SAMPLE PATH DERIVATIVES FOR (s, S) INVENTORY SYSTEMS

MICHAEL C. FU

University of Maryland, College Park, Maryland

(Received August 1990; revisions received February 1991, April 1992; accepted October 1992)

For (s, S) inventory systems, we derive sample path derivatives of performance measures with respect to the two parameters s and S. These derivatives yield derivative estimators which can be estimated from a single sample path or simulation of the inventory system, in some cases not even requiring actual knowledge of the underlying demand distribution. Such derivative estimates would be useful in sensitivity analysis or in gradient-based optimization techniques. We consider the nondiscounted periodic review system with general independent and identically distributed (i.i.d.) continuous demands, full backlogging, and general holding and shortage costs. For the infinite horizon model, consistency proofs are given for some special cases, although we argue why the estimators should be correct for the more general case.

Sample path derivatives have been found to be useful in the analysis and optimization of queueing systems. For an overview and extensive bibliography, see Ho and Cao (1991). The primary purpose of this paper is to introduce the idea of sample path derivatives for inventory models. For (s, S) inventory systems, we derive sample path derivatives of performance measures with respect to the two parameters s and S. Such derivatives could be useful in at least two applications: testing the sensitivity of costs to variations in s and S; and choosing optimal (or at least better) values for the levels s and S. Although this paper is a modest beginning, in that it deals with a system which is about as tractable as the $GI/G/1$ (e.g., integral equations must be solved in the most general case), it is hoped that by demonstrating the technique and proving its correctness for a fairly simple system, the way is paved for further avenues of research on more complicated systems. Thus, we stress that this technique should be applicable to more complicated systems, with the analysis, in general, having to be done on a case-by-case basis, depending on the system structure, performance measures of interest, and parameters of interest. For more complicated systems, however, actually *proving* correctness becomes a more formidable task.

The analysis of (s, S) inventory policies has been a topic of research for nearly four decades since the introduction of the multistage periodic review inventory model in the seminal work by Arrow, Harris and Marschak (1951). The system of interest involves a single item where once every period the inventory level is reviewed and, if necessary, orders are placed to replenish depleted inventory. An (s, S) ordering policy specifies that an order be placed when the level of inventory on hand plus that on order falls below the level s, and the amount of the order be the difference between S and the present level of inventory on hand and on order, i.e., order amounts are placed "up to S." The usual objective of inventory control is to find an ordering policy to minimize a cost function, where costs are associated with ordering, holding, and shortages. Specification of the system requires the demand characteristics, the order lead time characteristics, the policy for unmet demand (e.g., backlogged or lost), and the associated costs. For a certain class of models, Scarf (1960) and Iglehart (1963) showed that an optimal policy can be found within the class of (s, S) policies, and Veinott and Wagner (1965) gave a computational method to compute the corresponding optimal values of s and S for the case of discrete demands.

The usual means of finding the optimal values of s and S is either through dynamic programming or through stationary analysis. The dynamic programming method is a recursive means of finding the optimal values, but does not give any structural insight and is not well-suited to sensitivity analysis, whereas the stationary analysis approach gives structural insight, but analytic results on performance measures are available only for restricted cases and usually involve numerical methods which do not allow for easy sensitivity analysis or optimization. Our approach is sample path analysis, with the focus on

Subject classifications: Inventory/production: sample path derivatives, performance evaluation. Simulation: sensitivity analysis.
Area of review: MANUFACTURING, OPERATIONS AND SCHEDULING.

0030-364X/94/4202-0351$01.25

sample path derivatives. In terms of simulation, this means that estimates of derivatives can be obtained in a *single* simulation run, simultaneously with estimates of performance measures. In this paper, we consider only the technique of perturbation analysis. For a discussion of an alternative approach called the likelihood ratio method, see, e.g., L'Ecuyer (1990), who surmised that the likelihood ratio method was not applicable to the (s, S) inventory system, and that "to the best of our knowledge, a 'finite-differences' approach must be used" (p. 1380). In this paper, we show that one form of perturbation analysis can be applied successfully to this problem.

Since we are interested in derivative estimation, which requires continuity of the performance measure of interest, our model assumes that demands take on values from a continuum; thus, our model is along the lines of Iglehart (1963) and Sahin (1982), who address structural results, and in contrast to the discrete demand case considered by the majority of authors in addressing the problem of finding optimal values of s and S, e.g., Veinott and Wagner (1965), Federgruen and Zipkin (1984), Porteus (1985), and Zheng and Federgruen (1991). A continuous model is reasonable for products where the measures are continuous (e.g., weight or volume), or the demand reflects a cumulative amount of large quantities (e.g., the number of screws in the millions).

The rest of the paper is organized as follows. In Section 1, derivations of derivative estimators for functions of the inventory level, such as average (positive) inventory and backorder, are presented. The derivations are intended to illustrate the technique of perturbation analysis. The system is sufficiently complex that the simplest form of perturbation analysis known as infinitesimal perturbation analysis (IPA) does not suffice, and we are led to applying a more involved form of perturbation analysis known as smoothed perturbation analysis (SPA). The resulting estimators are novel in that they are the first to apply the SPA technique to derivatives with respect to *structural* parameters of the system (s and S), as opposed to parameters of probability distributions in the timing of events (e.g., interarrival and service times in queueing systems), such as found in the work of Glasserman and Gong (1990) and Fu and Hu (1992). Strong consistency is proven for the estimators by comparison with analytical results in Section 2. Since the derivation is via sample path analysis, distributional assumptions on the demand and the lead times are minimal; in general, the consistency of sample path derivative estimators is not very distribution-dependent. In Section 3, the average cost per period performance measure is tackled. For the special case of linear costs and constant lead time, consistency proofs are provided in Section 4. Section 5 concludes the paper with a brief summary.

1. AVERAGE INVENTORY AND BACKORDER

Our model is a periodic review inventory system with general independent and identically distributed (i.i.d.) continuous demands, full backlogging, and general holding and shortage costs. Intuitively, the idea behind the derivation of the sample path derivative estimators is the one used in deriving perturbation analysis estimators of queueing systems, i.e., a *thought* experiment of introducing a perturbation into the sample path and tracing its effect, as in IPA (see Ho and Cao 1983, Suri and Zazanis 1988); or tracing its expected effect, as in SPA (see Gong and Ho 1987). The former contribution is straightforward, while the latter involves computing the expected effect of a perturbation, causing a change in the ordering pattern along the sample path. In general, this conditional expectation may require additional simulation, but the quantity can be estimated easily from the original sample path for the periodic review inventory system with full backlogging. It is hoped that the figures that accompany the derivations prove useful in understanding the underlying concepts of the technique.

1.1. Description of the Problem

Let

Y_n = the inventory position in period n;
W_n = the inventory level in period n;
D_n = the demand in period n (i.i.d. for all n);
$F(\cdot)$ = the distribution function of D_n;
L = the order lead time;
$g(\cdot)$ = a continuous and piecewise differentiable, positive real-valued function of one variable.

By inventory position, we mean inventory on hand plus on order minus backorders; by inventory level, we mean inventory on hand minus backorders. Order lead time is the time between the placement of an order and its arrival; zero lead time means instantaneous replenishment. Since we assume full backlogging, Y_n and W_n may be negative. We assume that F is absolutely continuous with density function f and that g and its derivative have at most a countable number of discontinuities.

We assume that at each review period, the ordering decision is made at the beginning of the period and the demand for the period is subtracted at the end of the period, with inventory position for period n,

Y_n, defined *after* order placement but *before* demand satisfaction for the period. A typical sample path for Y_n is shown in Figure 1. Under this convention, the (s, S) ordering policy stipulates that if the inventory position after demand satisfaction falls below s, then at the beginning of the next period, an order is placed for an amount to bring the inventory position up to $S = s + q$; otherwise, no order is placed. Thus, the recursive dynamic equation for Y_n is given by

$$Y_{n+1} = \begin{cases} Y_n - D_n & \text{if } Y_n - D_n \geq s \\ S & \text{if } Y_n - D_n < s \end{cases}. \tag{1}$$

For convenience in the analysis that follows, we will assume that $Y_0 = S$.

Equation (1) is not as well-behaved as the Lindley equation for the $GI/G/1$ queue, in that it is not continuous across the breakpoint. This is basically why the straightforward sample path derivative estimate (the infinitesimal perturbation analysis estimate) is not enough for one of the derivatives.

Orders are assumed to arrive at the beginning of the period. We will assume throughout this paper that order lead time is a fixed constant, in which case we have the following relationship between the inventory level and the inventory position:

$$W_{n+L} = Y_n - \sum_{i=n}^{n+L-1} D_i. \tag{2}$$

Thus, for zero lead time, W_n coincides with Y_n. Extensions to stochastic lead times can be found in Fu and Hu (1994).

For the remainder of the paper, we define $q = S - s$, and consider the equivalent problem of derivative estimation with respect to q and s. Our performance measure in this section will be

$$g_n(s, q) = \frac{1}{n} \sum_{i=1}^{n} g(W_i), \tag{3}$$

i.e., the time average of a function of the inventory level in each period. In particular, we are interested in the following forms of g: $g(x) = x^+$, $g(x) = x^-$, where

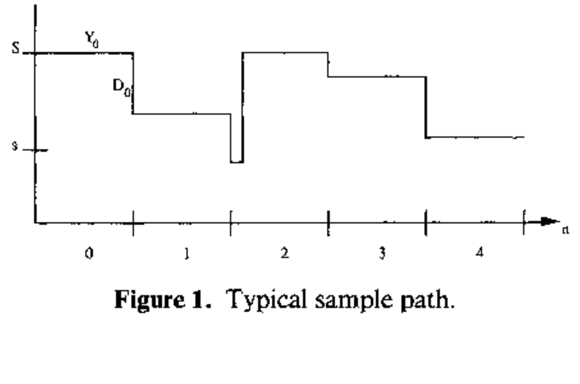

Figure 1. Typical sample path.

$x^+ = \max(0, x)$, and $x^- = \max(0, -x)$, in which g_n would denote the average (positive) inventory and average backorder, respectively.

Henceforth, for convenience, we let θ denote some parameter of the system, which for our example is either q or s (but in general could be, say, a parameter of the demand distribution). Our goal in this section is to estimate $\partial E[g_n(s, q)]/\partial\theta$, $\theta = s$ or q via perturbation analysis.

In general, the technique of IPA will suffice if w.p.1 a small change in the parameter of interest, say $\Delta\theta$, causes only a small (or possibly no) change in the performance measure of interest, i.e., of $O(\Delta\theta)$. If, however, there are cases (w.p. $O(\Delta\theta)$) where $\Delta\theta$ causes a *finite* change in the performance measure, then IPA alone will not give the correct answer, because it does not calculate the effects of these contributions, in which case conditional expectation can often be used to estimate the effects via a technique called smoothed perturbation analysis (Gong and Ho). Our analysis will show that the derivative estimator with respect to s consists of just an IPA contribution, whereas the derivative estimator with respect to q consists of both an IPA contribution and SPA contribution. The purpose of this section is to derive these two contributions. We then prove the strong consistency of the resulting estimators in the next section. Since consistency proofs are provided, liberties are sometimes taken in some of the derivation steps, sacrificing rigor in the name of clarity and insight (we hope).

1.2. IPA Estimation

The IPA estimator comes from taking the derivative in (1) while assuming that the event $\{Y_n - D_n \geq s\}$ (equivalently $\{Y_{n+1} = S\}$) is unchanged, so we have for both $\theta = s$ and $\theta = q$,

$$\frac{dY_{n+1}}{d\theta} = \begin{cases} \dfrac{dY_n}{d\theta} & \text{if } Y_n - D_n \geq s \\ 1 & \text{if } Y_n - D_n < s \end{cases}, \tag{4}$$

for all n. With the initial condition $Y_0 = S = s + q$, we have

$$\frac{dY_n}{d\theta} = 1 \tag{5}$$

for all n, and (2) gives

$$\frac{dW_n}{d\theta} = 1 \tag{6}$$

for all $n \geq L$.

Since the derivative of g has at most a countable number of discontinuities, and demand is a continuous random variable for any fixed values of s and q, the points of discontinuities in $g'(W_n)$ occur w.p.0. Thus, we apply the chain rule to get the derivative for g:

$$\frac{dg(W_n(\theta))}{d\theta}=\frac{\partial g(W_n(\theta),\theta)}{\partial\theta}+g'(W_n)\frac{dW_n}{d\theta} \quad \text{w.p.1.} \qquad (7)$$

The IPA estimator is thus

$$\frac{\partial g_n(s,q)}{\partial\theta}=\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial g(W_i(\theta),\theta)}{\partial\theta}+g'(W_i)\frac{dW_i}{d\theta}\right]$$

$$=\frac{1}{n}\sum_{i=L}^{n}\left[\frac{\partial g(W_i(\theta),\theta)}{\partial\theta}+g'(W_i)\right]. \qquad (8)$$

In the limit as $n\to\infty$, the effects of the lower limit L can be ignored; henceforth, we will replace the lower limit of the summation with 1 to ease the notational burden in all the analyses that follow.

In particular, for the inventory and backorder, we have

$$g(x)=x^+,\quad \frac{\partial g(W_i(\theta),\theta)}{\partial\theta}=0,$$

$$g'(W_n)=\begin{cases}1 & \text{if } W_n>0\\ 0 & \text{if } W_n<0\end{cases}, \qquad (9)$$

$$g(x)=x^-,\quad \frac{\partial g(W_i(\theta),\theta)}{\partial\theta}=0,$$

$$g'(W_n)=\begin{cases}0 & \text{if } W_n>0\\ -1 & \text{if } W_n<0\end{cases}, \qquad (10)$$

the derivative being undefined at the point $W_n=0$, which occurs w.p.0. Denoting the inventory level, (positive) inventory, and backorder averages by $\bar{Y}_n$, $\bar{W}_n$, $\bar{W}_n^+$, $\bar{W}_n^-$, we have

$$\left(\frac{\partial\bar{W}_n}{\partial s}\right)_{IPA}=1, \qquad (11)$$

$$\left(\frac{\partial\bar{W}_n^+}{\partial s}\right)_{IPA}=\frac{1}{n}\sum_{i=1}^{n}I\{W_i>0\}, \qquad (12)$$

$$\left(\frac{\partial\bar{W}_n^-}{\partial s}\right)_{IPA}=-\frac{1}{n}\sum_{i=1}^{n}I\{W_i<0\}, \qquad (13)$$

where $I\{*\}$ denotes the indicator function of the set.

We have taken $\theta=s$ above, because as we will prove later, the IPA estimator alone is correct for $\theta=s$ but not for $\theta=q$, where an additional SPA term must be added. Technically, the important condition that must be satisfied for IPA to be correct is that the sample performance be continuous on all of $\theta\in(0,\infty)$ w.p.1 (see, e.g., Glasserman 1991a, b). If (1) is used to rewrite Y_n as a function of Y_0, s, and q, the continuity with respect to s and the points of discontinuity with respect to q become clear. Here, we give an intuitive means to determine when IPA will suffice, intended as a useful heuristic for determining the applicability of the technique for more complicated systems.

If q is fixed, then a perturbation in s simply shifts both S and the entire sample path by that same perturbation, *resulting in no change in ordering decisions*, i.e., the periods in which an order is placed and those in which an order is not placed remain exactly the same after the perturbation. This is illustrated in Figure 2, where quantities for (s, S) are indicated by solid lines and quantities for $(s+\Delta s, S+\Delta s)$ are indicated by dashed lines. Consider the $(j-1)$th period. The distance between $Y_{j-1}(s+\Delta s)-D_{j-1}$ in the upper dashed sample path and its corresponding reorder point $s+\Delta s$ is the same as that between $Y_{j-1}(s)-D_{j-1}$ in the lower solid sample path and its corresponding reorder point s, because both of the former are shifted upwards by the same amount Δs. Thus, an infinitesimal change in s, with q fixed, causes only an infinitesimal change in the sample path. Such effects are calculated by IPA.

However, if a perturbation causes a change in the ordering decision in a period, the resulting shifts in the sample path would no longer be infinitesimal only. For the case when s is fixed and q is perturbed, the sample path is again shifted by the amount of the perturbation, but because s is fixed, an inventory position close to s before the perturbation could cause a change from order to not order, or vice versa, (depending on the sign of the change in q). This possibility is shown in Figure 3 for the case $\Delta q>0$, where $Y_{j-1}(q+\Delta q)-D_{j-1}$ in the upper dashed sample path is above its reorder point s (which is *unchanged* by the perturbation, because s is fixed), whereas $Y_{j-1}(q)-D_{j-1}$ in the lower solid sample path is below the reorder point. Thus, the perturbation Δq caused a

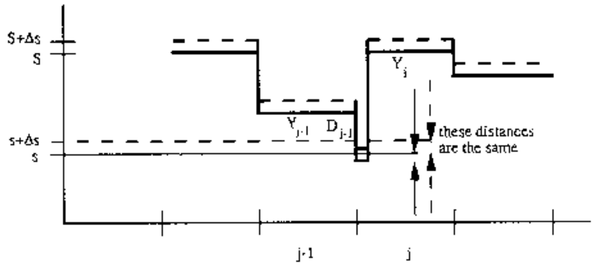

Figure 2. Effect on sample path with q fixed and s perturbed.

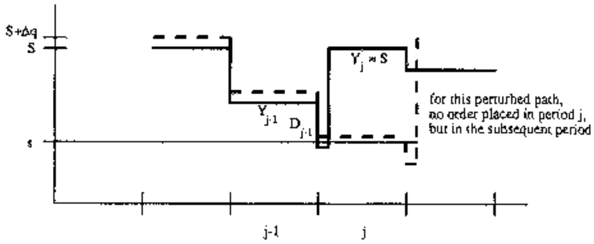

Figure 3. Effect on sample path with s fixed and q perturbed.

change in the ordering decision in period j, and it is clear from this period onwards that two paths are very different. The probability of this change in ordering decision occurring due to the perturbation Δq is $O(\Delta q)$ and the effect is finite, i.e., $O(1)$, so this is a first-order contribution which must be estimated. We will use conditional expectation to estimate this contribution.

1.3. The SPA Contribution

We again point out that the derivations done here are conducted in an informal manner, more for the benefit of insight into the thought process than for rigor; however, once we have the final estimators we will formally verify their correctness. The overall thought process is as follows: Given a *nominal* (the original) sample path at value q, we introduce a Δq perturbation to construct a sample path for $q + \Delta q$, called the *perturbed* sample path. Conditioned on some sample path quantities, we calculate a conditional expectation on the change in g_n, and then take the limit $\Delta q \to 0$. Note that since we take the limit in the end, no Δq is ever actually introduced into the sample path, the "value" of Δq being merely a useful tool for the derivation. In our derivation, we will take $\Delta q > 0$, so we are computing a right-hand derivative. Infinitesimal changes, i.e., those of $O(\Delta q)$, are computed via IPA. We use SPA to calculate any discrete changes in g_n of $O(1)$ which occur with a probability of $O(\Delta q)$. For $\Delta q > 0$, the only discrete changes that may occur are changes in the ordering decision in some periods from "order" in the nominal sample path to "not order" in the perturbed sample path. We will call such a change an *ordering change*. A *potential* ordering change occurs whenever a demand in the nominal sample path brings the inventory position to a point below s, e.g., in Figure 3, the demand in period $j - 1$ causes an order to be placed in period j. Thus, an ordering change may potentially occur for those periods j such that $Y_{j-1} - D_{j-1} < s$, i.e., in the periods where $Y_j = S$ (because the probability of zero demand in a period is zero), defined by the set:

$$M^*(n) = \{j \leq n\colon Y_j = s\}. \tag{14}$$

Whether or not an ordering change actually occurs depends on the size of Δq. In the case shown in Figure 3, Δq was large enough to cause an ordering change in the perturbed path in period j, i.e., $Y_{j-1}(q) - D_{j-1} < s$ and $Y_{j-1}(q + \Delta q) - D_{j-1} > s$.

Note that $\{Y_j\}$ is a regenerative process with its set of regenerative points given by $M^*(n)$. In our analysis, we will be considering the shifted process $Z_j = Y_{j-1} - s$, which is also regenerative. (Offsetting the index by 1 in the definition is merely for notational convenience, so that we can use Z_j, $j \in M^*(n)$, instead of having to use Z_{j-1} if the index were not offset.) We define one more sample path quantity, the random variable $\alpha_j = D_{j-1} - Z_j$.

Henceforth, we consider only $j \in M^*(n)$ and concentrate on two periods $j - 1$ and j. The quantities of interest are depicted in Figure 4. Conditioned on $Z_j = z_j$, the condition $j \in M^*(n)$ is equivalent to the condition $\alpha_j > 0$. Since an ordering change occurs if $\alpha_j \leq \Delta q$, our SPA estimator is expressed as

$$\left(\frac{\partial g_n}{\partial q}\right)_{SPA} = \sum_{j\in M^*(n)} \lim_{\Delta q\to 0} \frac{E_{z_j}[\Delta g_n | \alpha_j \leq \Delta q] P_{z_j}\{\alpha_j \leq \Delta q\}}{\Delta q} = \sum_{j\in M^*(n)} \lim_{\Delta q\to 0} E_{z_j}[\Delta g_n | \alpha_j \leq \Delta q] \cdot \lim_{\Delta q\to 0} \frac{P_{z_j}\{\alpha_j \leq \Delta q\}}{\Delta q}, \tag{15}$$

where E_{z_j} represents the conditional expectation given $Z_j = z_j$ and P_{z_j} represents the conditional probability given $Z_j = z_j$. Intuitively, the two terms in the summation represent the effect on g_n if an ordering change in period $j \in M^*(n)$ occurs and the probability of such an order change occurring in period $j \in M^*(n)$. The

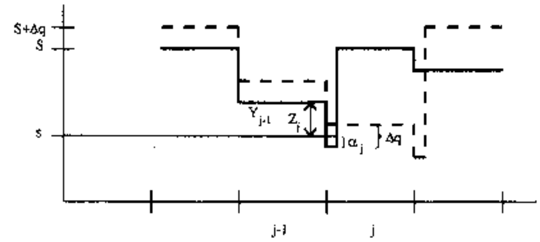

Figure 4. Sample path quantities for calculating conditional expectation.

former term is usually the more difficult term to handle, while the latter term can always be estimated explicitly from the original sample path if the demand distribution is known. However, even if the demand distribution is unknown, the form of the estimator for our system is such that it can still be estimated approximately from the sample path. This is in contrast to previous applications of SPA to queueing systems, e.g., the $GI/G/1$ queue, where no such nice decomposition results. In the rest of the derivation, we will often drop the subscripts for convenience. Due to the regenerative property, α_j and Z_j are i.i.d. for all $j \in M^*(n)$, corresponding to the values in the last period of each regenerative cycle. First, we calculate the conditional distribution for α:

$$P_z\{\alpha \leq x | D > z\} = P_z\{D - z \leq x | D > z\}$$

$$= \frac{P_z\{z < D \leq z + x\}}{P_z\{D > z\}}$$

$$= \frac{F(z+x) - F(z)}{1 - F(z)},$$

and so

$$\lim_{\Delta q \to 0} \frac{P_z\{\alpha \leq \Delta q\}}{\Delta q} = \frac{f(z)}{1 - F(z)}. \tag{16}$$

But if the underlying demand distribution function is unknown (e.g., using real data), we can always estimate via the finite difference approximation

$$\frac{1}{n^*} \frac{\sum_{j \in M^*(n)} I\{\alpha_j \leq \Delta\}}{\Delta} \to \frac{P_z\{\alpha \leq \Delta\}}{\Delta} \quad \text{w.p.1},$$

$$n^* = |M^*(n)|. \tag{17}$$

Note that in this approximate estimation scheme, an appropriate Δ has to be selected, trading off between the bias of a large Δ and the noise of a small Δ.

Now, we derive a way to estimate the term $\lim_{\Delta q \to 0} E_z[\Delta g_n | \alpha \leq \Delta q]$. First, for the nominal sample path, we have by the definition of g_n, (3):

$$g_n(q) = \frac{1}{n} \sum_{i=1}^{n} g(W_i), \tag{18}$$

where for notational ease we have omitted the argument s from g_n. For the perturbed sample path under the condition $\{\alpha \leq \Delta q\}$, we refer to Figure 4 to see that the perturbation causes an ordering change from order to not order in that period. However, because the inventory position is so (infinitesimally) close to the level s, any finite demand will cause an order to take place in the next period in the perturbed sample path, bringing it to the level S exactly one

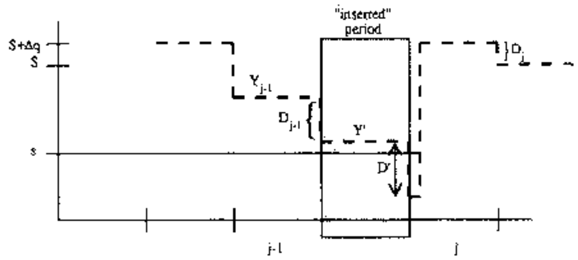

Figure 5. Perturbed sample path constructed from nominal path with inserted period.

period later than in the nominal path. Thus, the perturbed path for inventory position can be constructed from the nominal path with an appropriate extra period "inserted." This is shown in Figure 5, where only the perturbed path is actually depicted: The demand for period $j - 1$ leaves the inventory position of the perturbed path above s, so no order is placed in the subsequent period, whereas in the nominal path, the demand in period $j - 1$ brings the inventory position below s, causing an order up to S in period j (compare with Figure 3 or 4). The demand random variable in the "inserted" period is denoted by D' in Figure 5 and has the same distribution as demands in all other periods, so henceforth we will replace it with the generic demand random variable D. However, for the inventory level process, the effect of the order change is not felt until L periods later, so the "inserted" period is equal in distribution to level $W' = s - D_{(L)} - \alpha + \Delta q$, where $D_{(L)}$ denotes a r.v. representing the sum of L i.i.d. demands with distribution F, i.e., $D_{(L)} \sim F_L$, the L-fold convolution of F with itself.

We also note that the probability of this ordering change taking place occurs only once every regenerative cycle, and from (16) such a probability is $O(\Delta q)$. Since regenerative cycles are independent, the probability of there being $m > 1$ such ordering changes simultaneously is $O((\Delta q)^m)$, and hence can be ignored for the purpose of estimating first derivatives.

If we are interested in only n periods (i.e., the finite horizon model), then the last period in the nominal path must be discarded in the construction of the perturbed path. For the long-run average, such differences can be ignored, i.e., all $n + 1$ periods could be used (because in the limit, the two will be the same). In what follows, we will carry forward only the $n + 1$ period construction, denoting the estimator in the perturbed path construction by $E_z[g_{n+1}(q + \Delta q) | \alpha \leq \Delta q]$. The expected contribution to the performance measure from the inserted period is $E[g(s - D_{(L)} - \alpha + \Delta q)]$, because $W' = s - D_{(L)} - \alpha + \Delta q$ is the level of the inserted period. Then,

$$E_z[g_{n+1}(q + \Delta q) | \alpha \leqslant \Delta q] = \frac{1}{n+1}\left[\sum_{i=1}^{n} g(W_i + \Delta q) + E[g(s - D_{(L)} - \alpha + \Delta q) | \alpha \leqslant \Delta q]\right]. \tag{19}$$

Now taking the difference between the perturbed and the nominal paths,

$$\begin{aligned} E_z[\Delta g_n | \alpha \leqslant \Delta q] &= E_z[g_{n+1}(q + \Delta q) - g_n(q) | \alpha \leqslant \Delta q] \\ &= \frac{1}{n+1}\left[E[g(s - D_{(L)} - \alpha + \Delta q) | \alpha \leqslant \Delta q] + \sum_{i=1}^{n}(g(W_i + \Delta q) - g(W_i)) - \frac{\sum_{i=1}^{n} g(W_i + \Delta q)}{n}\right]. \end{aligned}$$

Taking the limit $\Delta q \to 0$, we get

$$\lim_{\Delta q \to 0} E_z[\Delta g_n | \alpha \leqslant \Delta q] = \frac{1}{n+1}\left[E[g(s - D_{(L)})] - \frac{\sum_{i=1}^{n} g(W_i)}{n}\right]. \tag{20}$$

Reinstalling our subscripts into (15), we have

$$\left(\frac{\partial g_n}{\partial q}\right)_{SPA} = \frac{1}{n+1}\sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[g(s - D_{(L)})] - \frac{\sum_{i=1}^{n} g(W_i)}{n}\right]. \tag{21}$$

1.4. The Estimation Algorithm

The final estimator for the derivative with respect to q is the sum of the IPA and SPA parts:

$$\left(\frac{\partial g_n}{\partial q}\right)_{PA} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial g(W_i(\theta), \theta)}{\partial \theta} + g'(W_i)\right] + \frac{1}{n+1}\sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[g(s - D_{(L)})] - \frac{\sum_{i=1}^{n} g(W_i)}{n}\right]. \tag{22}$$

In particular, we have

$$\left(\frac{\partial \bar{W}_n}{\partial q}\right)_{PA} = 1 + \frac{1}{n+1}\sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[s - E[D_{(L)}] - \frac{\sum_{i=1}^{n} W_i}{n}\right], \tag{23}$$

$$\left(\frac{\partial \bar{W}_n^+}{\partial q}\right)_{PA} = \frac{1}{n}\sum_{i=1}^{n} I\{W_i > 0\} + \frac{1}{n+1}\sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[(s - D_{(L)})^+] - \frac{\sum_{i=1}^{n} W_i^+}{n}\right], \tag{24}$$

$$\left(\frac{\partial \bar{W}_n^-}{\partial q}\right)_{PA} = -\frac{1}{n}\sum_{i=1}^{n} I\{W_i < 0\} + \frac{1}{n+1}\sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[(D_{(L)} - s)^+] - \frac{\sum_{i=1}^{n} W_i^-}{n}\right]. \tag{25}$$

An explicit algorithm estimating all of the above derivative estimators is given below.

PA Algorithm for Average Inventory Level, Inventory, and Backorder (Fixed Lead Time *L*)

INITIALIZE:

WSUM = WSUM1 = WSUM2 = WNUM1 = WNUM2 = HAZ = ORDAMT = ORDTME = 0;
$Z = S - s$; $Y = W = S$;

EACH PERIOD:
Generate *demand D* according to *F*;

If $Y - D < s$, Then ORDTME = L;
ORDAMT $= S - (Y - D)$;
HAZ = HAZ + $f(Z)/[1 - F(Z)]$;
$Y = S$;
Else $Y = Y - D$;
If ORDAMT > 0, Then If ORDTME = 0,
Then $W = W$ + ORDAMT; ORDAMT = 0;
Else ORDTME = ORDTME − 1;
$Z = Y - s$; $W = W - D$; WSUM = WSUM + W;
If $W > 0$, Then WSUM1 = WSUM1 + W;
WNUM1 = WNUM1 + 1;
Else WSUM2 = WSUM2 + W;
WNUM2 = WNUM2 + 1;

OUTPUT—At the end of N periods:

$(\partial W/\partial s)_{IPA} = 1$;

$(\partial W/\partial q)_{PA} = 1 + \text{HAZ} * (s - E[D_{(L)}] - \text{WSUM/N})/(\text{N} + 1)$;

$(\partial W^+/\partial s)_{IPA} = \text{WNUM1/N}$;

$(\partial W^+/\partial q)_{PA} = \text{WNUM1/N} + \text{HAZ} * (E[(s - D_{(L)})^+] - \text{WSUM1/N})/(\text{N} + 1)$;

$(\partial W^-/\partial s)_{IPA} = -\text{WNUM2/N}$;

$(\partial W^-/\partial q)_{PA} = -\text{WNUM2/N} + \text{HAZ} * (E[(D_{(L)} - s)^+] - \text{WSUM2/N})/(\text{N} + 1)$.

2. CONSISTENCY OF THE ESTIMATORS

Let Y and W denote the steady-state (or limiting) random variables for Y_n and W_n, respectively. Then, we have (see, e.g., Theorem 11, p. 92, Wolff 1989), $\sum_{i=1}^{n} Y_i/n \to E[Y]$ and $\sum_{i=1}^{n} W_i/n \to E[W]$ w.p.1. Since g is continuous, we also have $g_n \to E[g(W)]$ w.p.1 (assuming that $E[g(W)] < \infty$). Here, we prove that the estimators are strongly consistent, by which we mean that

$$\left(\frac{\partial g_n}{\partial \theta}\right)_{PA} \to \frac{\partial E[g(W)]}{\partial \theta} \quad \text{w.p.1.}$$

For the derivatives with respect to q, we need the help of the following lemma derived from an identity in the theory of point processes.

Lemma 1. *Let $\{D_i\}$ be a renewal process with distribution function $F(\cdot)$, density function $f(\cdot)$, and associated counting process $N(t)$. Define the cumulative process $S_n = \sum_{i=1}^{n} D_i$, the age process $Z(t) = t - S_{N(t)}$, and the renewal function $R(t) = \sum_{n=1}^{\infty} F_n(t)$. Then,*

$$E\left[\frac{f(Z(t))}{1 - F(Z(t))}\right] = r(t),$$

where $r(t) = dR(t)/dt$ is the renewal density.

Proof. Theorem T7 from Bremaud (1981) gives under mild conditions (using Martingale arguments):

$$E\left[\int_0^t C(s)\, dN(s)\right] = E\left[\int_0^t C(s)\lambda(Z(s))\, ds\right], \qquad (26)$$

where $\lambda(x) = f(x)/(1 - F(x))$. If we take $C(y) = 1$ for all y, we get

$$R(t) = E[N(t)] = E\left[\int_0^t \lambda(Z(s))\, ds\right]$$

$$= \int_0^t E[\lambda(Z(s))]\, ds,$$

where we have applied Fubini's Theorem to get the last equality. Differentiating both sides gives the desired result.

We begin with the easier result for average inventory level.

Theorem 1. *The estimators $(\partial \bar{W}_n/\partial s)_{IPA}$ and $(\partial \bar{W}_n/\partial q)_{PA}$ are strongly consistent.*

Proof. Here Y_n is regenerative with state-space $[s, S]$; thus, its stationary version Y has a distribution with an atom at S given by

$$P\{Y = S\} = \frac{1}{1 + R(q)},$$

and a density on its state space given by

$$f_Y(y) = \frac{r(S - y)}{1 + R(q)} \quad y \in [s, S],$$

where $R(x) = \sum_{n=1}^{\infty} F_n(x)$ is the renewal function for the demand sequence, $r(x) = \sum_{n=1}^{\infty} f_n(x) = dR/dx$ is the corresponding renewal density, and $\bar{R}(x) = \int_0^x R(y)\, dy$. Integrating to get the expectation of Y, we have

$$E[Y] = \frac{S}{1 + R(q)} + \frac{sR(q) + \bar{R}(q)}{1 + R(q)}$$

$$= s + \frac{q + \bar{R}(q)}{1 + R(q)}. \qquad (27)$$

The stationary version of (2) is given by

$$W = Y - D_{(L)}, \qquad (28)$$

from which we immediately have

$$E[W] = s + \frac{q + \bar{R}(q)}{1 + R(q)} - L \cdot E[D]. \qquad (29)$$

Differentiating $E[W]$ with respect to s, we get $\partial EW/\partial s = 1$, which trivially proves consistency for the derivative with respect to s.

Differentiating $E[W]$ with respect to q, we get

$$\frac{\partial EW}{\partial q} = 1 - \frac{r(q)[q + \bar{R}(q)]}{[1 + R(q)]^2}.$$

Noting that the Z_j for $j \in M^*(n)$ are i.i.d., because they occur once in every regenerative cycle, we let Z denote the generic r.v. We have w.p.1 that

$$\left(\frac{\partial \bar{W}_n}{\partial q}\right)_{PA} \to 1 + P\{Y = S\}E\left[\frac{f(Z)}{1 - F(Z)}\right]$$

$$\cdot [s - (L + 1)E[D] - E[W]]$$

$$= 1 - P\{Y = S\}E\left[\frac{f(Z)}{1 - F(Z)}\right]\left[\frac{q + \bar{R}(q)}{1 + R(q)}\right]$$

$$= 1 - E\left[\frac{f(Z)}{1 - F(Z)}\right]\left[\frac{q + \bar{R}(q)}{(1 + R(q))^2}\right],$$

by substituting for $E[W]$ from (29) in the second line and using the fact that $P\{Y = S\} = 1/(1 + R(q))$ in the third line. Via Lemma 1, we get the desired result, because

$$E\left[\frac{f(Z)}{1-F(Z)}\right] = r(q).$$

Next, we have the following theorem.

Theorem 2. *The estimators* $(\partial \bar{W}_n^+/\partial s)_{IPA}$, $(\partial \bar{W}_n^+/\partial q)_{PA}$, $(\partial \bar{W}_n^-/\partial s)_{IPA}$, *and* $(\partial \bar{W}_n^-/\partial q)_{PA}$ *are strongly consistent.*

Proof. The proofs for W^+ and W^- are a little more complicated. We have w.p.1 that

$$\left(\frac{\partial \bar{W}_n^+}{\partial s}\right)_{IPA} \to P\{W > 0\},$$

$$\left(\frac{\partial \bar{W}_n^+}{\partial q}\right)_{PA} \to P\{W > 0\} + P\{Y = S\}E\left[\frac{f(Z)}{1-F(Z)}\right] \cdot [E[(s - D_{(L)})^+] - E[W^+]],$$

$$\left(\frac{\partial \bar{W}_n^-}{\partial s}\right)_{IPA} \to -P\{W < 0\},$$

$$\left(\frac{\partial \bar{W}_n^-}{\partial q}\right)_{PA} \to -P\{W < 0\} + P\{Y = S\}E\left[\frac{f(Z)}{1-F(Z)}\right] \cdot [E[(D_{(L)} - s)^+] - E[W^-]].$$

We prove strong consistency for each of these four estimators.

Convolving via (28), we get the stationary density for W:

$$f_W(x) = \frac{1}{1+R(q)} \cdot \left\{ f_L(S-x) + \begin{matrix} \int_x^S r(S-u) f_L(u-x)\,du \text{ if } s \le x \le S \\ \int_s^S r(S-u) f_L(u-x)\,du \text{ if } x \le s \end{matrix} \right\}. \quad (30)$$

For the first two estimators, we see that $P\{W > 0\}$ is given by

$$P\{W>0\} = \frac{1}{1+R(q)}\left[F_L(s+q) + \int_s^S r(S-u)F_L(u)\,du\right] = \frac{1}{1+R(q)}\left[F_L(s+q) + F_L(s)R(q) + \int_0^q R(u) f_L(s+q-u)\,du\right]. \quad (31)$$

On the other hand, $E[W^+]$ is given by

$$E[W^+] = \frac{1}{1+R(q)}\left[\bar{F}_L(s+q) + \bar{F}_L(s)R(q) + \int_s^S R(S-u)F_L(u)\,du\right] = \frac{1}{1+R(q)}\left[\bar{F}_L(s+q) + \bar{F}_L(s)R(q) + \int_0^q R(u)F_L(s+q-u)\,du\right], \quad (32)$$

where $\bar{F}_L(x) = \int_0^x F_L(x)\,dx$. Differentiating with respect to s, we have

$$\frac{\partial E[W^+]}{\partial s} = \frac{1}{1+R(q)}\left[F_L(s+q) + F_L(s)R(q) + \int_0^q R(u) f_L(s+q-u)\,du\right] = P\{W>0\}, \quad (33)$$

completing the proof for the first estimator.

Differentiating with respect to q, we have

$$\frac{\partial E[W^+]}{\partial q} = \frac{1}{1+R(q)}\left[F_L(s+q) + \bar{F}_L(s)r(q) + F_L(s)R(q) + \int_0^q R(u) f_L(s+q-u)\,du - r(q)E[W^+]\right] = P\{W>0\} + P\{Y=S\}E\left[\frac{f(Z)}{1-F(Z)}\right] \cdot [E[(s-D_{(L)})^+] - E[W^+]], \quad (34)$$

where again Lemma 1 has been applied, completing the proof for the second estimator.

For the last two estimators, we use the fact that $E[W^+] - E[W^-] = E[W]$. We have

$$\frac{\partial E[W^-]}{\partial s} = \frac{\partial E[W^+]}{\partial s} - \frac{\partial E[W]}{\partial s} = P\{W>0\} - 1 = -P\{W<0\}, \quad (35)$$

completing the proof for the third estimator.

Finally, we have

$$\frac{\partial E[W^-]}{\partial q} = \frac{\partial E[W^+]}{\partial q} - \frac{\partial E[W]}{\partial q}$$

$$= P\{W > 0\} + \frac{r(q)}{1 + R(q)}$$

$$\cdot \{[E[(s - D_{(L)})^+] - E[W^+]]\}$$

$$- \left\{1 - \frac{r(q)[q + \bar{R}(q)]}{[1 + R(q)]^2}\right\}$$

$$= -P\{W < 0\} + \frac{r(q)}{1 + R(q)}$$

$$\cdot [E[(s - D_{(L)})^+] - E[W^+]$$

$$+ E[D_{(L)} - s] + E[W]]$$

$$= -P\{W < 0\} + \frac{r(q)}{1 + R(q)}$$

$$\cdot [E[(D_{(L)} - s)^+] - E[W^-]], \qquad (36)$$

and once more making use of Lemma 1, the proof for the last estimator is complete, as is the proof of the theorem.

Sample path proofs for these theorems can also be carried out for the IPA estimators using techniques similar to those in Glasserman (1991) or Hu (1992), and for the SPA estimators using the techniques of Fu and Hu (1992).

3. AVERAGE COST PER PERIOD

A cost function involving inventory, backorder, and ordering costs is a function of both the inventory position and the inventory level, and hence is a bit more complicated. Generalizing to such a function is quite straightforward for the IPA part, but must be handled on a case-by-case basis for the SPA part, depending both on the nature of the inventory level, which incorporates some kind of order lead time mechanism, and on the form of the function. Here, we consider the long-run average cost consisting of three components: ordering, holding, and shortage costs. We define the one-period cost function by

$$J(w, y) = H(w^+) + B(w^-) + K \cdot I\{\hat{y} - \hat{d} < s\} \qquad (37)$$

$$= H(w^+) + B(w^-) + K \cdot I\{y = S\}, \qquad (38)$$

where

$H(x)$ = the holding costs per period for inventory x;
$B(x)$ = the shortage costs per period for backorder x;
K = the setup cost for placing an order;
w = the inventory level in the period;
y = the inventory position in the period;
$\hat{y}$ = the inventory position in the previous period;
$\hat{d}$ = the demand in the previous period.

We have assumed a fixed setup cost with general, continuous holding and shortage cost functions. Since we will be interested in the long-run average cost, we have omitted the per-unit variable ordering cost without loss of generality. It is well known, but not our primary concern, that additional conditions on form of these cost functions are needed to be able to assert that a policy of the (s, S) type is optimal, e.g., convexity (Scarf 1960, Iglehart 1963, see also, Veinott 1966). Our performance measure of interest is:

$$J_n(s, q) = \frac{1}{n} \sum_{i=1}^{n} J(W_i, Y_i).$$

3.1. Derivation of the Estimators

Since $J(w, y)$ is piecewise differentiable w.p.1 with respect to both its arguments (the discontinuities in its derivatives at $w = 0$ and $y = s$ occur w.p.0 because demand is continuous), we apply the chain rule and the results of the previous section, (5) and (6), to get

$$\frac{dJ(W_n(\theta), Y_n(\theta))}{d\theta}$$

$$= \frac{\partial J(W_n(\theta), Y_n(\theta), \theta)}{\partial\theta} + \frac{\partial J}{\partial W_n}\frac{dW_n}{d\theta} + \frac{\partial J}{\partial Y_n}\frac{dY_n}{d\theta}$$

$$= \frac{\partial J(W_n(\theta), Y_n(\theta), \theta)}{\partial\theta} + \frac{\partial J}{\partial W_n} + \frac{\partial J}{\partial Y_n} \quad \text{w.p.1.}$$

From (38), we get

$$\frac{\partial J(W_n(\theta), Y_n(\theta), \theta)}{\partial\theta} = 0,$$

$$\frac{\partial J(W_n, Y_n)}{\partial Y_n} = 0,$$

$$\frac{\partial J(W_n, Y_n)}{\partial W_n} = \begin{cases} H'(W_n) & \text{if } W_n > 0 \\ -B'(-W_n) & \text{if } W_n < 0 \end{cases},$$

and the IPA part of the derivative estimator is quite simple:

$$\frac{dJ(W_n, Y_n)}{d\theta} = \begin{cases} H'(w_n) & \text{if } W_n > 0 \\ -B'(-W_n) & \text{if } W_n < 0 \end{cases}. \qquad (39)$$

Again, for the derivative estimator with respect to s (q fixed), the IPA part will suffice, so over n periods, our estimator is given by:

$$\left(\frac{\partial J_n}{\partial s}\right)_{IPA} = \frac{1}{n}\left[\sum_{i:W_i>0} H'(W_i) - \sum_{i:W_i<0} B'(-W_i)\right]. \qquad (40)$$

For the same reasons discussed in the previous section, the derivative with respect to q must include an additional term which we calculate via conditional expectation and call the SPA term. This term will depend on the nature of the order lead times. The methodology is essentially identical to the derivation in the previous section, i.e., the use of conditional expectation results in the construction of a perturbed sample path with an "inserted" period, the only difference being that the calculation of the additional contribution to the long-run average cost is slightly more difficult.

We wish to calculate $\lim_{\Delta q \to 0} E_z[\Delta J_n | \alpha \le \Delta q]$. For the nominal sample path, we have

$$J_n(q) = \frac{1}{n} \sum_{i=1}^{n} J(W_i, Y_i), \tag{41}$$

where for notational ease we have omitted the argument s. As before, for the perturbed sample path under condition $\{\alpha \le \Delta q\}$, Figure 4 demonstrates the perturbation that causes an ordering change from order to not order in period j. Again, because the inventory position is so (infinitesimally) close to the level s, any finite demand will cause an order to take place in the subsequent period in the perturbed sample path, bringing it to the level S one period later than in the nominal path. Thus, as before, the perturbed path for inventory position can be constructed from the nominal path with an extra period "inserted" before period j, as shown in Figure 5. This construction leads to a perturbed path for inventory level constructed from the nominal path by inserting an extra period L periods later, the level of the inserted period being $s - D_{(L)} - \alpha - \Delta q$.

Thus, the ordering, being a function of inventory position, is delayed by one more period in the perturbed path, and the amount ordered is increased by one extra period of demand. The inventory level of the inserted period occurs L periods later in the perturbed path and is of size $s - D_{(L)} - \alpha - \Delta q$, so calculating the expected holding or shortage cost incurred in the inserted period gives

$$E_z[J_{n+1}(q + \Delta q) | \alpha \le \Delta q] = \frac{1}{n+1} \left[\sum_{i=1}^{n} J(W_i, Y_i) + E[H([s - D_{(L)}]^+)] + E[B([D_{(L)} - s]^+)] \right], \tag{42}$$

where we have chosen to omit the $O(\Delta q)$ terms which go to 0 in the limit as $\Delta q \to 0$. Again, we include only the $n + 1$ period case. Parallel to the procedure in the previous section, but eschewing the straightforward but tedious details here, after taking the difference between the perturbed and the nominal paths and taking the limit $\Delta q \to 0$, we get

$$\lim_{\Delta q \to 0} E_z[\Delta J_n | \alpha \le \Delta q] = \frac{1}{n+1} \left[E[H([s - D_{(L)}]^+)] + E[B([D_{(L)} - s]^+)] - \frac{\sum_{i=1}^{n} J(W_i, Y_i)}{n} \right]. \tag{43}$$

So we have

$$\left(\frac{\partial J_n}{\partial q}\right)_{SPA} = \frac{1}{n+1} \sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[H([s - D_{(L)}]^+)] + E[B([D_{(L)} - s]^+)] - \frac{\sum_{i=1}^{n} J(W_i, Y_i)}{n} \right]. \tag{44}$$

Although K does not appear explicitly in (44), there is dependence on K through the average cost term $\sum_{i=1}^{n} J(W_i, Y_i)/n$, where, for example, an increase in q (with s held constant) would cause the number of orders to decrease, and hence the average per period cost of ordering would decrease.

3.2. The Estimation Algorithm

Again, the final estimator is the sum of the IPA and SPA parts:

$$\left(\frac{\partial J_n}{\partial q}\right)_{PA} = \frac{1}{n} \left[\sum_{i: W_i > 0} H'(W_i) + \sum_{i: W_i < 0} B'(-W_i) \right] + \frac{1}{n+1} \sum_{j \in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[E[H([s - D_{(L)}]^+)] + E[B([D_{(L)} - s]^+)] - \frac{\sum_{i=1}^{n} J(W_i, Y_i)}{n} \right]. \tag{45}$$

An explicit algorithm for both derivatives is given below.

PA Algorithm for $\partial J_n/\partial s$ and $\partial J_n/\partial q$ (Fixed Lead Time L)

INITIALIZE:
JSUM = HSUM = BSUM = HAZ = ORDAMT = ORDTME = 0;
$Z = S - s$; $Y = W = S$;

EACH PERIOD:
Generate demand D according to F;

If $Y - D < s$, Then ORDTME = L;
ORDAMT = $S - (Y - D)$;
HAZ = HAZ + $f(Z)/[1 - F(Z)]$;
JSUM = JSUM + K; $Y = S$;
Else $Y = Y - D$;
If ORDAMT > 0, Then If ORDTME = 0,
Then $W = W$ + ORDAMT; ORDAMT = 0;
Else ORDTME = ORDTME − 1;
$Z = Y - s$; $W = W - D$;
If $W > 0$, Then JSUM = JSUM + $H(W)$;
HSUM = HSUM + $H'(W)$;
Else JSUM = JSUM + $B(W)$;
BSUM = BSUM − $B'(-W)$;

OUTPUT: At the end of N periods:

$(\partial J_n/\partial s)_{IPA}$ = (HSUM + BSUM)/N;

$(\partial J_n/\partial q)_{PA}$ = (HSUM + BSUM)/N + HAZ

$*(H(E[(s - D_{(L)})^+])$

$+ B(E[(D_{(L)} - s)^+]) -$ JSUM/N)/(N+1).

4. CONSISTENCY PROOFS

Using analytic results, we prove consistency for the infinite horizon model, with linear holding and shortage costs. Defining the holding and shortage costs by

h = the per-unit holding cost (per period);
p = the per-unit shortage cost (per period);

i.e., $H(x) = hx$ and $B(x) = px$, the estimators given by (45) and (40) become

$$\left(\frac{\partial J_n}{\partial s}\right)_{IPA} = \frac{1}{n}\left[\sum_{i:W_i>0} h - \sum_{i:W_i<0} p\right], \tag{46}$$

$$\left(\frac{\partial J_n}{\partial q}\right)_{PA} = \frac{1}{n}\left[\sum_{i:W_i>0} h - \sum_{i:W_i<0} p\right] + \frac{1}{n+1}\sum_{j\in M^*(n)} \frac{f(Z_j)}{1 - F(Z_j)} \cdot \left[hE[(s - D_{(L)})^+] + pE[(D_{(L)} - s)^+] - \frac{\sum_{i=1}^{n} J(W_i, Y_i)}{n}\right], \tag{47}$$

and we have the following theorem.

Theorem 3. *The estimators $\partial J_n/\partial s$ and $\partial J_n/\partial s$ are strongly consistent.*

Proof. Let $\mathcal{J}$ denote the long-run average cost per period, which exists because W and Y exist and $J(\cdot,\cdot)$ has a countable number of discontinuities. We show that the estimators converge to the appropriate derivative of $\mathcal{J}$. We have w.p.1 that

$$\left(\frac{\partial J_n}{\partial s}\right)_{IPA} \to hP\{W > 0\} - pP\{W < 0\}.$$

The long-run average cost per period is given by

$$\mathcal{J}(s, q) = \frac{K}{1 + R(q)} + hE[W^+] + pE[W^-]. \tag{48}$$

Differentiating $\mathcal{J}$ with respect to s, we have

$$\frac{\partial \mathcal{J}}{\partial s} = h\frac{\partial E[W^+]}{\partial s} + p\frac{\partial E[W^-]}{\partial s} = hP\{W > 0\} - pP\{W < 0\}, \tag{49}$$

the second equality following from Theorem 2, completing the proof for the derivative with respect to s.

For the second estimator, we have w.p.1 that

$$\left(\frac{\partial J_n}{\partial q}\right)_{PA} \to hP\{W > 0\} - pP\{W < 0\} + P\{Y = S\}E\left[\frac{f(Z)}{1 - F(Z)}\right] \cdot [hE[(s - D_{(L)})^+] + pE[(D_{(L)} - s)^+] - \mathcal{J}].$$

Differentiating $\mathcal{J}$ with respect to q, we have

$$\frac{\partial \mathcal{J}}{\partial q} = -\frac{K}{1 + R(q)}\frac{r(q)}{1 + R(q)} + h\frac{\partial E[W^+]}{\partial q} + p\frac{\partial E[W^-]}{\partial q} \tag{50}$$

$$= -\frac{K}{1 + R(q)}\frac{r(q)}{1 + R(q)} + h\left[P\{W > 0\} + \frac{r(q)}{1 + R(q)} \cdot (E[(s - D_{(L)})^+] - E[W^+])\right] \tag{51}$$

$$+ p\left[-P\{W < 0\} + \frac{r(q)}{1 + R(q)} \cdot (E[(D_{(L)} - s)^+] - E[W^-])\right] \tag{52}$$

$$= hP\{W > 0\} - pP\{W < 0\} + \frac{r(q)}{1 + R(q)}$$

$$\cdot \Big[hE[(s - D_{(L)})^+] + p(E[(D_{(L)} - s)^+] \quad (53)$$

$$- \left(hE[W^+] + pE[W^-] + \frac{K}{1 + R(q)} \right) \Big] \quad (54)$$

$$= hP\{W > 0\} - pP\{W < 0\} + \frac{r(q)}{1 + R(q)}$$

$$\cdot [hE[(s - D_{(L)})^+] + p(E[(D_{(L)} - s)^+] - \mathscr{J}]. \quad (55)$$

The proof again follows with the application of Lemma 1.

Note that although the proofs are given for the special case of linear holding and shortage costs with fixed lead times, the derivation of the estimators was not based on these assumptions. Thus, we believe that the applicability of the estimators is far wider than the conditions under which consistency has been proved. For fixed lead times, analytical proofs could be provided for other forms of the holding and shortage cost functions, because the stationary distributions of W and Y are available. However, such proofs would simply be messy case-by-case exercises in integration of the form

$$\mathscr{J} = \frac{K}{1 + R(q)} + \int_0^S H(y) f_W(y)\, dy$$

$$+ \int_{-\infty}^{0} B(-y) f_W(y)\, dy, \quad (56)$$

which is why sample path proofs are preferable in establishing consistency of the estimators.

5. SUMMARY

In this paper, we have derived sample path derivative estimators for the average inventory, backorder, and cost per period in an (s, S) periodic review inventory model. The derivatives are taken with respect to the parameters s and $q = S - s$, which, primarily because they are *not* parameters of probability distributions, lead to estimators that do not involve derivatives of random variables, in contrast to most perturbation analysis algorithms. There are few papers in the perturbation analysis literature that address structural parameters, as opposed to parameters in the probability distributions of timing random variables; this is the first paper to apply the smoothed perturbation analysis technique. For certain special cases, we have given consistency proofs.

The basic motivation behind this work was to introduce the idea of sample path derivative estimators to inventory systems. Previous work in the area of sample path derivative estimation has focused on queueing systems. The results seem to indicate that using sample path derivative estimators for inventory systems has as much potential as the vast work done for queueing systems.

ACKNOWLEDGMENT

The author thanks Professor Jian-Qiang Hu for many useful discussions and Professor Paul Glasserman for the proof of Lemma 1. The author is also grateful to the anonymous referees for their detailed comments that resulted in an improved presentation of this paper.

REFERENCES

ARROW, K. J., T. HARRIS AND J. MARSCHAK. 1951. Optimal Inventory Policy. *Econometrica* **XIX,** 250–272.

BREMAUD, P. 1981. *Point Processes and Queues.* Springer-Verlag, New York.

FEDERGRUEN, A., AND P. ZIPKIN. 1984. An Efficient Algorithm for Computing Optimal (s, S) Policies. *Opns. Res.* **32,** 1268–1285.

FU, M. C., AND J. Q. HU. 1992. Extensions and Generalizations of Smoothed Perturbation Analysis in a Generalized Semi-Markov Process Framework. *IEEE Trans. Auto. Control* **37,** 1483–1500.

FU, M. C., AND J. Q. HU. 1994. (s, S) Inventory Systems With Random Lead Times: Harris Recurrence and Its Implications in Sensitivity Analysis. *Prob. Eng. and Info. Sci.* (to appear).

GLASSERMAN, P. 1991a. Structural Conditions for Perturbation Analysis Derivative Estimates: Finite-Time Performance Indices. *Opns. Res.* **39,** 724–738.

GLASSERMAN, P. 1991b. *Gradient Estimation Via Perturbation Analysis.* Kluwer Academic, Amsterdam.

GLASSERMAN, P., AND W. B. GONG. 1990. Smoothed Perturbation Analysis for a Class of Discrete Event Systems. *IEEE Trans. Auto. Control* **35,** 1218–1230.

GONG, W. B., AND Y. C. HO. 1987. Smoothed Perturbation Analysis of Discrete-Event Dynamic Systems. *IEEE Trans. Auto. Control* **AC-32,** 858–867.

HO, Y. C., AND X. R. CAO. 1983. Optimization and Perturbation Analysis of Queueing Networks. *J. Optim. Theory and Applic.* **40,** 559–582.

HO, Y. C., AND X. R. CAO. 1991. *Discrete Event Dynamic Systems and Perturbation Analysis.* Kluwer Academic, Amsterdam.

Hu, J. Q. 1992. Convexity of Sample Path Performances and Strong Consistency of Infinitesimal Perturbation Analysis Estimates. *IEEE Trans. Auto. Control* **37**, 258–262.

Iglehart, D. 1963. Optimality of (s, S) Policies in the Infinite Horizon Dynamic Inventory Problem. *Mgmt. Sci.* **9**, 259–267.

L'Ecuyer, P. 1990. A Unified View of the IPA, SF, and LR Gradient Estimation Techniques. *Mgmt. Sci.* **36**, 1364–1383.

Porteus, E. L. 1985. Numerical Comparisons of Inventory Policies for Periodic Review Systems. *Opns. Res.* **33**, 134–152.

Sahin, I. 1982. On the Objective Function Behavior in (s, S) Inventory Models. *Opns. Res.* **30**, 709–724.

Scarf, H. 1960. The Optimality of (S, s) Policies in the Dynamic Inventory Problem. *Mathematical Methods in the Social Sciences.* Stanford University Press, Stanford, California.

Suri, R., and M. Zazanis. 1988. Perturbation Analysis Gives Strongly Consistent Sensitivity Estimates for the $M/G/1$ Queue. *Mgmt. Sci.* **34**, 39–64.

Veinott, A. F., Jr. 1966. On the Optimality of (s, S) Inventory Policies: New Conditions and a New Proof. *SIAM J. Appl. Math.* **14**, 1067–1083.

Veinott, A. F., Jr., and H. M. Wagner. 1965. Computing Optimal (s, S) Inventory Policies. *Mgmt. Sci.* **11**, 525–552.

Wolff, R. W. 1989. *Stochastic Modeling and the Theory of Queues.* Prentice Hall, Englewood Cliffs, New Jersey.

Zheng, Y. S., and A. Federgruen. 1991. Finding Optimal (s, S) Policies Is About as Simple as Evaluating a Single Policy. *Opns. Res.* **39**, 654–665.