Proceedings of the 1996 Winter Simulation Conference
ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain

VARIANCE PROPERTY OF DISCONTINUOUS PERTURBATION ANALYSIS

Leyuan Shi

Department of Industrial Engineering
University of Wisconsin-Madison
Madison, WI 53706

ABSTRACT

In recent article we presented a Discontinuous Perturbation Analysis (DPA) to deal with discontinuous sample performance functions. DPA is applicable to derivative estimation with respect to most threshold type of problems in Discrete Event Dynamic Systems (DEDS). In this paper, we examine the variance property of DPA method through a few examples. We will demonstrate that a better variance reduction can be obtained via DPA approach even if IPA Infinitesimal Perturbation Analysis (IPA) is applicable. A numerical example is also provided to compare DPA with other derivative estimation techniques.

1 INTRODUCTION

In the traditional simulation setting, one simulation corresponds to an estimate of the performance measure at one value of the parameter. None of the existing variance reduction methods could change this simple fact. But the driving force behind the development of Perturbation Analysis and Likelihood Ratio methods is the question of whether or not there is more information obtainable from that one simulation in the sense of estimating performance at other values of the parameter. From efforts made to tackle this issue came the development of Perturbation Analysis and later Likelihood Ratio method, which, simply put, is a means by which we compute derivative estimates from a single simulation (or actual observation) of the system.

Perturbation Analysis (PA) was initiated by Ho et al. (Ho and Cao, 1991), (Cassandras,1993) in the context of production lines in manufacturing systems. Its infinitesimal version, Infinitesimal Perturbation Analysis (IPA), was for the sensitivity analysis of throughput in closed queueing network. Since then, it has been successfully applied to many different systems. There also has been a considerable effort to extend the theory of IPA to handle situations where IPA fails (Brémaud and Vázquez-Abad, 1992), (Fu and Hu, 1992), (Gong and Ho, 1987), (Cassandras, 1993), (Shi, 1996), (Dai and Ho 1994). Likelihood Ratio (LR), also known as Score Function (SF) estimator has been proposed recently in the area of importance sampling (Glynn, 1987), (Reiman and Weiss, 1989), (Rubinstein, 1989). The relations between PA and LR have been discussed in (L'ecuyer, 1990).

To give the basic idea of IPA and LR, let us consider a stochastic system with a parameter θ. Suppose the performance measure of the system is

$$J(\theta) = E_\omega[L(x(t,\theta,\omega))].$$

where x is a random variable whose c.d.f. (cumulative distribution function) is $F(\theta,x)$ depending on θ. To simplify the discussion, we assume that x is a scalar. Thus, we have

$$J(\theta) = \int_{-\infty}^{\infty} L(x)dF(\theta,x). \qquad (1)$$

Let $\xi = F(\theta,x)$. Then ξ is a random variable uniformly distributed on (0,1], and x is a function of ξ:

$$x = x(\theta,\xi) = F^{-1}(\theta,\xi) = inf\{x : f(\theta,\xi) \geq \xi\}. \qquad (2)$$

Equation (1) is equivalent to

$$J(\theta) = \int_0^1 L(x(\theta,\xi))d\xi. \qquad (3)$$

On the other hand, let $f(\theta,x)$ be the p.d.f. (probability density function), then (1) can be written as

$$J(\theta) = \int_{-\infty}^{\infty} L(x)f(\theta,x)dx. \qquad (4)$$

Taking the derivatives of both sides of (3) and (4), we obtain

$$\frac{\partial}{\partial\theta}J(\theta) \quad = \quad \frac{\partial}{\partial\theta}\int_0^1 L(x(\theta,\xi))d\xi$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} L(x(\theta,\xi))d\xi \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial x} L(x) \frac{\partial x}{\partial \theta} d\xi \\
&= E[\frac{\partial L(x)}{\partial x}\frac{\partial x}{\partial \theta}]. \qquad (5)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial \theta} J(\theta) &= \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} L(x) f(\theta, x) dx \\
&= \int_{-\infty}^{\infty} L(x) \frac{\partial}{\partial \theta} f(\theta, x) dx \\
&= \int_{-\infty}^{\infty} L(x) \frac{\partial}{\partial \theta} [\ln f(\theta, x)] f(\theta, x) dx \\
&= E[L(x) \frac{\partial}{\partial \theta} \ln f(\theta, x)], \qquad (6)
\end{aligned}
$$

respectively, provided that the two operations, "E" and "$\frac{\partial}{\partial \theta}$" are interchangeable in each case. Equation (6) also requires $f(\theta, x) \neq 0$. If the interchangeability in (5) or (6) holds, then

$$\frac{\partial L(x)}{\partial x}\frac{\partial x}{\partial \theta} \qquad (7)$$

or

$$L(x)\frac{\partial \ln f(\theta, x)}{\partial \theta} \qquad (8)$$

is an unbiased estimate of $\partial J(\theta)/\partial \theta$. Equation (7) corresponds to the Infinitesimal Perturbation Analysis (IPA) estimate, and Equation (8) is called the Likelihood Ratio (LR) estimate of $\partial J(\theta)/\partial \theta$.

In general, the IPA estimate has a smaller variance, enjoys the nice property of convergence with probability one and a certain amount of robustness with respect to uncertainties in the underlying distribution (Ho & Cao, 1991), but requires more restrictive conditions than the LR estimate. On the other hand, the LR estimate usually has a bigger variance, but applies to a wider range of systems since the requirement for interchangeability of two operators, "E" and "$\frac{\partial}{\partial \theta}$", is much milder.

As we can see in (2) that we use inversion to represent the random variable. It is well-known that there exists more than one way to represent a parametric family of random variables or stochastic processes with given marginal probability distributions. Different representations often give rise to different sample path derivatives. The problem of determining which sample path derivative has a smaller variance when multiple strongly consistent and unbiased sample path derivatives exist is important in practice. In this short paper we use an example to demonstrate that variance reduction can be obtained when we resort DPA approach. We also show that there is a trade-off between accuracy and computation effort.

2 SAMPLE PATH DERIVATIVE AND ITS VARIANCE

Sample path derivatives can be used as derivative estimates of performance measures in stochastic discrete event systems such as computer networks and communication systems. These derivative estimates can be calculated based on a single sample path using the technique of IPA or other derivative estimate techniques. Usually, there exists more than one way to represent a parametric family of random variables or stochastic processes with given marginal probability distributions. Different representations often give rise to different sample path derivatives which have the same mean but different variances. To see this, let us define the sample path derivative of a parametric random variable.

Consider a parametric family of cumulative distribution functions (c.d.f.'s) $\{F(x;\theta) : \theta \in \Theta\}$. A family of random variables $\{X(\theta)\}$ defined on a common probability space (Ω, F, P) is a representation of $\{F(x;\theta) : \theta \in \Theta\}$ if $X(\theta)$ is a representation of $F(x;\theta)$ for every $\theta \in \Theta$, i.e.,

$$Prob(X(\theta) < x) = F(x;\theta).$$

Given a family of $\{X(\theta) : \theta \in \Theta\}$ defined on (Ω, F, P). For simplicity of notation, we assume that $\theta \in \Theta = [a, b]$. Denote

$$\delta X(\theta, h) = \frac{X(\theta + h) - X(\theta)}{h},$$

where $\theta, \theta + h \in \Theta$.

Definition 1 $\{X(\theta)\}$ is a.s. differentiable at $\theta \in \Theta$ if there exists a random variable $Y(\theta)$ such that

$$\lim_{h\to 0} \delta X(\theta, h) = Y(\theta) \quad a.s.$$

We also call $Y(\theta)$ the sample path derivative of $X(\theta)$ and denote $dX(\theta)/d\theta = Y(\theta)$.

Definition 2 $\{X(\theta)\}$ is a.s. continuous (piecewise differentiable, differentiable) on Θ if $P(X(q)$ is continuous (piecewise differentiable, differentiable) on $\Theta) = 1$.

Definition 3 $\{X(\theta)\}$ is *admissible* (with respect to θ) if $\{X(\theta)\}$ is a.s. continuous and piecewise differentiable on Θ, a.s. differentiable for every $\theta \in \Theta$ and $E[\sup_{\theta\in\Theta} dX(\theta)/d\theta] < \infty$.

Definition 4 if $\{X(\theta)\}$ is admissible and if $\{X(\theta)\}$ is also a representation of $\{F(x;\theta)\}$, then we say it is an *admissible* representation.

Given a parametric family of c.d.f.'s $\{F(x;\theta)\}$, there may exist more than one admissible representation. To illustrate this, we first define the inverse transform for a c.d.f. $G(x)$ to be

$$G^{-1}(u) = \sup\{y : G(y) \le u\}.$$

It is well-known that $G^{-1}(U)$ is a representation of $G(x)$, where U is a uniform random variable over [0,1]. It then follows that $\{F^{-1}(U;\theta)\}$ is a representation of $\{F(x;\theta)\}$.

Example 1. Suppose $F_i(x) = 1 - e^{-\lambda_i x}$ $(x \geq 0, i = 1,2)$. Let $F(x;\theta) = \theta F_1(x) + (1-\theta)F_2(x)$, where $0 < \theta < 1$. Then it is easy to verify that $\{F^{-1}(U;\theta)\}$ is an admissible representation of $\{F(x;\theta)\}$. On the other hand, note that under the condition that $0 \leq U \leq \theta$ (or $\theta < U \leq 1$), $\frac{\theta - U}{\theta}$ (or $\frac{U-\theta}{1-\theta}$) is also uniformly distributed over [0,1]. Therefore,

$$\left\{ I_{[0 \leq U \leq \theta]} F_1^{-1}(\frac{\theta - U}{\theta}) + I_{[\theta \leq U \leq 1]} F_2^{-1}(\frac{U-\theta}{1-\theta}) \right\}$$

is also an admissible representation.

Now let ξ be another uniform random variable over [0,1] which is independent of U. We can verify that

$$\{ I_{[0 \leq \xi \leq \theta]} F_1^{-1}(U) + I_{[\theta \leq \xi \leq 1]} F_2^{-1}(U) \}$$

is also a representation of $F(x;\theta)$. But ?? is not an *admissible* representation because it is not a.s. continuous on Θ. In the next section we will see that variance reduction can be obtained by using such a representation.

Lemma 1(Hu, 1993) If $\{F^{-1}(U,\theta)\}$ is an admissible representation, then

$$\frac{dF^{-1}(U;\theta)}{d\theta} = \frac{\partial F(x;\theta)/\partial\theta}{\partial F(x;\theta)/\partial x} \mid_{x=F^{-1}(U;\theta)},$$

and

$$Var\left[\frac{dF^{-1}(U;\theta)}{d\theta}\right] \leq Var\left[\frac{dX(\theta)}{d\theta}\right],$$

where $\{X(\theta)\}$ is any admissible representation of $\{F(x;\theta)\}$.

The lemma tells us that if inversion gives admissible representation, its corresponding sample path derivative has the smallest variance. In other words, inversion is the best representation in all **admissible** representations.

3 VARIANCE REDUCTION VIA DPA APPROACH

Many experiments have shown that IPA, when applicable, produces smaller variance of estimation error (based on fixed simulation length) than others (Glasserman & Yao, 1991). However IPA fails for most threshold (or structure) type of problems such as routing problem and queue length. The reason is mainly due to the sample path discontinuity with respect to the parameter in question, i.e., infinitesimal change in parameter value produces finite change in the sample path of the system. There has been a considerable effort to extend the theory of IPA to handle the situation where IPA fails (Shi, 1996),(Fu and Hu, 1992),(Gong and Ho, 1987), (Brémaud and Vázquez-Abad, 1992), (Dai and Ho 1994).

In (Shi, 1996), we propose another alternative approach called Discontinuous Perturbation Analysis (DPA) which tries to deal with discontinuity of sample performance functions, especially for threshold type of parameters or structure parameters. In this approach we use the step function to represent a discontinues sample performance function. We then introduce the d-function to sample path derivative. With this format, we are able to provide a framework for constructing the derivative estimation of certain classes of discrete event systems. In the same paper we also showed that the DPA formulation derived from this approach is very closely related with other exiting methods such as SPA (Smooth Perturbation Analysis) (Gong and Ho, 1987), RPA (Rare Perturbation Analysis) (Brémaud and Vázquez-Abad, 1992) and LR. In some cases, the SPA formula can be derived from DPA formula. A major result of the DPA approach can be summarized as the following:

Let

$$L(\theta,\omega,n) = L_i^k(\theta,\xi,\eta),$$

if $g^{k-1}(\theta) < \xi_i \leq g^k(\theta), k = 1, \ldots M,$
and $0 = g^0(\theta) < g^1(\theta) < .. < g^k(\theta) < .. < g^M(\theta) = 1,$

Theorem (Shi, 1996) Assume that for all i and k, $L_i^k(\theta,\xi,\eta)$ satisfy
i) $L_i^k(\theta,\xi,\eta)$ is uniformly differentiable with respect to θ,
ii) $L_i^k(\theta,\xi,\eta)$ is a continuous function of ξ,
then we have

$$\begin{aligned}\frac{\partial}{\partial\theta}E[L(\theta,\omega,n)] &= E[\frac{\partial}{\partial\theta}L(\theta,\omega,n) \mid_{IPA}] \\ &+ E[\sum_{i=1}^{n}\sum_{k=1}^{M}(L_{i\theta}^k - L_{i\theta}^{k+1})\frac{dg_i^k(\theta)}{d\theta}]\end{aligned}$$

where $L_{i\theta}^k = L_i^k(\xi_i = g_i^k(\theta))$.

In the following we will provide an example to illustrate that when both IPA and DPA are applicable, DPA formula will provide a smaller variance than IPA formula, i.e., inversion is not the best representation in all **possible** representations.

Example 2. Consider a simple cyclic queueing network consisting of two servers and only one customer.

The service time distribution for server one is deterministic S1=m and for server two is $F(\theta, S_2) = \theta(1-e^{-\lambda_1 S_2}) + (1-\theta)(1-e^{-\lambda_2 S_2})$. The state process of this system can be decomposed into regenerative periods, each consisting of one busy period and one idle period for each server. The mean length of a regenerative period is

$$E[L(\theta)] = E[S_1 + S_2(\theta)] = \mu + \frac{\theta}{\lambda_1} + \frac{1-\theta}{\lambda_2}.$$

Its derivative with respect to θ is

$$\frac{\partial}{\partial\theta}E[L(\theta)] = \frac{1}{\lambda_1} - \frac{1}{\lambda_2}.$$

The IPA estimate of $\frac{\partial}{\partial\theta}E[L(\theta)]$ is

$$\begin{aligned} e\,|_{IPA} &\equiv \frac{\partial}{\partial\theta}E[L(\theta)]\,|_{IPA} \\ &= \frac{\partial}{\partial\theta}[S_1 + S_2(\theta)] \\ &= -\frac{\partial F(\theta, S_2)/\partial\theta}{\partial F(\theta, S_2)/\partial S_2} \\ &= \frac{e^{-\lambda_1 S_2} - e^{-\lambda_2 S_2}}{\theta\lambda_1 e^{-\lambda_1 S_2} + (1-\theta)\lambda_2 e^{-\lambda_2 S_2}}. \end{aligned}$$

The LR estimate of $\frac{\partial}{\partial\theta}E[L(\theta)]$ is

$$\begin{aligned} e\,|_{LR} &\equiv \frac{\partial}{\partial\theta}E[L(\theta)]\,|_{LR} \\ &= (\mu + S_2(\theta))\frac{\lambda_1 e^{-\lambda_1 S_2} - \lambda_2 e^{-\lambda_2 S_2}}{\theta\lambda_1 e^{-\lambda_1 S_2} + (1-\theta)\lambda_2 e^{-\lambda_2 S_2}}. \end{aligned}$$

On the other hand, as we mentioned in the previous section that $S_2(\theta)$ can also be represented as

$$\{I_{[0\le\xi\le\theta]}F_1^{-1}(\eta) + I_{[\theta\le\xi\le1]}F_2^{-1}(\eta)\}$$

That is

$$S_2(\theta) = \begin{cases} -\frac{1}{\lambda_1}\ln(\eta) & \text{if } 0 \le \xi \le \theta \\ -\frac{1}{\lambda_2}\ln(\eta) & \text{if } 1 \ge \xi > \theta \end{cases}$$

where ξ and h are two independent random variables with uniform distribution over [0,1]. Clearly $S_2(\theta)$ is not an admissible representation.

Using Discontinuous Perturbation Analysis (DPA) (Shi, 1996), we can use the step function to represent $S_2(\theta)$. We have

$$S_2(\theta) = -\frac{1}{\lambda_1}\ln(\eta) + (\frac{1}{\lambda_1}\ln(\eta) - \frac{1}{\lambda_2}\ln(\eta))H(\xi - \theta).$$

where

$$H(\xi - \theta) = \begin{cases} 0 & \text{if } \xi < \theta, \\ 1 & \text{if } \xi \ge \theta. \end{cases}$$

Then

$$\frac{d}{d\theta}S_2(\theta) = -(\frac{1}{\lambda_1}\ln(\eta) - \frac{1}{\lambda_2}\ln(\eta))\delta(\xi - \theta),$$

where δ-function is defined by

$$\delta(\xi - \theta) = \begin{cases} \infty & \text{if } \xi = \theta, \\ 0 & \text{if } \xi \ne \theta. \end{cases}$$

and for any continuous function $f(\xi)$

$$\int_0^1 f(\xi)\delta(\xi - \theta)d\xi = f(\theta).$$

To be able to calculate the sample path derivative we need to remove δ-function, hence the DPA estimate of $\frac{\partial}{\partial\theta}E[L(\theta)]$ is

$$\begin{aligned} e\,|_{DPA} &\equiv \frac{\partial}{\partial\theta}E[L(\theta)]\,|_{DPA} \\ &= E_\xi[\frac{dS_2(\theta)}{d\theta}] \\ &= \frac{1}{\lambda_2}\ln(\eta) - \frac{1}{\lambda_1}\ln(\eta). \end{aligned}$$

It is easy to verify that

$$E[e\,|_{DPA}] = E[e\,|_{IPA}] = E[e\,|_{LR}] = \frac{1}{\lambda_1} - \frac{1}{\lambda_2}.$$

However, their variances differ significantly. In fact,

$$Var[e\,|_{DPA}] = (\frac{1}{\lambda_1} - \frac{1}{\lambda_2})^2,$$

and

$$\begin{aligned} &Var[e\,|_{IPA}] \\ &= \int_0^\infty \frac{(e^{-\lambda_1 S_2} - e^{-\lambda_2 S_2})^2}{\theta\lambda_1 e^{-\lambda_1 S_2} + (1-\theta)\lambda_2 e^{-\lambda_2 S_2}}dS_2 \\ &\quad -(\frac{1}{\lambda_1} - \frac{1}{\lambda_2})^2, \end{aligned}$$

and

$$\begin{aligned} &Var[e\,|_{LR}] \\ &= \int_0^\infty (\mu + S_2)\frac{(\lambda_1 e^{-\lambda_1 S_2} - \lambda_2 e^{-\lambda_2 S_2})^2}{\theta\lambda_1 e^{-\lambda_1 S_2} + (1-\theta)\lambda_2 e^{-\lambda_2 S_2}}dS_2 \\ &\quad -(\frac{1}{\lambda_1} - \frac{1}{\lambda_2})^2. \end{aligned}$$

Proposition 1. When $2\lambda_1 \le \lambda_2$, we have

$$\begin{aligned} \lim_{\theta\to0} Var[e\,|_{IPA}] &= \infty, \\ \lim_{\theta\to0} Var[e\,|_{LR}] &= \infty. \end{aligned}$$

Proof:

Table 1: Comparison of the Variances of DPA and SPA

θ	$\rho(\lambda=1)$	N_r/N_c*	$dE[T]/d\theta$	DPA Mean(Var)	SPA Mean(Var)
0.2	0.62	100/500	1.4157	1.39619 (0.100597)	1.37187 (0.109193)
0.2	0.62	200/500	1.4157	1.40181 (0.099802)	1.39485 (0.129505)
0.5	0.65	30/1000	-0.798639	0.79770 (0.022511)	-0.81832 (0.026849)
0.5	0.65	200/500	-0.798639	-0.7951 (0.053081)	-0.7952 (0.072534)
0.7	0.57	30/1000	-1.78983	-1.71851 (0.034805)	-1.73729 (0.033423)
0.7	0.57	40/1000	-1.78983	-1.7359 (0.0332479)	-1.74807 (0.031902)

*N_r: Number of replications. *N_c: Number of customers simulated per replication.

We only prove when $\lambda_1 = 1$ and $\lambda_2 = 2$. When $\lambda_1 = 1$ and $\lambda_2 = 2$ we have

$$\int_0^\infty \frac{(e^{-\lambda_1 S_2} - e^{-\lambda_2 S_2})^2}{\theta\lambda_1 e^{-\lambda_1 S_2} + (1-\theta)\lambda_2 e^{-\lambda_2 S_2}} dS_2$$
$$= \int_0^\infty \frac{e^{-2S_2} - 2e^{-3S_2} + e^{-4S_2}}{\theta\lambda_1 e^{-S_2} + (1-\theta)\lambda_2 e^{2S_2}} dS_2$$
$$= \int_0^\infty \frac{e^{-2S_2}}{\theta\lambda_1 e^{-S_2} + (1-\theta)\lambda_2 e^{2S_2}} dS_2$$
$$+ \int_0^\infty \frac{-2e^{-3S_2} + e^{-4S_2}}{\theta\lambda_1 e^{-S_2} + (1-\theta)\lambda_2 e^{-2S_2}} dS_2.$$

It is easy to see that the second term is bounded by a finite number. Now let us take a look the first term. Since

$$\int_0^\infty \frac{e^{-2S_2}}{\theta\lambda_1 e^{-S_2} + (1-\theta)\lambda_2 e^{-2S_2}} dS_2$$
$$= \int_0^\infty \frac{1}{\theta\lambda_1 e^{S_2} + (1-\theta)\lambda_2} dS_2 \quad (\text{let } u = e^{S_2})$$
$$= \int_0^\infty \frac{1}{\theta\lambda_1 u + (1-\theta)\lambda_2} \cdot \frac{1}{u} du$$
$$= \frac{1}{(1-\theta)\lambda_2} \left[\ln((1-\theta)\lambda_2 + \theta\lambda_1) - \ln(\theta\lambda_1)\right] .$$

That is

$$\lim_{\theta\to 0} Var[e \mid_{IPA}] = \infty.$$

Use the same, way we can also prove that

$$\lim_{\theta\to 0} Var[e \mid_{LR}] = \infty.$$

Figure 1 gives the comparison of the variance of the DPA, IPA, and LR estimates when we choose $\lambda_1 = 1$ and $\lambda_2 = 2$. We can see that the variance of the DPA estimate is a constant and the variance of IPA and LR increases when θ approaches to zero.

This example provides some insight in sample path derivative techniques. Usually, when IPA is applicable, we will no longer look for new derivative estimate formula, given the believe that IPA would be the most efficient derivative estimation method. Little has been done in previous research to compare the IPA estimator with other methods except for LR. This topic certainly deserves further research. In the following we give another example to show that there is a trade-off between accuracy (smaller variance) and computation effort.

Example 3. An M/G/1 queue is simulated. The service time is given as the follows

$$S_i = \begin{cases} 0.8 - \xi_i & \text{if } 0 \le \xi_i \le \theta, \\ \xi_i & \text{if } 1 \ge \xi_i\, \theta, \end{cases}$$

where $0 < \theta \le 0.8$.

The performance measure we are interested in is the mean system time. Since S_i is discontinuous with respect to θ, IPA will not work. There has been a considerable effort to extend the theory of IPA to handle situations where IPA fails. DPA is one of such efforts. Another well-known method is Smooth Perturbation Analysis (SPA) (Ho and Gong, 1986). Here we compare these two methods.

For this problem, the DPA estimate formula is

$$\frac{\partial}{\partial\theta} E[L] \mid_{DPA} = \sum_{i=1}^{M} (L_{+i\theta} - L_{-i\theta}),$$

where M is the total number of customer simulated and the SPA estimate formula has the following form

$$\frac{\partial}{\partial\theta} E[L] \mid_{SPA} = \frac{1}{\theta} \sum_{i\in A} (L - L_{-i\theta}),$$

where $A = \{i : \xi_i > \theta \text{ and } 1 \le i \le M\}$.

Table 1 shows that overall the DPA estimate formula provides a slightly smaller variance than SPA. On the other hand, even though the DPA estimate has smaller variance than the SPA, the DPA estimate needs a slightly more computation effort than SPA (Shi, 1996). This clearly shows that there is a trade-off between accuracy and computation effort.

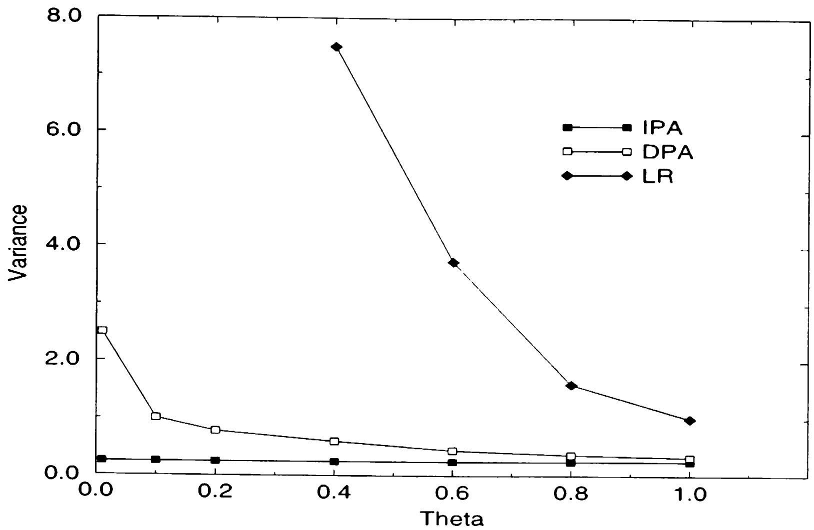

Figure 1: 90% Comparison of Variance of IPA, LR, and DPA

REFERENCES

Brémaud, P. and Vázquez-Abad F., 1992. On the pathwise computation of derivatives with respect to the rate of point process: The phantom RPA method. *Queueing Systems* 10: 249-270

Cassandras, C.G., 1993. *Discrete Event Systems: Modeling and Performance Analysis.* Irwin, March.

Dai, L. and Ho, Y.C., 1994. Structural infinitesimal perturbation analysis (SIPA) for derivative estimation in discrete event systems. Submitted to *Journal of Discrete Event Dynamic Systems.*

Fu, M.C. and Hu, J.Q., 1992. Extensions and generalizations of smoothed perturbation analysis in a generalized semi-markov framework. *IEEE Trans. on Automatic Control* 37: 1483-1500.

Hu. J.Q., 1993. Sample path derivatives: strong consistency and variance properties. Manuscript.

Glasserman, P., 1991. *Gradient Estimation Techniques for Discrete Event Systems.* Kluwer, MA.

Glasserman, P and Yao, D., 1991. Algebraic structure of some discrete event systems and their applications. *Discrete Event Dynamic Systems: Theory and Applications*, 1: 7-36.

Glynn, P.W., 1987. Likelihood ratio gradient estimation: an overview. *Proc. of 1987 Winter Simulation Conference* 366-375.

Gong, W.B. and Ho, Y.C., 1987. Smoothed perturbation analysis of discrete-event dynamic systems. *IEEE Transactions on Automatic Control* 33: 827-838.

Ho, Y.C. and Cao, X.R., 1991. *Discrete Event Dynamic Systems and Perturbation Analysis.* Kluwer Academic.

L'Ecuyer, P., 1990. A unified view of the IPA, SF, and LR gradient estimation techniques. *Management Science* 36: 1364-1383.

Reiman, M.I. and Weiss, A., 1989. Sensitivity analysis for simulation via likelihood ratios, *Operations Research* 37: 830-844.

Rubinstein, R. Y., 1989. Sensitivity analysis and performance extrapolation for computer simulation models. *Operations Research* 37: 72-81.

Shi, L, 1996. Discontinuous perturbation analysis for discontinuous sample performance functions. To appear in *IEEE Trans. on Automatic Control.*

AUTHOR BIOGRAPHY

LEYUAN SHI is an assistant professor in the Department of Industrial Engineering at University of Wisconsin-Madison. She holds a B.S. degree in mathematics from Nanjing Normal University, Chian (1982) and an M.S. degree in applied mathematics from Tsinghua University, China (1985), and M.S. and Ph.D degrees in applied mathematics from Harvard University (1990,1992). Her research interests include modeling, analysis, and optimization of discrete event systems, discrete-event simulation, and sensitivity analysis.