W|\Theta = \theta_0])^2,$$

for $\theta_0 \in P$, and $r_1(\theta) = E[W|\Theta = \theta]$.

Example 1. In order to illustrate our notation we present an inventory example that attracted our interest in the estimation of (4). Suppose that a retailer makes periodical orders (every N days) to satisfy its sales on the Internet (if there is no available inventory, the sale is lost). At the beginning of each period, the company has an initial inventory (I_0) and makes an order of size Q. Let U_i be the lead time for an order set at the beginning of period i. The random components in this model will be the time between customer arrivals and the number of items a customer orders. We assume that the time between customer arrivals are i.i.d. according to an exponential distribution:

$$g(x|\theta_2) = \begin{cases} (1/\theta_2)e^{-x/\theta_2}, & x > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta_2 \in P_2 = (0, \infty)$, in addition, every customer orders j items (independently of each other) with probability θ_{1j} ($j = 1, \ldots, k$). In this case we have $\theta_1 = (\theta_{11}, \ldots, \theta_{1k})$, $f(x|\theta_1) = \theta_{1x}, x = 1, 2, \ldots, k$, and

$$P_1 = \left\{ (\theta_{11}, \ldots, \theta_{1k}) : \sum_{i=1}^{k} \theta_{1i} = 1; \theta_{1i} \geq 0; i = 1, 2, \ldots, k \right\}.$$

In a particular application Θ_2 can be used to model a future scenario, for which sample data may not be available.

However, sample data for parameter Θ_1 may be available from past sales (assuming the same pattern), the sample data is $x_0 = (x_1, \ldots, x_n)$, where x_i is the number of items ordered by customer i, then we can compute $u = (u_1, \ldots, u_k)$, where u_j is the number of customers that ordered j items. Note that given $[\Theta = \theta_1]$, u follows a multinomial distribution, and as is well known [12], the conjugate family of a multinomial distribution is a Dirichlet distribution, which means that if $p_{\Theta_1}(\theta_1)$ corresponds to a Dirichlet(k, α_0) distribution, where $\alpha_0 = (\alpha_1, \ldots, \alpha_k)$, then $p_{\Theta_1}(\theta_1|x_0)$ corresponds to a Dirichlet(k, α_F) distribution, where:

$$\alpha_F = (\alpha_1 + u_1, \ldots, \alpha_k + u_k).$$

Note also that $\alpha_0 = (1, \ldots, 1)$ corresponds to a uniform distribution ($p_{\Theta_1}(\theta_1)$ is constant on P_1).

Performance measures of interest in this model are the type-I service level and the type-II service level [15]. For a type-I service level we have that

$$r_1(\theta) = P[D \leq I_1 + Q|\Theta = \theta],$$

where D is the demand from clients that arrived between $t = U_1$ and $t = N + U_2$, $I_1 = \max\{0, I_0 - D_1\}$, and D_1 is the demand from clients that arrived between $t = 0$ and $t = U_1$. It is also worth mentioning that in this case:

$$W = \begin{cases} 1, & \text{if } D \leq I_1 + Q, \\ 0, & \text{otherwise.} \end{cases}$$

Analytical expressions for $E[W|X = x_0]$ in both cases (type-I or type-II service levels), are not easy to obtain even for unrealistic assumptions (e.g., $k = 1$).

Example 2. The next example is a Markovian model that was used in [16] to forecast the demand for educational support material in an adult education program, and has an analytical solution that can be used to test the methodologies discussed in section 3. Let $U = \{U_i : i = 0, 1, \ldots\}$ be a Markov chain with finite state space $E = \{0, 1, 2, \ldots, k\}$, and suppose that there are q potential clients at time t, and the state of every client is known. State transitions for every client will occur independently and according to the Markov chain U. When potential client l is in state X_{t+1}^l at time $t + 1$, his demand for certain item is a function $f(X_{t+1}^l)$ of its state, and although in a general case it suffices that f be non-negative, we assume that:

$$f(i) = \begin{cases} 1, & \text{if } i \in A, \\ 0, & \text{otherwise,} \end{cases}$$

where $A \subseteq E$ is the set of states where an individual demands one unit of the item. We are interested in the total demand at time $t + 1$:

$$W = \sum_{l=1}^{n} f\left(X_{t+1}^{l}\right). \tag{10}$$

Since the state of each individual at time t is known, W can be simulated from the knowledge of:

$$p_i = P\left[f\left(X_{t+1}^{l}\right) = 1 \mid X_t^{l} = i\right] = \sum_{j \in A} P\left[X_{t+1}^{l} = j \mid X_t^{l} = i\right], \tag{11}$$

$i \in E$, $t = 1, 2, \ldots$, $l = 1, 2, \ldots, n$.

This model can be incorporated into our framework if we consider Y in discrete time with $T = 1$, since we wish to forecast the response for the next period ($t+1$) based on information for the initial period (t). In order to simulate W we let $\theta_1 = (p_1, \ldots, p_k)$, $\theta = (\theta_1, \theta_2)$,

and $Y(i) = \left(U_{t+i}^{1}, \ldots U_{t+i}^{n} X_{t+i}^{1}, \ldots, X_{t+i}^{n}\right)$, for $i = 0, 1$, since the knowledge of $Y(0)$ and $\theta = (\theta_1, \theta_2)$ is sufficient to generate W using simulation.

Information on θ comes from recording the behavior of potential clients in past periods, and can be summarized in a data vector $x_0 = (n_1, r_1, n_2, r_2, \ldots, n_k, r_k)$ that indicates that $n_i + r_i$ observed clients were in state i, from which n_i ordered one unit of the item (and r_i did not order). State 0 corresponds to new clients, for which there is no sample data. The likelihood function becomes:

$$f_n\left(x_0 \mid \theta\right) = \prod_{i=1}^{k} p_i^{n_i}\left(1 - p_i\right)^{r_i},$$

where θ y x_0 are as before. When $p_{\Theta}(\theta)$ is uniform on $[0,1]$, we can verify from (6) and (7) that:

$$p\left(\theta \mid x_0\right) = \frac{\prod_{i=1}^{k} p_i^{n_i}\left(1 - p_i\right)^{r_i}}{\prod_{i=1}^{k} B\left(1 + n_i, 1 + r_i\right)}, \tag{12}$$

for $0 < p_i < 1$, $i = 0, 1, \ldots, k$, where:

$$B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta) / \Gamma(\alpha + \beta),$$

for $\alpha, \beta > 0$. In other words, the posterior distribution for parameters $p_0, p_1, p_2, \ldots, p_k$ corresponds to $k+1$ independent random variables, where p_i is distributed as $Beta(1 + n_i, 1 + r_i)$ (where $n_0 = r_0 = 0$). It follows from (10)-(11) that $r_1(\theta) = \sum_{i=0}^{k} n_{it} p_i$, where n_{it} is the number of potential clients that are in state i at time t, and $r_1(\theta)$ is as in Proposition 1. Therefore using (12) and Proposition 1 we can verify that the parameter proposed in (4) is:

$$r\left(x_0\right) = \sum_{i=0}^{k} n_{it} \hat{p}_i, \tag{13}$$

where $\hat{p}_i = \left(1 + n_i\right) / \left(2 + n_i + r_i\right)$.

3. METHODOLOGY

In this section, we present estimators for the performance measure (4), and for each of these a Law of Large Numbers (LLN) and a Central Limit Theorem (CLT) [17] is obtained. A LLN for an estimator will ensure that the estimator is consistent, which means that it approaches the parameter as we increase the number of replications of the simulation experiment. A CLT allows us to establish an asymptotic confidence interval to assess the precision of the estimation procedure.

In section 3.1 and section 3.2 we present two different methodologies for the estimation of $r(x_0)$, respectively. The methodology presented in section 3.1 is a particular case of the one proposed in [3] adapted to our notation (in the mentioned article, inherent uncertainty in the model is also considered). Although the methodology described in section 3.1 produces a consistent estimator for $r(x_0)$, it requires a valid method to generate samples from the posterior distribution $p_{\Theta}(\theta \mid x_0)$, which is usually too difficult to obtain. In view of this limitation, in section 3.2 we propose a methodology for the estimation of $r(x_0)$ that does not require to sample from $p_{\Theta}(\theta \mid x_0)$.

3.1. Estimation of $r(x_0)$ by Sampling from the Posterior Distribution

In Fig. (**1**) we describe a first procedure for the estimation of $r(x_0)$. This methodology is similar to the classical method of replications, with the sole difference that on replication j a value of θ_j for the unknown parameter is sampled from the posterior distribution $p_{\Theta}(\theta \mid x_0)$ defined in (7).

1. For $j = 1$ to the number of replications m:
   a. Generate (independently) a value θ^{j} by sampling from $p_{\Theta}(\theta \mid x_0)$.
   b. Run (independently) a simulation experiment with $\Theta = \theta^{j}$ to obtain:
   $$W_j = g\left(Y(s), 0 \leq s \leq T\right)$$
   End Loop
2. Compute:
$$\hat{r}_m\left(x_0\right) = \frac{1}{m}\sum_{j=1}^{m} W_j, \quad S_m^{2} = \frac{1}{m-1}\sum_{j=1}^{m}\left[W_j - \hat{r}_m\left(x_0\right)\right]^{2}$$

Fig. (1). Estimation of $r(x_0)$ using the posterior distribution.

To apply the methodology described in Fig. (**1**), it is necessary to apply a valid method for the generation of random variables from the posterior density $p_{\Theta}(\theta \mid x_0)$; which can be available, for example, when the corresponding distribution has been identified from an analytical expression.

Note that the variables W_j in the algorithm of Fig. (**1**) are i.i.d. with $E[W_1] = r(x_o)$, and variance $\sigma_W^2 = \sigma_p^2 + \sigma_s^2$, so that the next proposition follows from the classical LLN and CLT ("$\rightarrow$" denotes convergence in probability and "$\Rightarrow$" denotes weak convergence).

Proposition 2. If $0 < \sigma_p^2 + \sigma_s^2 < \infty$, then:

$$\hat{r}_m(x_0) \rightarrow r(x_0),$$

as $m \rightarrow \infty$, and:

$$\frac{\sqrt{m}\left[\hat{r}_m(x_0) - r(x_0)\right]}{S_m} \Rightarrow N(0,1),$$

as $m \rightarrow \infty$, where σ_p^2 and σ_s^2 are defined in (8) and (9) respectively, and $\hat{r}_m(x_0)$, S_m^2 are defined in Fig. (**1**).

Corollary 1. An asymptotic $(1-\alpha)100\%$ confidence interval for $r(x_0)$ is given by:

$$\left[\hat{r}_m(x_0) - z_\alpha \frac{S_m}{\sqrt{m}}, \hat{r}_{21}^m(x_0) + z_\alpha \frac{S_m}{\sqrt{m}}\right].$$

3.2. Estimation of $r(x_0)$ by Using Ratio Estimation

In many cases an analytical expression for $p(\theta|x_0)$ cannot be found or is very difficult to obtain, and sampling from $p_\Theta(\theta|x_0)$ is not feasible because the corresponding family of distributions has not been identified. The methodology we will propose in this section can be used in this case, and is explained as follows.

From (6) and Proposition 1, $r(x_0)$ can be viewed as a ratio of the form:

$$r(x_0) = \int_P r_1(\theta) p_\Theta(\theta|x_0) d\theta = \frac{\int_P r_1(\theta) p_\Theta(\theta) f_n(x_0|\theta) d\theta}{\int_P p_\Theta(\theta) f_n(x_0|\theta) d\theta},$$

which allows us to apply the ratio estimation procedure used in the regenerative method of simulation [18]. In addition, we may try to reduce the variance of the point estimator by using importance sampling [19], i.e., we can sample from another density $q(\theta)$ that is positive whenever $p_\Theta(\theta)$ is positive, so that:

$$r(x_0) = \frac{\int_P r_1(\theta) p_\Theta(\theta) q(\theta)^{-1} f_n(x_0|\theta) q(\theta) d\theta}{\int_P p_\Theta(\theta) q(\theta)^{-1} f_n(x_0|\theta) q(\theta) d\theta}$$

$$= \frac{E_q\left[r_1(\theta) L(\theta)\right]}{E_q\left[L(\theta)\right]}, \tag{14}$$

where $L(\theta) = p_\Theta(\theta) q(\theta)^{-1} f_n(x_0|\theta)$.

The resulting point and variability estimation procedure is described in Fig. (**2**). Note that when importance sampling is not applied, then we sample from $p_\Theta(\theta)$ and $L(\theta) = f_n(x_0|\theta)$ in Fig. (**2**).

1. For $j = 1$ to the number of replications m:
   a. Generate (independently) a value θ^j by sampling from $q(\theta)$.
   b. Run (independently) a simulation experiment with $\Theta = \theta^j$ to obtain:
   $$W_j = g(Y(s), 0 \le s \le T)$$
   c. Compute:
   $$Z_j = p_\Theta(\theta^j) q(\theta^j)^{-1} f_n(x_0|\theta^j), \quad Y_j = W_j Z_j$$
   End Loop
2. Compute:
   $$\overline{Y}_m = \frac{1}{m} \sum_{j=1}^m Y_j, \quad \overline{Z}_m = \frac{1}{m} \sum_{j=1}^m Z_j$$
3. Compute:
   $$\hat{r}_m^r(x_0) = \frac{\overline{Y}_m}{\overline{Z}_m}, \quad S_{YZ} = \frac{1}{m-1} \sum_{j=1}^m (Y_j - \overline{Y}_m)(Z_j - \overline{Z}_m)$$
   $$S_Y^2 = \frac{1}{m-1} \sum_{j=1}^m (Y_j - \overline{Y}_m)^2, \quad S_Z^2 = \frac{1}{m-1} \sum_{j=1}^m (Z_j - \overline{Z}_m)^2$$
   $$S_{rm}^2 = S_Y^2 - 2\hat{r}_m^r(x_0) S_{YZ} + \hat{r}_m^r(x_0)^2 S_Z^2$$

Fig. (2). Estimation of $r(x_0)$ using Ratio Estimation.

In order to formulate a LLN and a CLT for the estimator proposed in Fig. (**2**), we introduce the following notation:

$$\sigma_r^2 = \sigma_{Y_1}^2 - 2r(x_0) \sigma_{Y_1 Z_1} + r(x_0)^2 \sigma_{Z_1}^2, \tag{15}$$

where $\sigma_{Y_1}^2 = E\left[(Y_1 - E[Y_1])^2\right]$, $\sigma_{Z_1}^2 = E\left[(Z_1 - E[Z_1])^2\right]$,

$\sigma_{Y_1 Z_1} = E\left[(Y_1 - E[Y_1])(Z_1 - E[Z_1])\right]$, $r(x_0)$ is defined in (4), and Y_1, Z_1 are defined in Fig. (**2**).

The following Proposition follows directly from (1.9) of Shedler (1993), we only need to consider $Y_k(f) = Y_k$ and $\tau_k = Z_k$ in Theorem 1.4 of [14].

Proposition 3. If $0 < \sigma_r^2 < \infty$, then:

$$\hat{r}_m^r(x_0) \rightarrow r(x_0),$$

as $m \rightarrow \infty$, and:

$$\frac{\overline{Z}_m \sqrt{m}\left[\hat{r}_m^r(x_0) - r(x_0)\right]}{S_{rm}} \Rightarrow N(0,1),$$

as $m \rightarrow \infty$, where σ_{rm}^2 and $r(x_0)$ are defined in (11) and (4) respectively, and $\hat{r}_m^r(x_0)$, $\overline{Z}_m$, S_{rm}^2 are defined in Fig. (**2**).

Corollary 2. An asymptotic $(1-\alpha)100\%$ confidence interval for $r(x_0)$ is given by:

$$\left[\hat{r}_{22}^{m}(x_0) - \frac{z_\alpha S_{rm}}{\overline{Z}_m \sqrt{m}}, \hat{r}_{22}^{m}(x_0) + \frac{z_\alpha S_{rm}}{\overline{Z}_m \sqrt{m}} \right].$$

4. EXPERIMENTAL RESULTS

In this section we present the results obtained from simulation experiments that we performed in order to verify the theoretical results presented in the previous section, as well as to test the performance of the proposed methodologies. We considered the Markovian model presented in Example 2 because the analytical expression obtained for the performance measurement (4) will allow us to verify the validity of our theoretical results.

In all the experiments reported in this section we produced 95% confidence intervals, and for $M = 1000$ independent replications per confidence interval estimation, we computed different performance measurements for the set of experiments (see Fig. **3**). We considered the model of Example 2 with $k = 5$, $n_{it} = 10$, $n_i = i$, $r_i = 10 - i$, $i = 1, \ldots, 5$. Under these assumptions we can verify from (13) that $r(x_0) \approx 21.667$. This value was considered as the true parameter value in order to compute the empirical coverage, mean square error and bias for each set of experiments.

In Table **1** we present the results of three sets of experiments with $m =$ 100, 500 and 1000 replications per experiment, respectively. For each set of experiments we considered $M = 1000$ independent replications of the corresponding estimation procedure. The results reported under "Posterior Sampling" correspond to the method of Fig. (**1**), and the results reported under "Ratio using $q(\theta)$" correspond to the method of Fig. (**2**), where using the same notation as in Example 2, $q(\theta) = p_{\Theta_2}(\theta_2) q_1(\theta_1 | x_0)$, $p_{\Theta_2}(\theta_2)$ is uniform on $[0,1]$, and $q_1(\theta_1 | x_0)$ corresponds to the limiting distribution (as $n \to \infty$) for $p_{\Theta_1}(\theta_1 | x_0)$, i.e., a $N(m, V^{-1})$ distribution (negative values or greater than 1 were rejected), where m is the maximum likelihood estimator for Θ_1, and V is the influence matrix (see [12]). It can be verified that the i-th element of m is $m_i = n_i / (n_i + r_i)$, and V is a diagonal matrix with $V_i = (n_i + r_i) / m_i (1 - m_i)$ as its i-th diagonal element. The results reported under "Ratio using Prior" correspond to the method of Fig. (**2**) when no importance sampling is applied (we sample from $p_\Theta(\theta)$ and $L(\theta) = f_n(x_0 | \theta)$).

As we can see from Table **1**, sampling from the posterior distribution provided very good coverages (close to the nominal 0.95) even for a relatively small number of replications. This method exhibited also the best performance from the point of view of halfwidth, mean square error and bias. When sampling from the prior distribution, the ratio estimation method had the worst performance from the point of view of coverage, halfwidth, mean square error and bias, and although the coverage was poor, it gets close to the nominal 0.95 as the number of replication increases, which shows

that the method is valid but needs a larger number of replications to provide reliable confidence intervals.

Halfwidth

For $i = 1, 2, \ldots, M$, let h_i be the halfwidth obtained in experiment i:

$$\overline{h} = \frac{1}{M} \sum_{i=1}^{M} h_i, \qquad S_h = \sqrt{\frac{\sum_{i=1}^{M} \left(h_i - \overline{h} \right)^2}{M - 1}}$$

Empirical Mean Square Error

For $i = 1, 2, \ldots, M$, let r_i be the true (analytical) value for a performance measure and let $\hat{r}_i$ be the point estimator obtained in experiment i:

$$MSE = \frac{\sum_{i=1}^{M} \left(\hat{r}_i - r_i \right)^2}{M}$$

Empirical Coverage

$$EC = \frac{1}{M} \sum_{i=1}^{M} C_i,$$

where: $\quad C_i = \begin{cases} 1, & \text{if } |\hat{r}_i - r_i| < h_i, \\ 0, & \text{otherwise.} \end{cases}$

Fig. (3). Performance measures for a set of experiments.

When applying importance sampling, the ratio estimation method performed better, providing a reasonable coverage even for a small number of replications. Although it performed worse than posterior sampling, its performance was very close, with the advantage that this method can be applied in a very general setting without sampling from the posterior distribution.

Table 1.　Performance of 95% Confidence Intervals from 1000 Replications

Method	EC	Halfwidth		MSE	Bias
		$\overline{h}$	S_h		
m = 100					
Posterior Sampling	.934	1.026	.065	0.296	.028
Ratio using Prior	.641	3.297	1.849	9.867	.614
Ratio using $q(\theta)$	.915	1.389	.546	0.818	.001
m = 500					
Posterior Sampling	.950	.460	.014	0.057	.006
Ratio using Prior	.846	2.608	.929	2.861	.045
Ratio using $q(\theta)$	.937	.777	.464	0.351	.091
m = 1000					
Posterior Sampling	.950	.325	.007	0.027	-.008
Ratio using Prior	.898	2.092	.594	1.429	.102
Ratio using $q(\theta)$	.942	.616	.392	0.215	.104

5. CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK

Our main objective in the development of the theoretical framework presented in section 2 has been to propose a performance measure that is able to incorporate both priors and sample data in the analysis of simulation experiments. This performance measure allows the incorporation of priors and/or sample data only as special cases. We discussed two estimation methodologies (posterior sampling and ratio estimation) that allows to assess the precision of the point estimations (through asymptotic confidence intervals). On the other hand, it is important to note that both methodologies are easy to implement since the main difference with the classical method of replications is the generation of an additional sample for the parameter in each replication.

Our experimental results show that, when using importance sampling, the ratio estimation method may perform well. It should be mentioned that we applied a very general methodology (using the limiting distribution) when applying importance sampling, and the results were close to the posterior sampling method. Although further research is required, it is expected that a good choice of a sampling distribution may provide an even better performance (variance reduction) for the ratio estimation method.

A direction for future research could be the incorporation of the proposed methodologies in the development of decision support systems by considering forecasts provided by the decision makers. However, the usefulness of the decision making support systems is very closely related to its capacity to model the real system, which depends on an adequate selection of the parameters that model the different scenarios and the probability models for the random components that incorporate the stochastic uncertainty. In our opinion, it is precisely in this area, the search of models and their particular applications, in which the greatest research potential of these methodologies can be found.

ACKNOWLEDGMENTS

REFERENCES

[1]   Kraan B, Bedford T. Probability inversion of expert judgment in the quantification of model uncertainty. Manag Sci 2005; 51(6): 995-1006.

[2]   Fisher ML, Raman A, McClelland AS. Rocket science retailing is almost here. Harv Bus Rev 2000; 78(4): 115-24.

[3]   Chick SE. Input distribution selection for simulation experiments: accounting for input uncertainty. Oper Res 2001; 49(5): 744-58.

[4]   Chick SE. Bayesian methods for discrete event simulation, Proceedings of the 2004 Winter Simulation Conference. Ingalls RG, Rossetti MD, Smith JS, Peters BA, Eds. 2004; 89-100.

[5]   Zouaoui F, Wilson JR. Accounting for parameter uncertainty in simulation input modeling. IIE Trans 2003; 35(9): 781-792.

[6]   Law AM, Kelton WD. Simulation modeling and analysis, 3rd ed. New York: McGraw-Hill 2000.

[7]   Cheng RCH, Holland W. Two-point methods for assessing variability in simulation output. J Stat Comp Simul 1998; 60: 183-205.

[8]   Cheng RCH, Holland W. Calculation of confidence intervals for simulation output. ACM Trans Model Comput Simul 2004; 14: 344-62.

[9]   Lee SH, Glynn PW. Computing the distribution function of a conditional expectation *via* monte carlo: discrete conditioning spaces. In: Farrington PA, Nembhard HB, Sturrock DT, Evans GW, Eds. Proceedings of the 1999 Winter Simulation Conference, 1999; 1654-63.

[10]   Muñoz DF. A Bayesian framework for modeling demand in supply chain simulation experiments. In: Chick LS, Sanchez PJ, Ferrin D, Morrice DJ, Eds. Proceedings of the 2003 Winter Simulation Conference, 2003; 1319-25.

[11]   Berger JO. Statistical decision theory and Bayesian analysis, 2nd ed. New York: Springer-Verlag 1985.

[12]   Bernardo JM, Smith AFM. Bayesian theory. Chichester: John Wiley & Sons 2000.

[13]   Asmussen S. Applied probability and queues. New York: Springer-Verlag 2003.

[14]   Shedler GS. Regenerative stochastic simulation. San Diego: Academic Press 1993.

[15]   Anupindi R, Chopra S, Deshmukh SD, Van Mieghem JA, Zemel E. Managing business process flows, 2nd ed. New Jersey: Pearson-Prentice Hall 2006.

[16]   Detta JE. Estudio de un caso comparativo de pronósticos basados en regresión con pronósticos basados en simulación. Eng. diss., Department of Industrial & Operations Engineering, Instituto Tecnológico Autónomo de México 2006.

[17]   Chung KL. A course in probability theory, 2nd ed. San Diego: Academic Press 1974.

[18]   Iglehart DL. The regenerative method for simulation analysis. In: Chandy KM, Yeh RT, Eds. Current trends in programming methodology vol. III software engineering, Prentice Hall 1978; pp. 52-71.

[19]   Glynn PW, Iglehart DL. Importance sampling for stochastic simulations. Manag Sci 1989; 35(11): 1367-92.