Proceedings of the 2004 Winter Simulation Conference
R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

SENSITIVITY ANALYSIS FOR TRANSIENT SINGLE SERVER QUEUING MODELS USING AN INTERPOLATION APPROACH

Mohamed A. Ahmed
Talal M. Alkhamis

Department of Statistics and Operations Research
Kuwait University
P. O. Box 5969, Safat, KUWAIT

ABSTRACT

Simulation is an essential tool for performance evaluation of many practical systems where planners typically want to know how the system will perform under various parameter settings. Since large-scale simulation may require great amount of computer time and storage, appropriate statistical analysis can become quite costly. In this paper, we develop an interpolation technique as an effective tool for estimating system respones to parametric perturbations in simulation. We also analyze the usefulness of the continuous–time Markov chains frame-work to find the likelihood ratio (Radon- Nikodym derivative) for Markovian single server queueing models. We provide numerical experiments that demonstrate how the interpolation technique significantly outperform the likelihood ratio performance extrapolation technique in the context of the Markovian queueing models in transient analysis.

1 INTRODUCTION

In designing, analyzing and operating real-life complex systems, we are interested, however, not only in performance evaluation but in sensitivity analysis and optimization as well. Planners typically want to know how the system will perform under various parameter settings. To determine this, a computer simulation model may be developed and then run for these parameter settings. One problem with Monte Carlo analysis is its expensive use of computer time. To address this problem, we propose an efficient technique for estimating the expected performance of a stochastic system for various values of the parameters from a single simulation of the nominal system. Given the performance measure at two values of the input distribution parameters, the proposed technique provides the ability to interpolate the simulation results at different values of these parameters. This technique is based on the likelihood ratio performance extrapolation (LRPE). Arsham et al. (1989) showed that using the likelihood ratio (Radon-Nikodym derivative) approach, one can estimate simultaneously the performance measure at various parameter values from a single simulation run.

Consider a stochastic simulation system parameterized by a real vector $\boldsymbol{\theta} \in \Theta$ of continuous parameters, where Θ is some open subset of R^n. We are interested in performance measures that are based on the behavior of the stochastic system in some time interval *T*, where *T* is a stopping time. Suppose we have independent simulation results of the system at parameter $\boldsymbol{\theta}_1 \in \Theta$ and want to estimate the transient performance measure of that system at parameter $\boldsymbol{\theta}_0 \in \Theta$, $\ell(\boldsymbol{\theta}_0 \mid \mathrm{I})$, where I represents the initial conditions used to start the simulation at time 0.

The basic idea of LRPE is that $\ell(\boldsymbol{\theta}_0 \mid \mathrm{I})$ can usually be viewed as the expectation of some function of $\boldsymbol{\theta}_0$ and the sample path ω, say $h(\boldsymbol{\theta}_0, \omega)$, with respect to a probability measure $P_{\boldsymbol{\theta}_0}$. Suppose that $P_{\boldsymbol{\theta}_0}$ is absolutely continuous with respect to $P_{\boldsymbol{\theta}_1}$,i.e., for every measurable set *B*, if $P_{\boldsymbol{\theta}_1}(B) = 0$ then $P_{\boldsymbol{\theta}_0}(B) = 0$. In this case, one can write

$$\begin{aligned}\ell(\boldsymbol{\theta}_0 \mid \mathrm{I}) &= \mathrm{E}_{\boldsymbol{\theta}_0}\left[h(\boldsymbol{\theta}_0, \omega)\right] \\ &= \int h(\boldsymbol{\theta}_0, \omega)\, dP_{\boldsymbol{\theta}_0}(\omega) \\ &= \int [h(\boldsymbol{\theta}_0, \omega) \frac{dP_{\boldsymbol{\theta}_0}}{dP_{\boldsymbol{\theta}_1}}(\omega)]\, dP_{\boldsymbol{\theta}_1}(\omega) \\ &= \mathrm{E}_{\boldsymbol{\theta}_1}\left[h(\boldsymbol{\theta}_0, \omega) L(T, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1)\right],\end{aligned}$$

where $L(T, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) = (dP_{\boldsymbol{\theta}_0} / dP_{\boldsymbol{\theta}_1})(\omega)$ is the Radon-Nikodym derivative of $P_{\boldsymbol{\theta}_0}$ with respect to $P_{\boldsymbol{\theta}_1}$ or the likelihood ratio of the process up to stopping time T. The subscript $\boldsymbol{\theta}_0$ in $\mathrm{E}_{\boldsymbol{\theta}_0}\left[h(\boldsymbol{\theta}_0, \omega)\right]$ means that the expectation operator is induced by $P_{\boldsymbol{\theta}_0}$.

Typically ω could be the set of values taken by a finite sequence of independent (possibly multivariate) random variables $\boldsymbol{Y}$ with probability density function $f(\boldsymbol{y},\boldsymbol{\theta})$. For example, consider an M/M/1 queue and let $\ell(\boldsymbol{\theta}_0 \mid \mathrm{I})$ be the expected mean waiting time in the system for the first T customers in the system, provided that the initial conditions used to start the simulation at time 0 is I. In this case, ω could be the set of actual interarrival and service times and $dP_{\boldsymbol{\theta}_0}(\omega)$ is the product of their densities. We have

$$\begin{aligned}\ell(\boldsymbol{\theta}_0 \mid \mathrm{I}) &= \mathrm{E}_{\boldsymbol{\theta}_0}[h(\boldsymbol{Y},\boldsymbol{\theta}_0)] \qquad (1)\\ &= \int h(\boldsymbol{y},\boldsymbol{\theta}_0) f(\boldsymbol{y},\boldsymbol{\theta}_0)\,d\boldsymbol{y}\\ &= \int h(\boldsymbol{y},\boldsymbol{\theta}_0)\frac{f(\boldsymbol{y},\boldsymbol{\theta}_0)}{f(\boldsymbol{y},\boldsymbol{\theta}_1)} f(\boldsymbol{y},\boldsymbol{\theta}_1)\,d\boldsymbol{y}\\ &= \mathrm{E}_{\boldsymbol{\theta}_1}[h(\boldsymbol{Y},\boldsymbol{\theta}_0)L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)]\ , \qquad (2)\end{aligned}$$

where

$$L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)=\frac{f(\boldsymbol{y},\boldsymbol{\theta}_0)}{f(\boldsymbol{y},\boldsymbol{\theta}_1)} \text{ and } f(\boldsymbol{y},\boldsymbol{\theta}_0)=\prod_{j=1}^{T} f_j(\boldsymbol{y}_j,\boldsymbol{\theta}_0).$$

It is important to note that the original expectation of $h(\boldsymbol{Y})$ in (1) is taken with respect to the underlying pdf $f(\boldsymbol{y},\boldsymbol{\theta}_0)$, whereas that given in (2) is taken with respect to the pdf $f(\boldsymbol{y},\boldsymbol{\theta}_1)$. It follows that changing the probability density from $f(\boldsymbol{y},\boldsymbol{\theta}_0)$ to $f(\boldsymbol{y},\boldsymbol{\theta}_1)$, we can express the performance measure $\ell(\boldsymbol{\theta} \mid I)$ for all $\boldsymbol{\theta} \in \Theta$ as an expectation with respect to $f(\boldsymbol{y},\boldsymbol{\theta}_1)$ and then estimate it accordingly. In terms of simulation, this means that in principle, one simulation at $\boldsymbol{\theta}_1$ can produce estimates of the performance measure at all "valid" values of $\boldsymbol{\theta}$.

Estimating $\ell(\boldsymbol{\theta}_0 \mid I)$ using the Radon-Nikodym approach yields computational savings, but reduces precision. By generating a sample $\boldsymbol{Y_1},\boldsymbol{Y_2},...,\boldsymbol{Y_n}$ from $f(\boldsymbol{y},\boldsymbol{\theta}_1)$,we can estimate $\ell(\boldsymbol{\theta}_0 \mid I)$ by the corresponding sample mean $\tilde{\ell}(\boldsymbol{\theta}_0 \mid I) = \frac{1}{n}\sum_{i=1}^{n}[h(\boldsymbol{y}_i)\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)]$. The accuracy of the estimator $\tilde{\ell}(\boldsymbol{\theta}_0 \mid I)$ is determined by its variance, $\mathrm{Var}\,\tilde{\ell}(\boldsymbol{\theta}_0 \mid I) = \frac{1}{n}\mathrm{Var}_{\boldsymbol{\theta}_1}[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)]$. Note that the farther $\boldsymbol{\theta}_1$ is from $\boldsymbol{\theta}_0$, the higher variance of the estimator $\tilde{\ell}(\boldsymbol{\theta}_0 \mid I)$, i. e., the variance of the LRPE estimators grows quite fast as the length of the perturbed parameter increases. For simplicity, in the sequel of the paper we suppress the initial condition I from the transient performance measures and their estimates.

Implementation of the LRPE approach requires computation of the likelihood ratio (Radon-Nikodym derivative) of the underlying stochastic system. In this paper we use the continuous-time Markov chain frame-work to find the likelihood ratio for a class of queuing model (see Nakayama et al. (1994)). Continuous-time Markov chains are good models for many stochastic systems, including certain queuing systems, inventory systems, and reliability and maintenance systems. While the basis of the LRPE technique (Rubinstein (1986, 1989), Glynn (1986, 1987, 1990), Reiman and Weiss (1989), L'Ecuyer (1990, 1995)) has been known for some time, the technique works only for perturbations of limited size due to its high variance.

In this paper, we develop an interpolating technique as an effective tool for estimating system response to parametric perturbations in simulation. We show that the proposed technique is an effective tool for measuring parameter sensitivity in the context of the Markovian queueing models in transient analysis. There are many instances in which the transient behavior of stochastic systems is important. Since the characteristics of most real systems change over time, the stochastic processes for those systems do not have steady-state distribution. For example, in a manufacturing system the production scheduling rules and the facility layout (e.g., number and location of machines) may change from time to time.

The rest of the paper is organized as follows. Section 2 develops the proposed technique as an efficient method for estimating the transient performance measures of stochastic systems. Section 3 gives the Radon-Nikodym derivative for the M/M/1 queueing systems and provides our computational experiments that demonstrate the efficiency of the proposed technique. Finally, section 4 contains some concluding remarks.

2 THE INTERPOLATION APPROACH

In this section, we discuss the interpolation technique as an efficient method for estimating the transient performance measures. Suppose we have independent simulation results of a stochastic system at two values of the parameter $\boldsymbol{\theta} \in \Theta$, say $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, and consider the expected performance of the system at $\boldsymbol{\theta}_0$. Then, we have

$$\begin{aligned}\ell(\boldsymbol{\theta}_0) &= \mathrm{E}_{\boldsymbol{\theta}_0}[h(\boldsymbol{Y},\boldsymbol{\theta}_0)]\\ &= \int h(\boldsymbol{y},\boldsymbol{\theta}_0) f(\boldsymbol{y},\boldsymbol{\theta}_0)\,d\boldsymbol{y}\\ &= \alpha\int h(\boldsymbol{y},\boldsymbol{\theta}_0) f(\boldsymbol{y},\boldsymbol{\theta}_0)\,d\boldsymbol{y}\\ &\qquad +(1-\alpha)\int h(\boldsymbol{y},\boldsymbol{\theta}_0) f(\boldsymbol{y},\boldsymbol{\theta}_0)\,d\boldsymbol{y}\\ &= \alpha\int h(\boldsymbol{y},\boldsymbol{\theta}_0)\frac{f(\boldsymbol{y},\boldsymbol{\theta}_0)}{f(\boldsymbol{y},\boldsymbol{\theta}_1)} f(\boldsymbol{y},\boldsymbol{\theta}_1)\,d\boldsymbol{y}\\ &\qquad +(1-\alpha)\int h(\boldsymbol{y},\boldsymbol{\theta}_0)\frac{f(\boldsymbol{y},\boldsymbol{\theta}_0)}{f(\boldsymbol{y},\boldsymbol{\theta}_2)} f(\boldsymbol{y},\boldsymbol{\theta}_2)\,d\boldsymbol{y}\\ &= \alpha\,\mathrm{E}_{\boldsymbol{\theta}_1}[h(\boldsymbol{Y},\boldsymbol{\theta}_0)L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)]\\ &\qquad +(1-\alpha)\mathrm{E}_{\boldsymbol{\theta}_2}[h(\boldsymbol{Y},\boldsymbol{\theta}_0)L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_2)].\end{aligned}$$

An estimator of $\ell(\boldsymbol{\theta}_0)$ is

$$\tilde{\ell}(\boldsymbol{\theta}_0) = \frac{\alpha}{n}\sum_{i=1}^{n}\left[h(\boldsymbol{y}_i)\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)\right] + \frac{1-\alpha}{n}\sum_{i=1}^{n}\left[h(\boldsymbol{y}_i)\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_2)\right]$$

and its variance is given by

$$\mathrm{Var}\,\tilde{\ell}(\boldsymbol{\theta}_0) = \frac{\alpha^2}{n}\mathrm{Var}_{\boldsymbol{\theta}_1}\left[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)\right] + \frac{(1-\alpha)^2}{n}\mathrm{Var}_{\boldsymbol{\theta}_2}\left[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_2)\right].$$

It can be shown that the value of α that minimizes $\mathrm{Var}\,\tilde{\ell}(\boldsymbol{\theta}_0)$ is given by

$$\alpha^* = \frac{\mathrm{Var}_{\boldsymbol{\theta}_2}\left[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_2)\right]}{\mathrm{Var}_{\boldsymbol{\theta}_1}\left[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)\right] + \mathrm{Var}_{\boldsymbol{\theta}_2}\left[h(\boldsymbol{Y})\,L(T,\boldsymbol{\theta}_0,\boldsymbol{\theta}_2)\right]}.$$

2.1 Illustrative Example (Poisson Rates)

Consider a system with a Poisson arrival process that runs for a time T. Assume that the Poisson process can be constructed to be independent of the rest of the system. Given the simulation results at arrival rate θ_1, $\ell(\theta_1)$, we want to estimate the performance measure for arrival rate θ_0, say, $\ell(\theta_0)$. That is, we choose to simulate at arrival rate θ_1 and use LRPE to estimate the performance measure for arrival rate θ_0 as follows:

$$\ell(\theta_0) = \mathrm{E}_{\theta_1}\left[h(\boldsymbol{Y})\;L(T,\theta_0,\theta_1)\right],$$

where $L(T,\theta_0,\theta_1) = (\frac{\theta_0}{\theta_1})^{\mathrm{N(T)}}\exp-[(\theta_0-\theta_1)T]$.

Here T is the duration of the simulation and N(T) is the number of Poisson events up to time T. Consider for simplicity the case where $h(\boldsymbol{Y}) = 1$, then a point estimate for $\ell(\theta_0)$ is given by $\tilde{\ell}(\theta_0) = \frac{1}{n}\sum_{\mathrm{i}=1}^{\mathrm{n}} L_i(T,\theta_0,\theta_1)$ and its variance is given by $\mathrm{Var}\,\tilde{\ell}(\theta_0) = \frac{1}{n}\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right]$.

The following lemma presents the variance of the LRPE estimator, $\mathrm{Var}\,\tilde{\ell}(\theta_0)$.

Lemma 1 For the system with Poisson arrival process of example 2.1, the variance of the LRPE estimator is given by

$$\mathrm{Var}\,\tilde{\ell}(\theta_0) = \frac{\exp\left(T(\theta_0-\theta_1)^2\right)-1}{n}.$$

Proof

$$\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] = \mathrm{E}_{\theta_1}\left[L(T,\theta_0,\theta_1)^2\right] - \left[\mathrm{E}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right]\right]^2$$

and

$$\mathrm{E}_{\theta_1}\left[L(T,\theta_0,\theta_1)^2\right] = \exp(-2(\theta_0-\theta_1)T)\mathrm{E}_{\theta_1}\left[\left(\frac{\theta_0}{\theta_1}\right)^{2N(T)}\right]$$

where

$$\mathrm{E}_{\theta_1}\left[\left(\frac{\theta_0}{\theta_1}\right)^{2N(T)}\right] = \sum_{\mathrm{n}=0}^{\infty}\left(\frac{\theta_0}{\theta_1}\right)^{2\mathrm{n}}\frac{(\theta_1 T)^{\mathrm{n}}}{\mathrm{n}!}\exp(-\theta_1 T)$$

$$= \exp(-\theta_1 T)\sum_{\mathrm{n}=0}^{\infty}\frac{\left(\theta_0^2 T/\theta_1\right)^n}{n!}$$

$$= \exp(-\theta_1 T)\exp\left(\frac{\theta_0^2 T}{\theta_1}\right)$$

$$= \exp\left[-T\left(\frac{\theta_1^2-\theta_0^2}{\theta_1}\right)\right],$$

and thus,

$$\mathrm{E}_{\theta_1}\left[L(T,\theta_0,\theta_1)^2\right] = \exp(-2(\theta_0-\theta_1)T)\exp\left(-T\left(\frac{\theta_1^2-\theta_0^2}{\theta_1}\right)\right)$$

$$= \exp\left(\mathrm{T}(\theta_0-\theta_1)^2/\theta_1\right)$$

and since $\mathrm{E}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] = 1$, then

$$\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] = \exp\left(T(\theta_0-\theta_1)^2/\theta_1\right) - 1.$$

It appears that attempts to extrapolate for large values produce misleading estimates. Take for example, $T = 1$, $\theta_1 = 1$, and $\theta_0 = 4$, then $Var_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] = 8102.084$.

Using the interpolation technique, suppose we have two independent simulation runs of the system at θ_1 and θ_2. We can estimate the performance measure at θ_0 as follows:

$$\ell(\theta_0) = \alpha\mathrm{E}_{\theta_1}\left[h(Y,\theta_0)L(T,\theta_0,\theta_1)\right] + (1-\alpha)\mathrm{E}_{\theta_2}\left[h(Y,\theta_0)L(T,\theta_0,\theta_2)\right],$$

where $L(T,\theta_0,\theta_1) = \left(\frac{\theta_0}{\theta_1}\right)^{\mathrm{N}(T)}\exp\left[-(\theta_0-\theta_1)T\right]$ and

$$L(T,\theta_0,\theta_2) = \left(\frac{\theta_0}{\theta_2}\right)^{\mathrm{N}(T)}\exp\left[-(\theta_0-\theta_2)T\right].$$

Consider for simplicity the case where $h(\mathbf{Y}) = 1$, then a point estimate for $\ell(\theta_0)$ is given by

$$\tilde{\ell}(\theta_0) = \alpha^* \sum_{i=1}^{n} L_i(T,\theta_0,\theta_1) + (1-\alpha^*)\sum_{i=1}^{n} L_i(T,\theta_0,\theta_2) \text{ and}$$

its variance is given by $\mathrm{Var}\,\tilde{\ell}(\theta_0)$

$$= \frac{\alpha^{*2}}{n}\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] + \frac{(1-\alpha^*)^2}{n}\mathrm{Var}_{\theta_2}\left[L(T,\theta_0,\theta_2)\right]$$

$$= \frac{\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right]\mathrm{Var}_{\theta_2}\left[L(T,\theta_0,\theta_2)\right]}{\mathrm{Var}_{\theta_1}\left[L(T,\theta_0,\theta_1)\right] + \mathrm{Var}_{\theta_2}\left[L(T,\theta_0,\theta_2)\right]}$$

$$= \frac{\left(\exp\left(T(\theta_0-\theta_1)^2/\theta_1\right)-1\right)\left(\exp\left(T(\theta_0-\theta_2)^2/\theta_2\right)-1\right)}{\exp\left(T(\theta_0-\theta_1)^2/\theta_1\right)+\exp\left(T(\theta_0-\theta_2)^2/\theta_2\right)-2}.$$

Consider the case with $T = 1$, $\theta_1 = 1$, $\theta_0 = 4$, and $\theta_2 = 7$. Then $\mathrm{Var}\,\tilde{\ell}(\theta_0) = 2.616$, and thus very substantial decrease in the variance compared with the LRPE from θ_1 ($\mathrm{Var}_{\theta_1}\tilde{\ell}(\theta_0) = 8102.084$). Table 1 presents a comparison of LRPE and the interpolation variances for example 2.1 with $T = 1, \theta_1 = 1$, and $\theta_2 = 1.1$, for different values of θ_0. Note that, in all cases significant variance reductions were achieved using the interpolation approach.

Table 1: Comparisons of LRPE and Interpolation Approach (INT) for Example 2.1

θ_0	Variance of			Ratio of Var's	
	LR_1	LR_2	INT	INT/LR_1	INT/LR_2
1.2	0.041	0.009	0.007	0.183	0.817
1.3	0.094	0.037	0.026	0.282	0.718
1.4	0.174	0.085	0.057	0.329	0.651
1.5	0.284	0.157	0.101	0.355	0.645
1.6	0.433	0.255	0.161	0.371	0.629
1.7	0.632	0.387	0.240	0.380	0.620
1.8	0.896	0.561	0.345	0.395	0.615
1.9	1.248	0.789	0.483	0.387	0.613
2.0	1.718	1.088	0.666	0.388	0.612

3 RADON-NIKODYM DERIVATIVE FOR THE M/M/1 QUEUING SYSTEM

In this section, we present the Radon Nikodym derivative for the M/M/1 queueing system in transient analysis. Also, we present our computational experiments that show that the proposed interpolation approach is an efficient way to estimate transient measures of performance.

The M/M/1 queue can be analyzed as a birth–death process (see, e.g. Cooper (1981), Gross and Harris (1998) and Kleinrock (1975)) by selecting the birth–death coefficients as follows:

$$\lambda_n = \lambda \qquad n = 0, 1, 2, \ldots$$
$$\mu_n = \mu \qquad n = 1, 2, 3, \ldots$$

For this queuing system, let X_k represent the state of the system at the k^{th} transition, and define the following

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \ldots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \ldots \\ 0 & \mu & -(\lambda+\mu) & \lambda & \ldots \\ . & . & . & . & . \end{bmatrix}$$

$$q(X_k) = \begin{cases} \lambda & X_k = 0 \\ (\lambda+\mu) & \text{otherwise} \end{cases}$$

The matrix Q is called the infinitesimal generator (or rate matrix or intensity matrix) of the process; and its elements, q_{ij}, give the "rates" of going from state i to state j. The ith diagonal element is usually denoted by $-q_i$; q_i, gives the rate of leaving state i to any other states. The elements in each row of Q thus sum to zero. Given that the system has entered state i, the holding time in state i, is an exponential random variable with parameter $\lambda + \mu$, since it is the minimum of two exponential random variables, namely, arrival time with parameter λ and service time with parameter μ. Given that a transition occurs from state i, the probability that the transition is due to an arrival (state increases to $i+1$) is $\frac{\lambda}{\lambda+\mu}$ and the probability that is due to a service completion (state decreases to i-1) is $\frac{\mu}{\lambda+\mu}$. Thus we have a process that stays in state i for a time that is exponential random variable and jumps to either state $i+1$ or state i -1 with transition probabilities of

$$P(X_k, X_{k+1}) = \begin{cases} 1 & X_k = 0, X_{k+1} = 1 \\ \frac{\lambda}{\lambda+\mu} & X_{k+1} = X_k + 1 \\ \frac{\mu}{\lambda+\mu} & X_{k+1} = X_k - 1 \end{cases}$$

The following lemma gives the Radon-Nikodym derivative for the M/M/1 system. First, let

N_o = number of state transitions from state 0,
N_u = number of state transitions from state X_k to X_k+1,
N_d = number of state transitions from state X_k to $X_k - 1$.

Note that N_o+N_u represent the number of arrivals by time T and N_d represents the number of departures by time T.

Lemma 2 *For the M/M/1 system, let T be the stopping time. The likelihood ratio with respect to parameter λ_0 is given by*

$$L(T,\lambda,\lambda_0) = \left(\frac{\lambda}{\lambda_0}\right)^{N_o+N_u} exp\{-T(\lambda-\lambda_0)\}$$

and with respect to μ_0 is given by

$$L(T,\mu,\mu_0) = \left(\frac{\mu}{\mu_0}\right)^{N_d} \exp\{-(\mu-\mu_0) \times (\sum_{k\in N-N_0} t_k + 1_{\{X_{N(T)}\neq 0\}}[T-T_{N(T)}])\}.$$

Proof Let $N(T)$ denote the number of transitions up to time T. The likelihood of the sample path up to time T under parameter λ and μ is

$$\Gamma(T,\mu,\lambda) = \prod_{k=0}^{N(T)-1} q(X_k,\lambda,\mu)\exp\{-q(X_k,\lambda,\mu)t_k\} \times P(X_k,X_{k+1})\exp\{-q(X_{N(T)},\lambda,\mu)(T-T_{N(T)})\}$$

$$= \lambda^{N_0}\exp\left\{-\lambda\sum_{k\in N_0} t_k\right\}(\lambda+\mu)^{N_u}\exp\left\{-(\lambda+\mu)\sum_{k\in N_u} t_k\right\}$$

$$\times\left(\frac{\lambda}{\lambda+\mu}\right)^{N_u}(\lambda+\mu)^{N_d}\exp\left\{-(\lambda+\mu)\sum_{k\in N_d} t_k\right\}\left(\frac{\mu}{\lambda+\mu}\right)^{N_d}$$

$$\times\left[\exp\left\{-(\lambda+\mu-1_{\{X_{N(T)}=0\}}\mu)\right\}(T-T_{N(T)})\right]$$

$$= \lambda^{N_0}\exp\{-\lambda T\}\cdot\exp\left\{-\mu\sum_{k\in N(T)-N_0} t_k\right\}$$

$$\times\lambda^{N_u}\mu^{N_d}\exp\left\{-\mu 1_{\{X_{N(T)}\neq 0\}}(T-T_{N(T)})\right\}.$$

For a given parameter value λ_0, the Likelihood ratio is given by

$$L(T,\lambda_0,\lambda) = \frac{\Gamma(T,\lambda)}{\Gamma(T,\lambda_0)} = \left(\frac{\lambda}{\lambda_0}\right)^{N_o+N_u}\exp\{-T(\lambda-\lambda_0)\}.$$

For a given parameter value μ_0, the Likelihood ratio is given by

$$L(T,\mu_0,\mu) = \frac{\Gamma(T,\mu)}{\Gamma(T,\mu_0)} = \left(\frac{\mu}{\mu_0}\right)^{N_d}$$

$$\times\exp\left\{-(\mu-\mu_0)(\sum_{k\in N(T)-N_0} t_k + 1_{\{X_{N(T)}\neq 0\}}[T-T_{N(T)}]\right\}.$$

3.1 Example

Consider an M/M/1 system which starts empty and runs for time T. We are interested in knowing the average number in the system at time T, for various arrival rates. Assume we are not aware that this can be solved analytically. We simulate at arrival rate λ and use LRPE to estimate the average number in the system at time T for several other arrival rates $\lambda_0 = \lambda + \Delta$. It has been shown that as Δ increases, the variance of the LRPE estimate increases very rapidly (Rubinstein (1986); Arsham et al. (1989)). For the interpolation approach we simulate at arrival rates λ_1 and λ_2 to estimate the average number in the system at time T for several other arrival rates λ_0. The M/M/1 model was simulated using the following parameters. Arrival rates λ_1 = 1 and λ_2 = 2, service rate μ = 2, and T = 3. The LRPE was applied to estimate the average number in the system at time T for 9 perturbed arrival rates between 1 and 2. Table 2 presents the crude Monte Carlo simulation for the average number in the system at time T, LRPE estimates for the nine rates from λ_1, and the interpolation estimates with their corresponding variances. It is clear from this table that the interpolation approach leads to smaller variances.

Table 2: M/M/1 Example

λ_o	CMS		LRPE		INT	
	$\tilde{\ell}(\lambda_0)$	Var	$\tilde{\ell}(\lambda_0)$	Var	$\tilde{\ell}(\lambda_0)$	Var
1.1	0.88	1.31	0.88	1.67	0.84	0.84
1.2	1.07	1.87	1.00	2.60	1.01	0.98
1.3	1.22	2.08	1.12	4.18	1.13	0.99
1.4	1.34	2.39	1.20	6.50	1.21	1.08
1.5	1.44	2.35	1.32	10.4	1.41	1.14
1.6	1.64	2.83	1.33	16.2	1.55	1.27
1.7	1.77	3.18	1.46	25.0	1.72	1.58
1.8	1.96	3.66	1.57	37.2	1.87	2.05
1.9	2.16	4.01	1.58	55.0	2.05	2.80

4 CONCLUSION

We have presented an interpolation technique that use the likelihood performance extrapolation approach to estimate the expected performance of a stochastic system for various values of the input parameters from a single simulation of the nominal system. We have shown, through extensive experimentation, that the proposed technique is an effective tool for measuring parameter sensitivity in the context of the M/M/1 queueing models in transient analysis. We have presented the usefulness of the continuous time Markov Chain for computing the likelihood ratio (Radom-Nikodym derivative) of the underlying stochastic system.

REFERENCES

Arsham, H. , A. Feuerverger, D. L. Mcleish, J. Kreimer, and R.Y. Rubinstein. 1989. Sensitivity Analysis and the “What If” Problem in Simulation Analysis, *Mathl. Comput. Modeling* 12: (2) 193-219.

Cooper, R. B. 1981. *Introduction to Queueing Theory*. 2nd ed. New York: North Holland .

Glynn, P. W. 1986. Stochastic Approximation for Monte Carlo Optimization. *In Proceedings of the 1986 Winter Simulation Conference,* J. Wilson, J. Henriksen and S Roberts (eds.). IEEE Press, 356-364.

Glynn, P. W. 1987. Likelihood Ratio Gradient Estimation: An Overview. *In Proceedings of the 1987 Winter Simulation Conference,* A. Thesen, H. Grant and W D Kelton (eds.). IEEE Press, 366-375.

Glynn, P. W. 1990. Likelihood Ratio Derivative Estimators for Stochastic Systems. *Comm. ACM* :33, 75-84.

Gross, D. and C. M. Harris. 1998. *Fundamentals of Queueing Theory*. 3rd ed. New York: John Wiley & Sons.

Kleinrock, L. 1975. *Queueing Systems*. New York: John Wiley & Sons.

L'Ecuyer, P. 1990. A Unified View of the IPA, SF and LR Gradient Estimation Techniques. *Management Science* 36: 1364-1383.

L'Ecuyer, P. 1995. Note: On the Interchange of Derivative and Expectation for Likelihood Ratio Derivative Estimators, *Management Science* 41: 738-748.

Nakayama, M. K., A. Goyal, and P.W. Glynn. 1994. Likelihood Ratio Sensitivity Analysis for Markovian Models of Highly Dependable Systems. *Operations Research* 42: 137-157.

Reiman , M. I. , and A. Weiss. 1989. Sensitivity Analysis for Simulations Via Likelihood Ratios. *Operations Research* 37: 830-844.

Rubinstein, R.Y. 1986. The Score Function Approach for Sensitivity Analysis of Computer Simulation Models. *Math Comput Sim* 28: 351-379.

Rubinstein, R. Y. 1989. Sensitivity Analysis and Performance Extrapolation for Computer Simulation Models. *Operations Research* 37: 72-81.

AUTHOR BIOGRAPHIES

MOHAMED A. AHMED is an associate professor in the Department of Statistics and Operations Research at Kuwait University. He received a B.S. in Production Engineering from Helwan University-Egypt, a M.S. and a Ph.D. in Operations Research from George Washington University. His research interests lie in the area of inventory control, simulation, stochastic optimization, and applied probability. His e-mail address is `<wakeel@kuc01.kuniv.edu.kw>`.

TALAL M. ALKHAMIS is an associate professor in the Department of Statistics and Operations Research at Kuwait University. He received a B.S. in Computer Science & Statistics from Kuwait University, a M.S. in Computer Science, and a Ph.D. in Operations Research from the Florida Institute of Technology. His research interests include discrete Optimization, meta-heuristics, and stochastic optimization. His e-mail address is `<alkhamis@kuc01.kuniv.edu.kw>`.