A UNIFIED VIEW OF INFINITESIMAL PERTURBATION ANALYSIS AND LIKELIHOOD RATIOS

by

Pierre L'Ecuyer

TECHNICAL REPORT No. 32

February 1989

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A Unified View of
Infinitesimal Perturbation Analysis
and Likelihood Ratios

Pierre L'Ecuyer[1]

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

ABSTRACT

We present a view of the likelihood ratio (LR) gradient estimation technique (also called the score function (SF) method), under which infinitesimal perturbation analysis (IPA) can be viewed as a (degenerate) special case, by selecting appropriately what the random component ω effectively represents. Varying the actual meaning of ω (i.e. defining the underlying sample space in different ways) might define different variants of the LR method, some of them mixing IPA with more traditional LR. We illustrate this by many examples. We also give general conditions under which the gradient estimators are unbiased.

Keywords: simulation, gradient estimation, likelihood ratios, perturbation analysis.

---

[1]Operations Research Department, Stanford University, on leave from Département d'informatique, Université Laval, Ste-Foy, Québec, Canada, G1K 7P4.

1. The LR gradient estimation technique

Consider a stochastic simulation model parametrized by a real vector $\theta \in \mathbf{R}^d$ of continuous parameters, and suppose we want to estimate the gradient $\nabla \alpha(\theta)$ of the (differentiable) expected value $\alpha(\theta)$ of some real-valued objective function. Two techniques have been proposed recently to estimate such a gradient by simulation: infinitesimal pertubation analysis (IPA) [5, 6, 7, 10, 12, 19] and likelihood ratios (LR) [8, 14, 15, 16].

The basic idea of LR is that $\alpha(\theta)$ can usually be viewed as the expectation of some function of θ and of the "sample path" ω, say $h(\theta, \omega)$, with respect to a probability measure $P_\theta(\cdot)$ over some measurable space (Ω, Σ). Here, Ω is the sample space, and $\omega \in \Omega$ represents all the "random elements" in the simulation, so that when it is fixed, the evolution of the system becomes deterministic. More specifically, we assume that $h(\theta, \cdot)$ is Σ-measurable. Usually, one cannot differentiate this expectation directly by differentiating inside the integral, and one of the reasons is that $P_\theta(\cdot)$ typically depends on θ. That dependence can be eliminated if one can take a probability measure $G(\cdot)$ on the same measurable space, independent of θ, and that dominates the $P_\theta(\cdot)$'s for θ in the region of interest. In that case, one can rewrite:

$$\alpha(\theta) = \int_\Omega h(\theta, \omega) P_\theta(d\omega) = \int_\Omega \left[h(\theta, \omega) \frac{P_\theta(d\omega)}{G(d\omega)} \right] G(d\omega) = \int_\Omega H(\theta, \omega) G(d\omega) \qquad (1)$$

where $H(\theta, \omega) = h(\theta, \omega) P_\theta(d\omega)/G(d\omega)$. The ratio $P_\theta(d\omega)/G(d\omega)$ is called the *Radon-Nikodym* derivative of $P_\theta(\cdot)$ with respect to $G(\cdot)$. By the Radon-Nikodym theorem ([1], theorem 32.2), it exists and (1) is valid if and only if $G(\cdot)$ dominates $P_\theta(\cdot)$, i.e. iff $P_\theta(\cdot)$ is absolutely continuous with respect to $G(\cdot)$, i.e. iff for every measurable set B, $G(B) = 0$ implies $P_\theta(B) = 0$. Note that if sampling is done using $G(\cdot)$, the bracketted term in (1) can be used as an estimator of $\alpha(\theta)$. This "change of measure" approach is called *importance sampling*, and is often used as a variance reduction technique [9, 16, 17].

Under appropriate *regularity conditions* (that permit to interchange the derivative and expectation), one can differentiate $\alpha(\cdot)$ by differentiating the bracketted term with respect to θ inside the integral:

$$\nabla \alpha(\theta) = \int_\Omega \psi(\theta, \omega) G(d\omega). \qquad (2)$$

where

$$\psi(\theta, \omega) = \nabla_\theta H(\theta, \omega) = \nabla_\theta h(\theta, \omega) \frac{P_\theta(d\omega)}{G(d\omega)} + h(\theta, \omega) \frac{\nabla_\theta P_\theta(d\omega)}{G(d\omega)}. \qquad (3)$$

Sufficient regularity conditions are given in [5, 6, 7, 14, 16] for special cases. More general conditions are given in section 3 of this paper. The need for these regularity conditions

and for the existence of the Radon-Nikodym derivative certainly impose some limitations on the method (see the examples in section 4), but there are also many practical cases where it works well.

When (2) holds, $\psi(\theta, \omega)$ can be used to estimate $\nabla\alpha(\theta)$. Note that only one simulation experiment (using $G(\cdot)$) is required to estimate the gradient. In principle, $\psi(\cdot, \omega)$ can be evaluated at any value of θ for which (2) holds, permitting to estimate of the gradient everywhere by a single simulation. But the variance of the gradient estimator is sometimes dramatically high for some values of θ (see e.g. [15]). The method can also be generalized to higher order derivatives (see [14, 15]).

How do we choose $G(\cdot)$? Among those $G(\cdot)$ for which (1-2) hold, one would like to choose one for which the variance is low. But this is not always easy to do. In practice, generating values of ω according to $G(\cdot)$ and computing $\psi(\theta, \omega)$ for any generated ω should also be done easily and at reasonably low cost. This usually limits the set of interesting choices.

To estimate $\nabla\alpha(\theta_0)$ at a single point θ_0, an easy choice for $G(\cdot)$ (when it is admissible) is $P_{\theta_0}(\cdot)$. In that case, things simplify in (3) and we obtain:

$$\nabla\alpha(\theta_0) = \int_\Omega \left[(\nabla_\theta h(\theta, \omega) + h(\theta, \omega)\nabla_\theta \ln P_\theta(d\omega)) \Big|_{\theta=\theta_0} \right] P_{\theta_0}(d\omega). \tag{4}$$

The expression $\nabla_\theta \ln P_\theta(d\omega)$ is called the *score function* (SF). The LR gradient estimator has been introduced [8, 14, 15] as the bracketted expression in (4), sometimes in a less general setting. In this paper, we take the following definition.

DEFINITION 1. A LR gradient estimator for $\nabla\alpha(\theta)$ is one that is defined by the bracketted expression in (4), where ω obeys P_{θ_0}, provided that this expression exists for almost all ω, and $h(\theta, \cdot)$ is a measurable function of ω whose expectation is $\alpha(\theta)$. ∎

The estimator defined in (3) can be viewed as a combination of LR with importance sampling. It is also a generalization of LR (see also [16]).

We have not talked much yet about the choice of the sample space Ω, but this is one of the key points in this paper. In fact, for a given simulation model, there can be different ways of defining the sample space and the meaning of ω. In many simulations, all random variables are generated by first generating $U(0,1)$ variates, and transforming them in the appropriate way. Hence, ω can be viewed as a sequence of independent $U(0,1)$ variates, and the value of the objective function $h(\theta, \omega)$ is a deterministic function of this sequence.

But this is only *one* way of viewing it. In fact, there is no need to assume that $U(0,1)$ variates are used to drive the simulation in the first place. For instance, the LR technique can be used to estimate the gradient not only for a simulation model, but also for a real system (provided $P_\theta(\cdot)$ is known and ω can be observed). In that case, ω will usually *not* be a sequence of $U(0,1)$ variates. Consider for instance a M/G/1 queue. One possibility is to view ω as the sequence of interarrival and service times, and define $\Omega = [0,\infty)^\infty$ as the sample space. Note that in that case, the distribution of ω depends on θ, whereas it does not when ω is defined as a sequence of $U(0,1)$ variates.

So, there might be different ways of defining the sample space (and the associated probability space $(\Omega, \Sigma, P_\theta(\cdot))$). Different choices may lead to different gradient estimators, some being more efficient than others. We will come back to this point in the next section.

Some might feel more confortable, for some reason, with an *underlying* probability space in which the "basic" random element is a sequence of independent $U(0,1)$ variates. Let $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P}(\cdot))$ be such a space. If $(\Omega, \Sigma, P_\theta(\cdot)) \neq (\tilde{\Omega}, \tilde{\Sigma}, \tilde{P}(\cdot))$, we assume that there is a measurable transformation $\phi_\theta : \tilde{\Omega} \to \Omega$ such that $\omega = \phi_\theta(\tilde{\omega})$, and such that $P_\theta(\omega \in \cdot) = \tilde{P}(\tilde{\omega} \in \phi_\theta^{-1}(\cdot))$ Note that ω *may* contain less information than $\tilde{\omega}$, and that its probability law *may* depend on θ. We can also define the $\tilde{P}(\cdot)$-measurable function $\tilde{h}(\theta, \cdot)$ by $\tilde{h}(\theta, \tilde{\omega}) = h(\theta, \phi_\theta(\tilde{\omega}))$ (note that $h(\theta, \cdot)$ is $P_\theta(\cdot)$-measurable). We have

$$\alpha(\theta) = \int_\Omega h(\theta, \omega) P_\theta(d\omega) = \int_{\tilde{\Omega}} \tilde{h}(\theta, \tilde{\omega}) \tilde{P}(d\tilde{\omega}).$$

Note that this notion of an underlying sample space $\tilde{\Omega}$ and transformation ϕ_θ is not really necessary. We introduced it here just to clarify some links with the common practice (in simulation) of viewing the sample space that way. What we really have in mind in to just define the sample space as Ω and forget about $\tilde{\Omega}$.

Usually, for convenience, ω is defined as a sequence of independent (univariate) random variables that represent the stochastic aspects of the simulation: $\omega = (\omega_1, \omega_1, \ldots)$. $P_\theta(d\omega)/d\omega$ is then a product of univariate density or/and probability mass functions, and $\nabla_\theta \ln P_\theta(d\omega)$ is a sum whose number of terms is typically the number of such univariate functions that depend on θ. Assuming that the variances of the sample performance $h(\theta, \omega)$ and gradient $\nabla_\theta h(\theta, \omega)$ are bounded, the variance of the LR gradient estimator increases linearly (in general) with that number (which is typically a linear function of the simulation length). From this reasoning, we should expect LR to work much better for terminating simulations for which only a small number of random variates are generated with probability laws that depend on θ.

3

Note that since regenerative simulations can be analyzed in a way very similar to terminating simulations [8, 9], the above remark also applies to steady-state regenerative simulations for which a small number of θ-dependent variates are generated per regenerative cycle. A version of the LR method specially adapted for regenerative systems is presented in [8].

2. Choosing what ω should represent

Suppose that ω is defined as a sequence of independent $U(0,1)$ variates. Then, $P_\theta(\cdot)$ is independent of θ, $H(\theta,\omega) = h(\theta,\omega)$, and the last term in (3) vanishes. In that case, equation (4) becomes (under the appropriate regularity conditions and with $\theta = \theta_0$):

$$\nabla\alpha(\theta) = \int_\Omega \nabla_\theta h(\theta,\omega)P_\theta(d\omega) \tag{5}$$

and $\nabla_\theta h(\theta,\omega)$ is the usual IPA gradient estimator ([2, 5, 6, 12, 16]).

DEFINITION 2. An IPA estimator for $\nabla\alpha(\theta)$ is defined as $\nabla_\theta h(\theta,\omega)$ (provided that this quantity exists for almost all ω), where ω is a sequence of independent $U(0,1)$ random variables, and $h(\theta,\cdot)$ is a measurable function of ω whose expectation is $\alpha(\theta)$. ∎

When (5) is satisfied, we have an *unbiased* IPA gradient estimator. Note that there often exists different functions $h(\theta,\cdot)$ that satisfy (1) and (5), and thus different unbiased IPA gradient estimators for the same $\alpha(\cdot)$. In practice, the function $h(\theta,\cdot)$ is usually defined or implied by the simulation model.

The basic idea of IPA is to generate a sample path ω, viewed as a sequence of $U(0,1)$ variates, and for ω fixed, observe the effect of an infinitesimal perturbation on θ (around θ_0) by propagating it over the sample path, assuming that the sequence of events do not change, and that the events only "slide" in time. The gradient estimation is taken as the gradient of the objective function for that fixed value of ω. Note that the propagation rules permit to evaluate $\nabla_\theta h(\theta,\omega)$ only at $\theta = \theta_0$, and thus to estimate the gradient only at θ_0. If that definition of ω is used in (2) with $G(\cdot) \neq P_\theta(\cdot)$, one gets a combination of IPA with importance sampling.

According to the above definitions, IPA can be viewed as a special case of LR. One big advantage of IPA is that since no component of ω depends on θ, the variance does not increase any more with the simulation length. But the function $h(\theta,\omega)$ must absorb all the tranformations and may become overly complex, sometimes making the actual computation of $\nabla_\theta h(\theta,\omega)$ intractable, or invalidating (2). In fact, a large part of the IPA literature deals with the development of effective techniques to compute $\nabla_\theta h(\theta,\omega)$ during the simulation (see e.g. [10, 18]). One might even associate the term IPA more with these techniques than with equation (5), and these "IPA" techniques can be used to implement LR as well. Some people might also argue that since the likelihood ratio has disappeared in (5), this is no more LR. In fact, we just view it as a degenerate case.

As we said before, ω can also represent something else than a sequence of $U(0,1)$ variables. For example, ω can be viewed as representing the whole history of the system, including all events with their types and occurence times, etc. In many cases, ω can carry enough θ-dependent high level information so that for a given value of ω, $h(\theta,\omega)$ does not depend any more on θ, and the first term of the right hand side expression in (3) vanishes. But then, one might be unable to write down $P_\theta(\cdot)$ explicitely, preventing the actual computation of the estimate. For this reason, ω will usually be taken as a sequence of independent random variables. In Glynn [8], for example, when simulating a Markov chain, ω is taken as the sequence of states visited by the chain.

For the extreme case where ω is defined as the value of the objective function itself, i.e. $h(\theta,\omega) = \omega$ (and $\Omega = \mathbb{R}$) by definition, $P_\theta(\cdot)$ is actually the distribution function of the cost. When we can write it down and write down the likelihood ratio, there is usually no need to simulate, since $\nabla \alpha(\theta)$ can be computed directly. Monte-Carlo methods are precisely useful for the cases where we *cannot* efficiently compute the expression directly.

Between these extremes, there is often different other possibilities. For instance, if a set of $U(0,1)$ values must go through many levels of transformation, one may choose any one of the levels to define ω. Also, ω might contain the original $U(0,1)$ values for some of the generated random variables, and the transformed values for others. This gives rise to hybrid methods, "mixing" in some way IPA with LR. According to our definitions, this is still LR. In section 4, we give examples for which one might think that LR does not apply, but for which LR effectively applies if the sample space and ω are defined appropriately.

But what is the best way, then, to define ω ? There is no easy answer to this question. There is no straightforward recipe. Of course, one would like (2) to be valid. There are examples for which (2) is valid if ω represents higher level information, like e.g. the set of actual interarrival times, service times and transitions between nodes in a queueing network, and not valid if ω represents the sequence of $U(0,1)$ variates. But for other examples, the opposite is true (see [14, 5, 6] and the examples in section 4). In certain situations, (2) might be valid for none of the extreme cases, but for some intermediate definition of ω (see example 4.5). If it is valid for many possible definitions of ω, one will then try to minimize the variance. This is certainly problem-dependent, but from the last two paragraphs of the previous section, trying to put the least number of θ-dependent components in ω appears to be a good strategy.

One consequence of the above discussion is that many properties of the IPA method also apply to LR, and vice-versa. For example, the validity of interchanging the derivative

6

and expectation is a problem for the LR method in general. Various (problem dependent) devices have been suggested to "smooth out" or transform some problems for which IPA doesn't apply directly, into problems for which IPA will work correctly (see e.g. [10] and the references in [5, 6, 12]). In principle, one could think of developping such devices for LR in general.

Note that for the case where $\alpha(\theta)$ is a steady-state performance measure and ω contains an infinite sequence of θ-dependent random variables, the Radon-Nikodym derivative in (1) typically does not exists. However, (5) might be valid in that case, and then, one typically has $\nabla_\theta h(\theta, \omega) = \nabla \alpha(\theta)$ with probability one (this is when IPA is strongly consistent).

3. Interchanging the Derivative and Expectation

In this section, we give sufficient conditions for the interchange of derivative and expectation leading to (2) to be valid. Conditions for specific cases are also given in [5, 6, 7, 14, 16], and are sometimes more direct to verify. But often, the conditions below can be verify directly, as we will see in the examples of the next section.

Note that each component of the gradient can be dealt with separately. For $i = 1, \ldots, d$, to study the i-th component of the gradient, we look at what happens when only the component i of θ is allowed to change and all other components of θ are fixed. To simplify things, in this section, we assume that $d = 1$. For the more general case, just apply the results below to each component of θ (while the other components are fixed). All probabilistic statements in this section are made with respect to the probability measure $G(\cdot)$. The lemma below is an adaptation of lemma 1 in [5]. It uses the following assumption:

A1. Let $d = 1$. There is a neighbourhood Υ of θ such that for almost all ω, $H(\cdot, \omega)$ exists and is continuous over Υ, and is differentiable everywhere in $D(\omega) \subseteq \Upsilon$, where $\Upsilon \setminus D(\omega)$ is at most a denumerable set. Assume that (1) is satisfied for all $\theta \in \Upsilon$. Also,

$$\sup_{v \in D(\omega)} |\psi(v, \omega)| \tag{6}$$

is integrable with respect to $G(\cdot)$ (note that $D(\omega)$ is the set where $\psi(\cdot, \omega)$ exists). ∎

LEMMA 1. *Under A1, equation (2) is valid.*

PROOF. The proof is largely inspired by the proof of lemma 1 in [5]. From a generalized version of the mean value theorem (see e.g. theorem 8.5.2 in [3]), if θ and $\theta + h$ are in Υ and ω satisfies the requirements of A1,

$$\left| \frac{H(\theta + h, \omega) - H(\theta, \omega)}{h} \right| \leq \sup_{v \in D(\omega)} |\psi(v, \omega)| \, .$$

Hence, from the dominated convergence theorem,

$$
\begin{aligned}
\int_\Omega \psi(\theta, \omega) G(d\omega) &= \int_\Omega \left(\lim_{h \to 0} \frac{H(\theta + h, \omega) - H(\theta, \omega)}{h} \right) G(d\omega) \\
&= \lim_{h \to 0} \left(\frac{\alpha(\theta + h) - \alpha(\theta)}{h} \right) \\
&= \nabla \alpha(\theta). \; \blacksquare
\end{aligned}
$$

8

4. Examples

In this section, we give a number of examples to illustrate the main ideas of the paper. For some of them, we also give numerical results. The first five examples deal with a simple M/G/1 queue that evolves until a certain (fixed) number of departures have occured. The next to examples consider the lifetime of a k-out-of-N reliability system, without repairs. The following one discuss a general continuous time Markov chain, while the last one looks at sensitivity with respect to thresholds. In particular, we look at replacement policies defined by thresholds in a multicomponent system. In the latter case, we actually *don't* know how LR can be used efficiently to estimate the gradient.

Consider a M/G/1 queue, initially empty, and let $\alpha(\theta)$ be the expected mean system time (waiting + service times) for the first T customers in the system, where θ is a parameter of the servive time distribution. The arrival rate is $\lambda = 1$. We want to estimate the derivative $\alpha'(\theta)$ at a given point $\theta = \theta_0$ by simulating at that point. For a given realization ω, $h(\theta, \omega)$ represents the observed average waiting time for the T customers. We have

$$h(\theta, \omega) = \frac{1}{T} \sum_{i=1}^{T} (W_i + S_i) \tag{7}$$

where W_i and S_i are respectively the actual (observed) waiting time and service time of customer i (these are deterministic functions of θ and ω). Let A_i denotes the interarrival time between customers $i - 1$ and i (A_1 is the arrival time of customer 1 and the system starts at time 0). We have $W_1 = 0$, and $W_{i+1} = \max(0, W_i - A_{i+1} + S_i)$ for $i > 0$. The first five examples below are variants of this one; only the service time distributions differ. Application of IPA to this system has been analyzed in [19] when the objective function is the *steady-state* average system time per customer. It has been shown that under some conditions on the service time distribution, IPA gives an asymptotically unbiased and strongly consistent gradient estimate. For the case of a finite number of customers (terminating simulation), the validity of IPA has been analyzed e.g. in [5], example 4.

4.1. A M/M/1 queue

Let the service time distribution be exponential with mean θ, $a \le \theta \le b$, where $0 < a < b$. IPA is known to work for that case: assuming that the interarrival and service times are generated by inversion, one can take ω as the sequence of $U(0,1)$ values used to generate them, i.e. $\omega = (U_1, \ldots, U_{2T})$, $A_i = -\ln(1 - U_{2i-1})$ and $S_i = -\theta \ln(1 - U_{2i})$. Here,

9

$G(\cdot) = P_\theta(\cdot)$. An infinitesimal perturbation on S_i affects the system time of customer i and of all the customers that follow him in the same busy period (if any). Therefore,

$$\psi(\theta, \omega) = \nabla_\theta h(\theta, \omega) = \frac{1}{T} \sum_{i=1}^{T} \sum_{j \in B_i} \frac{\partial S_j}{\partial \theta} \qquad (8)$$

where B_i is the set containing customer i and all the customers that precede him in the same busy period (if any), and $\partial S_i / \partial \theta = S_i / \theta$. This can be computed during the simulation as described in [19]. We can easily verify assumption A1. In fact, for any ω, S_i is continuous and differentiable in θ, and $h(\cdot, \omega)$ is continuous in the S_i's (and in the W_i's, which are continuous in the S_i's). Also, $h(\cdot, \omega)$ fails to be differentiable at θ only when two events (arrival or departure) occur simultaneously, and this happens at most for a finite number of values of θ. Since $\sup_{\theta \geq a} \psi(\theta, \omega)$ is clearly integrable, lemma 1 applies and IPA provides an unbiased estimate for that case.

Another choice is to take ω as the set of actual interarrival and service times: $\omega = (A_1, S_1, \ldots, A_T, S_T)$. In this case, $P_\theta(d\omega)/d\omega$ is the product of their densities:

$$P_\theta(d\omega) = \prod_{i=1}^{T} \left(\frac{1}{\theta} e^{-S_i/\theta} e^{-A_i} dS_i dA_i \right),$$

$G(d\omega) = P_{\theta_0}(d\omega)$, and $\nabla_\theta h(\theta, \omega) = 0$. Note that only the service time densities appear in the likelihood ratio, since the interarrival times are independent of θ (in fact, taking either the actual interarrival times or the corresponding $U(0, 1)$ values in ω makes no difference here). One has

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{i=1}^{T} \frac{\partial}{\partial \theta} \ln \left(\frac{1}{\theta} e^{-S_i/\theta} dS_i \right) = \frac{1}{\theta^2} \sum_{i=1}^{T} (S_i - \theta).$$

This can be computed easily together with $h(\theta, \omega)$ during the simulation. For any ω, $H(\cdot, \omega)$ is continuous and differentiable in $[a, b]$ (note that $H(\theta, \omega)$ depends on θ only throught $P_\theta(d\omega)$). Also, the gradient estimator is

$$\psi(\theta, \omega) = \frac{1}{T\theta^2} \left(\sum_{i=1}^{T} (W_i + S_i) \right) \left(\sum_{i=1}^{T} (S_i - \theta) \right)$$

and $\sup_{\theta \in [a,b]} |\psi(\theta, \omega)|$ is P_{θ_0}-integrable, so that lemma 1 applies.

In principle, one can also combine the two approaches and take S_i for some of the customers and the corresponding $U(0, 1)$ values for others. This would give rise to more complex expressions but could be implemented in practice without too much difficulty, and lemma 1 would still apply. There might be no practical advantage of doing such a

10

combination in that case, but there are other examples where it can be helpful. For the case where ω contains only the set of waiting times, one faces the problem of expressing $P_\theta(\cdot)$. In fact, the waiting times are *dependent* random variables whose distributions are quite complex in general. Except for small values of T, this is not practical. Finally, for the extreme case where $\omega = h(\theta, \omega)$, one has to compute the distribution of the average cost using numerical methods!

4.2. Discrete law with θ-dependent value of the probability mass

Let $0 < a < b$, $0 < \theta < 1$, and suppose that the service time is b with probability θ, and a with probability $1 - \theta$. In this case, IPA doesn't apply (see also [19]), but if ω contains the set of actual service times, then LR do apply.

Suppose $U_1, \ldots, U_T$ are the $U(0,1)$ variates used to generate the service times. Let $C_i = 1$ and $S_i = b$ if $U_i \leq \theta$, $C_i = 0$ and $S_i = a$ otherwise (C_i is Bernouilli (θ)). For IPA, ω contains $(U_1, \ldots, U_T)$, and $H(\cdot, \omega) = h(\cdot, \omega)$ is discontinuous since S_i jumps from b to a at $\theta = U_i$. This is why IPA doesn't work. But suppose $\omega = (A_1, S_1, \ldots, A_T, S_T)$. In this case, the likelihood ratio can be expressed in terms of the variables $C_1, \ldots, C_T$. Their joint probability mass is

$$p_\theta(C_1, \ldots, C_T) = \prod_{i=1}^{T} \theta^{C_i}(1 - \theta)^{1-C_i},$$

and $H(\theta, \omega) = K(\omega) p_\theta(C_1, \ldots, C_T)$, where $K(\omega) = h(\theta, \omega)/p_{\theta_0}(C_1, \ldots, C_T)$ do not depends on θ (because ω contains all the information to compute $h(\theta, \omega)$ independently of θ). $H(\cdot, \omega)$ is continuous and differentiable on $(0, 1)$, and since $K(\omega)$ is integrable and $p_\theta(\cdot) \leq 1$, A1 is satisfied and we get an unbiased estimate.

4.3. Discrete law with θ-dependent support

Suppose that the service time is θ with probability p, and 2θ with probability $1 - p$, where $\theta > 0$ is the parameter and p is a constant, $0 < p < 1$. Here, the "naive" application of LR, where ω contains the set of actual service times and $G(\cdot) = P_{\theta_0}(\cdot)$, doesn't apply because there is no neighbourhood of θ_0 in which the Radon-Nikodym derivative exists ($H(\theta, \omega) = 0$ everywhere except at $\theta = \theta_0$, and so it is discontinuous). For IPA, use $U_1, \ldots, U_T$ to generate the service times: $S_i = \theta$ if $U_i < p$, $S_i = 2\theta$ otherwise. $\psi(\theta, \omega)$ can be computed as in (8), again with $\partial S_i / \partial \theta = S_i / \theta$. The arguments to verify A1 are the same as in example 1, and so IPA applies.

4.4. A mixture of IPA and LR

Let q be a constant, $0 < q < 1$, and suppose that the service time is generated as in example 4.2 with probability q, and as in example 4.3 with probability $1-q$. Let $D_1, \ldots, D_T$ be the corresponding Bernouilli (q) random variables, i.e. $D_i = 1$ if the service time of the i-th customer is generated from the first distribution (example 4.2), $D_i = 0$ otherwise. In that case, we can define ω as the set of values of D_i, the values of C_i for the customers i for which $D_i = 1$, and all the $U(0,1)$ values used to generate the rest (the interarrival times and the other S_i's). When $D_i = 0$, define $C_i = 0$. Since the likelihood ratio will depend on ω only throught the D_i and C_i, let us define $\tilde{\omega} = (D_1, C_1, \ldots, D_T, C_T)$. The probability mass of $\tilde{\omega}$ is given by

$$p_\theta(\tilde{\omega}) = \prod_{i=1}^{T} q^{D_i}(1-q)^{1-D_i} \left(\theta^{C_i}(1-\theta)^{1-C_i}\right)^{D_i}.$$

and one has

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{i=1}^{T} D_i \left(\frac{C_i}{\theta} - \frac{1-C_i}{1-\theta}\right) = \sum_{i=1}^{T} \frac{D_i(C_i - \theta)}{\theta(1-\theta)}.$$

This can be computed with $h(\theta, \omega)$ during the simulation. To compute $\nabla_\theta h(\theta, \omega)$ for a fixed ω, one applies a mixture of the more traditional IPA and LR techniques: the service times of the customers for which $D_i = 0$ are "perturbed" using the usual IPA technique, while the perturbations for the other service times are considered to be zero. More specifically, for a fixed ω, $\nabla_\theta h(\theta, \omega)$ is computed using the right hand side of (8), but with

$$\frac{\partial S_i}{\partial \theta} = \begin{cases} 0 & \text{if } D_i = 1; \\ 1 & \text{if } D_i = 0 \text{ and } S_i = \theta; \\ 2 & \text{if } D_i = 0 \text{ and } S_i = 2\theta. \end{cases}$$

Again, assumption A1 is easily verified, by combining the arguments of the two previous examples. Therefore, by mixing IPA with LR, we obtain an unbiased gradient estimate, despite the fact that neither IPA alone nor "naive" LR (putting all the S_i's in ω) works.

Table 1 gives the results of a numerical experiment for this example. We used $T = 10$, $q = p = a = 1/2$, $b = 3/2$, and estimated the derivative at $\theta = 0.2, 0.5$ and 0.9. We used two gradient estimation techniques: symmetric finite differences with common random numbers (FDC), and the "hybrid" method described above (LR). For FDC, simulations were made at $\theta \pm 0.01$, starting from the same (empty) state, and the same $U(0,1)$ values were used on both sides, with proper synchronization. In each case, we made 100000 replications and computed a 95% confidence interval. Note that the same streams of random numbers were used for the six different entries of table 1. As expected, the results from the two

techniques agree. They also agree with the exact values, which were computed using dynamic programming. [To compute these exact values, we can write recursive equations to compute $V_n(s)$ and $\nabla_\theta V_n(s)$ (in terms of $V_{n+1}(\cdot)$ and $\nabla_\theta V_{n+1}(\cdot)$), where $V_n(s)$ represents the expected total system time spent *from now on* by the next $T - n$ customers to depart, given that there are s customers in the system, $1 \leq s \leq T - n$, one of which is beginning its service. V_{T-1} is the expected service time, and $V_0(1)/T$ is the expected average system time for the first T customers.] Note that for FDC, the gradient estimator has some bias, due to the finite differences, but here, that bias is "lost in the noise", since the confidence intervals cover the exact values.

θ	True grad.	LR	FDC
0.2	2.551	$2.53 \pm .04$	$2.55 \pm .03$
0.5	3.980	$3.96 \pm .06$	$4.01 \pm .04$
0.9	5.525	$5.53 \pm .16$	$5.53 \pm .05$

Table 1: Numerical results for example 4:
95% confidence intervals for the gradient, based on 10^5 replications.

4.5. θ times a Bernouilli (θ)

Suppose that the service time is $k\theta$ with probability θ, and 0 with probability $1 - \theta$, for some constant $k > 0$. For each service time, one can generate a $U(0,1)$ variate U_i, and define $S_i = k\theta$ if $U_i < \theta$, $S_i = 0$ otherwise. In that case, it is easy to see that neither $\omega = (A_1, U_1, \ldots, A_T, U_T)$, nor $\omega = (A_1, S_1, \ldots, A_T, S_T)$ will work. However, if one takes $\omega = (A_1, C_1, \ldots, A_T, C_T)$, where $C_i = S_i/(k\theta)$, then LR works. In fact, C_i is Bernouilli (θ). The score function is

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{i=1}^{T} \frac{(C_i - \theta)}{\theta(1 - \theta)},$$

and $\nabla_\theta h(\theta, \omega)$ can be computed using the right hand side of (8), with $\partial S_i/\partial\theta = kC_i = S_i/\theta$ (because $S_i = k\theta C_i$, and because for ω fixed, C_i is fixed).

For this example, we made the same numerical experiment as for the previous example, with $k = 2$, and the results appear in table 2. Again, they agree very well with the exact values (computed by dynamic programming).

θ	True grad.	LR	FDC
0.2	1.038	$1.03 \pm .01$	$1.03 \pm .01$
0.5	4.529	$4.49 \pm .05$	$4.51 \pm .04$
0.9	16.671	$16.61 \pm .30$	$16.75 \pm .12$

Table 2: Numerical results for example 5:
95% confidence intervals for the gradient, based on 10^5 replications.

4.6. Lifetime of a k-out-of-N system

Consider a k-out-of-N reliability system with identical components. The N components have independent (random) lifetimes durations $X_1, \ldots, X_N$, each with distribution $F_\theta(\cdot)$. For simplicity, assume that this distribution is continuous, with density $f_\theta(\cdot)$, and that for each $x > 0$, both $F_\theta(x)$ and $f_\theta(x)$ are differentiable w.r.t. θ. The system is down (failed) when less than k components are still alive. For a given realization ω, $h(\theta, \omega)$ represents the system's lifetime, and $\alpha(\theta)$ its expectation. We will examine four of the (many) possible choices for ω. We want to estimate $\nabla \alpha(\theta_0)$, using (4). Let s and X_s denote the number and lifetime of the last component that fails (the $(N - k + 1)$-th failure). One has $h(\theta, \omega) = X_s$.

One can take $\omega = (X_1, \ldots, X_N)$, in which case $\nabla_\theta h(\theta, \omega) = 0$,

$$H(\theta, \omega) = X_s \prod_{i=1}^{N} \frac{f_\theta(X_i)}{f_{\theta_0}(X_i)},$$

and

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{i=1}^{N} \frac{\partial}{\partial \theta} \ln f_\theta(X_i).$$

This is the most straightforward application of LR.

A second choice is to take ω as s and X_s, plus the set A of components that are still alive at time X_s. During the simulation, the N lifetimes can be generated as above (as usual), and $\nabla_\theta h(\theta, \omega) = 0$, but now,

$$P_\theta(d\omega)/d\omega = f_\theta(X_s) \prod_{i \in A} (1 - F_\theta(X_s)) \prod_{i \notin A \cup \{s\}} F_\theta(X_s),$$

and the score function is

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{i=1}^{N} L_i$$

14

where

$$L_i = \begin{cases} \dfrac{\partial}{\partial\theta}\ln f_\theta(X_s) & \text{if } i = s; \\[2mm] \dfrac{\partial}{\partial\theta}\ln(1 - F_\theta(X_s)) & \text{if } i \in A; \\[2mm] \dfrac{\partial}{\partial\theta}\ln F_\theta(X_s) & \text{otherwise.} \end{cases}$$

A third choice is to use the IPA technique: suppose $X_i = F_\theta^{-1}(U_i)$ where the U_i's are independent $U(0,1)$ variates, for $i = 1,\ldots,N$, and take $\omega = (U_1,\ldots,U_N)$. Here, $\nabla_\theta h(\theta,\omega) = \partial F_\theta^{-1}(U_s)/\partial\theta$, where s is defined as above, and $\nabla_\theta \ln P_\theta(d\omega) = 0$. Note that there are many cases where we can obtain equivalent results by taking ω as something else than the U_i's. Consider for example the Weibull distribution: $F(x) = 1 - \exp(\lambda x^\alpha)$ and $F^{-1}(U) = (-\lambda\ln(1 - U))^{1/\alpha}$. In this case, one can take $\omega = (V_1,\ldots,V_N)$ where $V_i = -\ln(1 - U_i)$ and this is equivalent to IPA (for $\theta = \lambda$ or $\theta = \alpha$). The expression for $\nabla_\theta h(\theta,\omega)$ can be obtained with slightly less manipulations in the latter case.

For exponential lifetimes, there is the following fourth choice. Let $1/\theta$ be the failure rate for each component, and let the *state* of the system be defined as the number of components that are still alive. That system evolves as a continuous time Markov chain. It goes from N to $N - 1,\ldots,$ to k, and finally to $k - 1$ where it dies. The jump rate from $N - i$ to $N - i - 1$ is $(N - i)/\theta$. Hence, ω can be defined as $\omega = (Y_N,\ldots,Y_k)$, where Y_j is exponential with mean θ/j, and $h(\theta,\omega) = Y_N + \cdots + Y_k$. One has in this case

$$\nabla_\theta \ln P_\theta(d\omega) = \sum_{j=k}^{N} \left(\frac{jY_j - \theta}{\theta^2} \right)$$

and $\nabla_\theta h(\theta,\omega) = 0$. Note also, that for the exponential case, analytical formulas are obtained readily: $\alpha(\theta) = E[h(\theta,\omega)] = E[Y_N + \cdots + Y_k] = \theta(1/N + \cdots + 1/k)$, and $\nabla\alpha(\theta) = (1/N + \cdots + 1/k)$. For many other distributions, analytical formulas can also be obtained by exploiting the fact that X_s is an order statistic. Often, one can write down the density (or prob. mass) and the expectation of X_s explicitly, and differentiate. This is what we did to compute the exact gradient for the Weibull case in table 4 below.

All this can be adapted easily to more general reliability networks, with more complex structures, components that have different lifetime distributions (with possibly different parameters), repair possibilities, etc. For the last case (exponential lifetimes and fourth choice of ω), the state description of the Markov chain would get more complicated in general. A "practical" analytical formula is not always available for the exponential case, but replacing transition times by their expectations is certainly a good idea (see also

	Exponential	Weibull
$f_\theta(x)$	$(1/\theta)e^{-x/\theta}$, $x > 0$	$\lambda\theta x^{\theta-1}e^{-\lambda x^\theta}$, $x > 0$
$F_\theta(x)$	$1 - e^{-x/\theta}$	$1 - e^{-\lambda x^\theta}$
$F_\theta^{-1}(U)$	$-\theta\ln(1-U)$	$(-\ln(1-U)/\lambda)^{1/\theta}$
$\dfrac{\partial}{\partial\theta}X$ for $X = F_\theta^{-1}(U)$	X/θ	$-X(\ln X)/\theta$
$\dfrac{\partial}{\partial\theta}\ln f_\theta(x)$	$(x-\theta)/\theta^2$	$1/\theta + (1-\lambda x^\theta)\ln x$
$\dfrac{\partial}{\partial\theta}\ln F_\theta(x)$	$\dfrac{-xe^{-x/\theta}}{\theta^2(1-e^{-x/\theta})}$	$\dfrac{\lambda x^\theta \ln x e^{-\lambda x^\theta}}{1-e^{-\lambda x^\theta}}$
$\dfrac{\partial}{\partial\theta}\ln(1-F_\theta(x))$	x/θ^2	$-\lambda x^\theta \ln x$

Table 3: Expressions used in the gradient estimators,
for two distributions.

example 7). In practice, importance sampling (when it works) can also be very effective when simulating such networks (see [17]).

We made some numerical experiments, first with the exponential distribution with mean θ: $F_\theta(x) = 1 - \exp(-x/\theta)$, then with a Weibull distribution with θ as the form parameter: $F_\theta(x) = 1 - \exp(-x^\theta)$, $x \geq 0$. We tried different values of N and k. The results appear in table 4, under the form of 95% confidence intervals on $\nabla\alpha(\theta)$. The "methods" LR1, LR2, IPA and MC correspond to the four choices of ω described above, in the same order. Note that MC does not apply for the Weibull distribution, but for the seven other cases, A1 can be verified easily for all $\theta > 0$. We leave that as an (easy) exercise to the reader. Expressions used in the estimators are given in table 3 for these two distributions.

From these numerical results, LR2 appears to be generally better than LR1, and MC is not much better than even LR1. The most efficient by far is certainly IPA.

4.7. A density with parameter-dependent support

In the previous example, suppose that the component lifetimes are uniformly distributed, between 0 and θ. The density of X_i is $f_\theta(x) = 1/\theta$ for $0 \leq x \leq \theta$. Since the support of $f_\theta(\cdot)$ depends on θ, if we take $\omega = (X_1, \ldots, X_N)$, the Radon-Nikodym derivative $P_\theta(\cdot)/P_{\theta_0}(\cdot)$ does not exists at $\theta > \theta_0$. For $\theta < \theta_0$, it exists in a neighbourhood of θ, and (1) is valid.

N	k	distrib.	θ	True grad.	LR1	LR2	IPA	MC
8	2	Expon.	1.0	1.7179	1.694±.042	1.705±.036	1.718±.004	1.732±.040
8	6	Expon.	1.0	0.4346	0.430±.011	0.436±.009	0.436±.001	0.440±.009
8	2	Expon.	5.0	1.7179	1.694±.042	1.705±.036	1.718±.004	1.732±.040
20	5	Expon.	1.0	1.5145	1.485±.048	1.483±.041	1.513±.002	1.523±.045
20	10	Expon.	1.0	0.7688	0.755±.025	0.764±.020	0.769±.001	0.771±.020
20	15	Expon.	1.0	0.3462	0.340±.012	0.347±.008	0.347±.001	0.350±.008
50	10	Expon.	1.0	1.6703	1.669±.024	1.674±.020	1.671±.001	1.650±.022
50	40	Expon.	1.0	0.2457	0.246±.004	0.246±.002	0.246±.001	0.246±.002
1	1	Weibull	1.0	-0.4228	-0.404±.033	-0.404±.033	-0.420±.010	
8	2	Weibull	1.0	-1.0756	-1.027±.057	-1.052±.041	-1.074±.008	
8	6	Weibull	1.0	0.2930	0.308±.012	0.296±.004	0.294±.001	
20	5	Weibull	1.0	-0.6843	-0.618±.063	-0.669±.031	-0.682±.004	
20	10	Weibull	1.0	0.1662	0.197±.030	0.169±.007	0.167±.001	
20	15	Weibull	1.0	0.3389	0.350±.014	0.342±.006	0.340±.001	

Table 4: 95% confidence intervals for the gradient,
based on 10^5 replications (example 5).

One may think of using that to estimate the gradient at θ, but the problem is that since $f_\theta(x) = 1/\theta$ for $x < \theta$ and 0 for $x > \theta$, the continuity assumption on $H(\cdot, \omega)$ does not holds. In fact, for any neighbourhood of θ, $H(\cdot, \omega)$ is discontinuous in that neighbourhood whenever $\max_i X_i$ is in it. Therefore, A1 is not satisfied. This illustrates the fact that even when $\nabla_\theta h(\theta, \omega) = 0$, the existence of the Radon-Nikodym derivative is not a sufficient condition for LR to apply.

Note that IPA applies for this case: the gradient estimator is X_s/θ, the same as for the exponential case.

4.8. A continuous-time Markov chain

Consider a continuous time Markov chain with finite state space S. Let λ_i denote the jump rate out of state i, and p_{ij} be the transition probability from i to j. There is also a cost incurred continuously at rate c_i when in state i. Suppose that these quantities depend on some parameter vector $\theta \in \mathbb{R}^d$. Let $\alpha(\theta)$ be the total expected cost for the first T transitions, where T is fixed.

Simulation is often the most convenient tool to analyze such chains, particularly for very large state spaces (see e.g. [9]). Note that here, an event list is not necessary to run the simulation; one can just use the transition probabilities to jump from state to state (see [4]). Typically, the transition matrix is very sparse, and from any given state i, the number of reachable states is small. In fact, there is usually no need to write down that matrix, neither to enumerate S. Take for instance a closed Jackson network with say 20 nodes and 100 customers (one server per node, one class of customers): the state space is huge, but from any given state with say B busy nodes, there are only B possibilities for the node where the departure occurs and at most 20 possibilities for the destination node of the departing customer. It is quite easy to generate the two corresponding discrete variables and there is no need to generate explicitly even a row of the transition matrix. (See also [4] for a slightly different approach.)

Let $\{X_n, n \geq 0\}$ be the embedded Markov chain (the sequence of visited states), $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$ the transition times (the system jumps into X_n at time τ_n), and for $n \geq 0$, $\zeta_n = \tau_{n+1} - \tau_n$. Note that ζ_n is exponential with mean $1/\lambda_{X_n}$.

A simple choice for ω is $\omega = (X_0, \zeta_0, X_1, \zeta_1, \ldots, X_{T-1}, \zeta_{T-1})$. Except for very small T, the variance of the gradient estimator is then usually quite high, and as the previous examples suggest, one would usually prefer to use IPA if it is applicable. Unfortunately, IPA rarely works for the transition probabilities, but it can be used here for the times between jumps: take $\omega = (X_0, U_0, \ldots, X_{T-1}, U_{T-1})$, where U_i is the $U(0,1)$ variate used to generate ζ_i (by inversion). This could certainly help, but there is a better choice: just take $\omega = (X_0, \ldots, X_{T-1})$. The ζ_i's can simply be replaced by their expectations. This reduces simultaneously the variance of the cost estimate and the variance of the likelihood ratio. In fact, there is no need to generate any ζ_i. The cost estimate is simply

$$\sum_{n=0}^{T-1} c_{X_n}/\lambda_{X_n}. \tag{9}$$

If all the p_{ij} depend on θ, there are still $T - 1$ terms in the score function (assuming X_0 fixed), and the variance could kill us for large T. There might be cases, however, where only *some* of the p_{ij}'s depend on θ, and this can makes a big difference in the variance of the gradient estimator. If only the transition times (and not the transition probabilities) are influenced by θ, then IPA applies and the gradient estimator is readily obtained by differentiating (9).

Note that on the other hand, terms in the likelihood ratios *cannot* be replaced (in general) by their expectations. For example, the score function in (4) has zero expectation,

but is correlated with $h(\theta, \omega)$ and the expectation of the product is *not* zero. There are special cases, however, where it can be done: see e.g. Algorithm B in [8].

With minor adaptations, the above approach also applies to semi-Markov processes, where the inter-jump times are no more exponential, and to the case where T is a random stopping time.

4.9.   Components replacement

(Taken from [11]). Consider a system comprised of N identical components, that evolve independently. Each component has a random lifetime distribution with increasing failure rate. Whenever a component fails, it must be replaced instantly by a new one. Other components may be replaced (preventively) at the same occasion. The repaiman can also halt the system at any moment and replace preventively any number of working components. All replacements are assumed instantaneous. A failure cost c_f is incurred every time a component fails. At each intervention, there is also a fixed cost c_i, and a replacement cost which is c_r times the number of components replaced. Preventive replacements are made to avoid some of the failures, and replacements are sometimes lumped together to pay the fixed cost less often.

Here, we restrict our attention to the (generally suboptimal) class of policies defined by two thresholds: $\theta_1 > \theta_2 > 0$. Whenever a component fails or reaches age θ_1, the repaiman intervenes and replaces all components older than θ_2. We are interested in the total cost for a fixed duration T, assuming that all components are new at the beginning. The parameter here is $\theta = (\theta_1, \theta_2)$.

Unfortunately, applying LR to estimate the gradient in this case is still an open problem. Suppose ω is the set of generated component lifetimes (for fixed θ, this is enough to compute the cost). Then, the likelihood ratio is always one, since the component lifetimes do not depend on θ, but $h(\theta, \omega)$ is discontinuous in θ. More specifically, for any neighbourhood Υ of $\theta = (\theta_1, \theta_2)$, $h(\cdot, \omega)$ will be discontinuous in Υ if some component lifetimes are near enough θ_1 or θ_2 to change the sequence of failures when θ changes inside Υ (for fixed ω), and the set of values of ω for which this happens has positive probability. Exactly the same problem occurs with IPA (ω is the set of $U(0,1)$ values used to generate the component lifetimes). Suppose now that ω includes the sequence of all failures and replacements, with their times. Now, $h(\theta, \omega)$ becomes independent of θ, but in general, the Radon-Nikodym derivative doesn't exists in a neighbourhood of θ, since a typical ω will have non-zero probability for only *one* value of θ.

For the moment, we don't know how to apply LR or IPA to estimate the gradient for this example. Many other examples, most of them involving "threshold" parameters, fall into this category. For instance, think of a (s, S) inventory systems, where $\theta = (s, S)$, or a time-sharing computer system where the parameter is the quantum size, or a checkpoint–rollback–recovery system (for databases) where θ is the time between checkpoints (or is used in a rule to decide the next checkpoint time, based on the state of the system), etc. At present, for all these examples, to the best of our knowledge, a "finite differences" approach (preferably with common random numbers) must be used.

5. Conclusion

We pointed out the strong relationship that exists between IPA and LR. Section 3 also provides easily verifiable conditions under which IPA and/or LR apply. When they do not apply, these conditions sometimes permit us to understand why. We have illustrated with examples some ideas related to this approach. In particular, there are often many different ways to implement LR, some being much more efficient than others. IPA and more "traditional" LR (i.e. as used for instance in [8, 14, 15, 16]) can sometimes be combined on the same problem, and for the same parameter. However, when IPA applies, it is typically the most efficient method.

In practice, the change of measure used to define (1–3) can sometimes be used to reduce the variance (importance sampling [17]). We have not explored that issue in this paper. For all the examples in section 4, we have used $G(\cdot) = P_{\theta_0}$ to estimate $\nabla \alpha(\theta_0)$, but substantial variance reductions can sometimes be obtained by using a different (and appropriate) $G(\cdot)$.

For some examples, it appears that finite differences (FD) remains the only applicable approach at this time. Some experimental evidence ([13] and example 4 in this paper) suggests that FD with common random numbers might be practically as good as IPA when θ has only one component ($d = 1$). But for large numbers of parameters (large value of d), performing all the simulations required for FD becomes rather time consuming, and a good LR implementation might beat FD significantly.

Acknowledgements

References

[1] Billingsley, P., *Probability and Measure*, Wiley, New York, 1979.

[2] Cao, X. R., "Sensitivity Estimates Based on One Realization of a Stochastic System", *J. Stat. Comput. Simul.*, **27** (1987), 211–232.

[3] Dieudonné, J., *Foundations of Modern Analysis*, Second edition, Academic Press, New York, 1969.

[4] Fox, B. L., "Beating Future–Event Schedules", Report no. 761, School of Op. Res. and Indus. Eng., Cornell Univ. (1987).

[5] Glasserman, P., "Performance Continuity and Differentiability in Monte Carlo Optimization", *Proceedings of WSC'88*, The Society for Computer Simulation (1988), 518–524.

[6] Glasserman, P., "Structural Conditions for Perturbation Analysis Derivative Estimation: Finite Time Performance Indices", Manuscript, Harvard University (1988).

[7] Glasserman, P., "Structural Conditions for Perturbation Analysis of Queueing Systems", Manuscript, Harvard University (1988).

[8] Glynn, P. W. "Likelihood Ratio Gradient Estimation: an Overview", *Proceedings of WSC'87*, The Society for Computer Simulation (1987), 366–375.

[9] Glynn, P. W. and Iglehart, D. L., "Simulation Methods for Queues: an Overview", *Queueing Systems*, **3** (1988), 221–255.

[10] Gong, W. B. and Ho, Y. C. "Smoothed (Conditional) Perturbation Analysis of Discrete Event Dynamical Systems", *IEEE Trans. on Automatic Control*, **AC-32**, 10 (1987), 858–866.

[11] Haurie, A. and L'Ecuyer, P., "Approximation and Bounds in Discrete Event Dynamic Programming", *IEEE Trans. on Automatic Control*, **AC-31**, 3 (1986), 227–235.

[12] Heidelberger, P., Cao, X.-R., Zazanis, M. A. and Suri, R., "Convergence Properties of Infinitesimal Perturbation Analysis Estimates", to appear in *Management Science* (1989).

[13] L'Ecuyer, P., Giroux, N. and Glynn, P. W., "Stochastic Optimization by Simulation: Some Experiments With Simple Steady-State Queues", in preparation (1989).

[14] Reiman, M. I. and Weiss, A., "Sensitivity Analysis for Simulation via Likelihood Ratios", to appear in *Operations Research* (1989).

[15] Rubinstein, R. Y., "The Score Function Approach for Sensitivity Analysis of Computer Simulation Models", *Math. and Computers in Simulation*, **28** (1986), 351–379.

[16] Rubinstein, R. Y., "Sensitivity Analysis of Computer Simulation Models via the Efficient Score", to appear in *Operations Research* (1989).

[17] Shahabuddin, P., Nicola, V. F., Heidelberger, P., Goyal, A. and Glynn, P. W., "Variance Reduction in Mean Time to Failure Simulation", *Proceedings of WSC'88*, The Society for Computer Simulation (1988), 491–499.

[18] Suri, R., "Infinitesimal Perturbation Analysis for General Discrete Event Systems", *J. of the ACM*, **34**, 3 (1987), 686–717.

[19] Suri, R. and Zazanis, M., "Perturbation Analysis Gives Strongly Consistent Estimates for the M/G/1 Queue", *Management Science*, **34**, 1 (1988), 39–64.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
Technical Report NO. 32	ARO 25839.5-MA

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Dept. of Operations Research		U. S. Army Research Office

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Stanford, CA 94305-4022	P. O. Box 12211 Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
U. S. Army Research Office		DAAL03-88-K-0063

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
P. O. Box 12211 Research Triangle Park, NC 27709-2211	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

A Unified View of Infinitesimal Perturbation Analysis and Likelihood Ratios

12. PERSONAL AUTHOR(S) Pierre L'Ecuyer

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	February 1989	23

16. SUPPLEMENTARY NOTATION The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	simulation, gradient estimation, likelihood ratios, perturbation analysis.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We present a view of the likelihood ratio (LR) gradient estimation technique (also called the score function (SF) method), under which infinitesimal perturbation analysis (IPA) can be viewed as a (degenerate) special case, by selecting appropriately what the random component ω effectively represents. Varying the actual meaning of ω (i.e. defining the underlying sample space in different ways) might define different variants of the LR method, some of them mixing IPA with more traditional LR. We illustrate this by many examples. We also give general conditions under which the gradient estimators are unbiased.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted   SECURITY CLASSIFICATION OF THIS PAGE
All other editions are obsolete   UNCLASSIFIED