Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

ON THE REGULARITY CONDITIONS AND APPLICATIONS FOR GENERALIZED LIKELIHOOD RATIO METHOD

Yijie Peng

School of Management
Fudan University
Shanghai, 200433, CHINA

Michael C. Fu

Robert H. Smith School of Business
Institute for Systems Research
University of Maryland
College Park, MD 20742, USA

Jianqiang Hu

School of Management
Fudan University
Shanghai, 200433, CHINA

ABSTRACT

We compare different sets of regularity conditions required to derive a generalized likelihood ratio method (GLRM) proposed by Peng et al. (2016a), and present additional applications of GLRM. A numerical experiment substantiates that the GLRM can address a broad set of sensitivity estimation problems in a unified framework.

1 INTRODUCTION

Stochastic derivative estimation is an important topic in simulation optimization, because it plays a central role in both sensitivity analysis and gradient-based optimization. See Fu (2006, 2008, 2015) for general overviews. Likelihood ratio method (LRM) is one of the most popular stochastic derivative estimation techniques, providing an unbiased estimator that does not require continuity of the payoff. Please refer to Glynn (1990) and Rubinstein and Shapiro (1993) for details.

For a discontinuous payoff function dependent on a structural parameter, push-out LRM, which moves the parameter out of the payoff and into a distribution, can be a potential choice. See Rubinstein (1992), Rubinstein (1997), and Pflug and Rubinstein (2002) for applications in different problems, and Rubinstein and Shapiro (1993) as a review. However, the push-out technique basically applies a change of variables and requires an explicit inverse to be implemented efficiently. Unfortunately, this requirement is not always satisfied in application, which is shown in this work and Peng et al. (2016a).

Peng et al. (2016a) put the push-out LRM in a general framework, and show the generalized LRM (GLRM) proposed in Peng et al. (2016a) is a generalization of the push-out LRM when the inverse of the transformation exists but is not available in closed form. Peng et al. (2016a) derive GLRM in two different ways under different sets of regularity conditions. One derivation is by function smoothing and integration by parts; and another one is through change of variables and differentiation of implicit function. In this work, we compare regularity conditions required for the two derivations and provide additional applications of GLRM.

Other methods closely related to LRM include the combined infinitesimal perturbation analysis (IPA)-LRM of L'Ecuyer (1990) and the support independent unified likelihood ratio and infinitesimal perturbation analysis (SLRIPA) in Wang et al. (2012). The former puts the IPA and LRM in a unified framework,

978-1-5090-4486-3/16/$31.00

while the latter can treat discontinuity in both sample performance and distribution function by using a combination of IPA and LRM. Peng et al. (2016a) show that the framework where GLRM applies is a generalization of the framework of both combined IPA-LRM and SLRIPA, although the actual estimators are different. Chen and Glasserman (2007) provides a LRM-based estimator for Greeks of financial derivatives with underlying asset prices governed by a discretized diffusion process, and proves that the proposed estimator converges to a Malliavin calculus estimator (Fournié et al. 1999) as discretiziation size goes to zero. Peng et al. (2016a) show the estimator in Chen and Glasserman (2007) is a special case of GLRM.

The rest of the paper is organized as follows. Section 2 introduces some background on LRM and push-out LRM. GLRM is introduced in Section 3. Applications of GLRM can be found in Section 4. A numerical experiment is provided in Section 5. The last section offers conclusions.

2 BACKGROUND

The problem is to estimate the derivative $\frac{\partial}{\partial\theta}E[V(X;\theta)]$ by simulation output, where $V(\cdot;\cdot):\ \mathbb{R}^n\times\Theta\rightarrow\mathbb{R}$, and

$$\frac{\partial}{\partial\theta}E[V(X;\theta)]=\frac{\partial}{\partial\theta}\int\int_{\mathbb{R}^n}V(x;\theta)\ f(x;\theta)\ dx,$$

where $x=(x_1,..,x_n)$ and $f(x;\theta)$ is the joint density of $X=(X_1,..,X_n)$. With appropriate integrability and smoothness conditions, the differentiation and expectation can be exchanged, and we have the LRM estimator as follows:

$$\begin{aligned}\frac{\partial}{\partial\theta}E[V(X;\theta)]&=\int\int_{\mathbb{R}^n}\left(\frac{\partial V(x;\theta)}{\partial\theta}\ f(x;\theta)+V(x;\theta)\ \frac{\partial f(x;\theta)}{\partial\theta}\right)\ dx\\&=E\left[\frac{\partial V(X;\theta)}{\partial\theta}+V(X;\theta)\frac{\partial\ln f(X;\theta)}{\partial\theta}\right]\ .\end{aligned}$$

When X_i, $i=1,..,n$, are mutually independent with PDF f_i, $\frac{\partial\ln f(X;\theta)}{\partial\theta}=\sum_{i=1}^n\frac{\partial\ln f_i(X_i;\theta)}{\partial\theta}$.

From the derivation above, we know if $V(x;\theta)$ is independent of parameter θ, then the first term in the expectation is zero, therefore unlike the pathwise derivative methods, the LRM estimator does not need to differentiate the payoff function V, an attractive property of LRM when V is discontinuous with respect to x. For the case when $V(x;\theta)$ is dependent on structural parameter θ without appropriate regularity conditions, e.g. continuity, both IPA and LRM fail. In this situation, the push-out LRM, which moves the parameter of interest into distributions, provides a potential solution.

In the rest of the section, we use the framework proposed in Peng et al. (2016a) to give a general discussion on the push-out LRM. Suppose $V(x;\theta)$ has the following structure:

$$V(x;\theta)=\varphi(G_1(x;\theta),..,G_m(x;\theta)),\quad m\leq n,$$

where $\varphi(y_1,..,y_m)$ could be a function without continuity condition and the dependence on parameter θ for $V(\cdot;\theta)$ only comes from $G_i(\cdot;\theta)$, $i=1,..,m$.

Remark. The framework can be easily extended as follows:

$$Q(x;\theta)=\vartheta(V_1(x;\theta),..,V_{n'}(x;\theta);\theta),$$

where $\vartheta(v_1,..,v_{n'},\theta)$ is differentiable with respect to $(v_1,..,v_{n'},\theta)$ and $V_i(\cdot)$, $i=1,..,n'$, are defined similarly as $V(\cdot)$.

The push-out LRM requires that there exists an inversion for the following change of variables:

$$y_1=G_1(x;\theta),\ldots,y_m=G_m(x;\theta),\quad y_{m+1}=x_{m+1},\ldots,y_n=x_n,\tag{1}$$

such that we can explicitly write it out as

$$x_1 = G_1^{-1}(y;\theta), \dots, x_m = G_m^{-1}(y;\theta), \quad x_{m+1} = y_{m+1}, \dots, x_n = y_n, \tag{2}$$

where $y \doteq (y_1,..,y_n)$. The subscript index of the variables $x_1,..,x_n$ are given arbitrarily, therefore we can always assume that the last $n-m$ transformations, if $n>m$, are identical without loss of generality. Generally, transformation (1) would map $\mathbb{R}^n$ onto some space $\Omega_Y(\theta)$ dependent on parameter θ. It is essential that $\Omega_Y(\theta)$ is independent of parameter θ, i.e. $\Omega_Y(\theta) = \Omega_Y$, to apply LRM. The most frequently used transformation in the literature is a linear transformation which has the form

$$y = A(\theta)x + B(\theta), \quad x = A^{-1}(\theta)y - A^{-1}(\theta)B(\theta), \tag{3}$$

assuming the matrix $A(\theta)$ is invertible. In this case, $\Omega_Y(\theta) = \mathbb{R}^n$.

We introduce the following notation: $G_k(x;\theta) \doteq x_k$, $G_k^{-1}(y;\theta) = y_k$ for $k = m+1,..,n$, $x_{1:k} \doteq (x_1,..,x_k)$, $y_{1:k} \doteq (y_1,..,y_k)$, $G_{1:k}(x;\theta) \doteq (G_1(x;\theta),..,G_k(x;\theta))$, $G_{1:k}^{-1}(y;\theta) \doteq (G_1^{-1}(y;\theta),..,G_k^{-1}(y;\theta))$ for $k = 1,..,n$, and $G(x;\theta) \doteq G_{1:n}(x;\theta)$. In the setting assumed above, we successfully push out the parameter θ into distribution by noticing that

$$E[V(X;\theta)] = \int\int_{\mathbb{R}^n} \varphi(G_{1:m}(x;\theta))\ f(x;\theta)\ dx = \int\int_{\mathbb{R}^n} \varphi(y_{1:m})\ f(G^{-1}(y;\theta);\theta)\ |\det(J(y;\theta))|\ dy, \tag{4}$$

where $G^{-1}(y;\theta) \doteq G_{1:n}^{-1}(y;\theta)$, $\det(\cdot)$ means the determinant, and J is the Jacobian matrix of the inverse transform (2) defined by

$$J(y;\theta) \doteq \begin{pmatrix} \frac{\partial G_1^{-1}(y;\theta)}{\partial y_1} & \frac{\partial G_1^{-1}(y;\theta)}{\partial y_2} & \cdots & \frac{\partial G_1^{-1}(y;\theta)}{\partial y_n} \\ \frac{\partial G_2^{-1}(y;\theta)}{\partial y_1} & \frac{\partial G_2^{-1}(y;\theta)}{\partial y_2} & \cdots & \frac{\partial G_2^{-1}(y;\theta)}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_n^{-1}(y;\theta)}{\partial y_1} & \frac{\partial G_n^{-1}(y;\theta)}{\partial y_2} & \cdots & \frac{\partial G_n^{-1}(y;\theta)}{\partial y_n} \end{pmatrix}.$$

Under appropriate regularity conditions, ordinary LRM can be applied, with the only difference being that the density of the distribution has been changed from $f(x;\theta)$ to $f(G^{-1}(x);\theta)\ |\det(J(x;\theta))|$. Notice that the procedure above relies on the existence of an explicit formula for inversion (2).

3 GENERALIZED LIKELIHOOD RATIO METHOD

In this section, we introduce the GLRM proposed in Peng et al. (2016a), which generalizes the push-out LRM in the case when the explicit inversion (2) cannot be found, although it exists and can be determined by an implicit function. For $k, k' = 1,..,m$, $h(x) = (h_1(x),..,h_{k'}(x))$, we define the following notation and operations: $\frac{\partial}{\partial x_{1:k}} \doteq (\frac{\partial}{\partial x_1},..,\frac{\partial}{\partial x_k})$,

$$\frac{\partial}{\partial x_{1:k}} \otimes h(x) \doteq \begin{pmatrix} \frac{\partial}{\partial x_1}h_1(x) & \frac{\partial}{\partial x_1}h_2(x) & \cdots & \frac{\partial}{\partial x_1}h_{k'}(x) \\ \frac{\partial}{\partial x_2}h_1(x) & \frac{\partial}{\partial x_2}h_2(x) & \cdots & \frac{\partial}{\partial x_2}h_{k'}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_k}h_1(x) & \frac{\partial}{\partial x_k}h_2(x) & \cdots & \frac{\partial}{\partial x_k}h_{k'}(x) \end{pmatrix},$$

and

$$h(x) \otimes \frac{\partial}{\partial x_{1:k}} \doteq \left(\frac{\partial}{\partial x_{1:k}} \otimes h(x)\right)^T,$$

$\frac{\partial^2}{\partial\theta\partial x_{1:m}} \doteq \left(\frac{\partial^2}{\partial\theta\partial x_1},..,\frac{\partial^2}{\partial\theta\partial x_n}\right)$ and $\frac{\partial^2}{\partial x_i\partial x_{1:m}} \doteq \left(\frac{\partial^2}{\partial x_i\partial x_1},..,\frac{\partial^2}{\partial x_i\partial x_m}\right)$, $i = 1,..,m$. Define p-norm as follows:

$$||h||_p \doteq \int |h(x)|^p dx \ .$$

The following regularity conditions are used to justify the unbiasedness of GLRM in Peng et al. (2016a).

(A.1) The set of discontinuity points for $\varphi(y_1,..,y_m)$ has zero measure.

(A.2) Invertibility Condition: $G_{1:m}(x;\theta)$ is differentiable with respect to x, $\forall\ x \in \mathbb{R}^n$ the matrix $\frac{\partial}{\partial x_{1:m}} \otimes G_{1:m}(x;\theta)$ is invertible, and denote its inversion as

$$(\Upsilon_{i,j}(x;\theta))_{m\times m} \doteq \left(\frac{\partial}{\partial x_{1:m}} \otimes G_{1:m}(x;\theta)\right)^{-1} = \begin{pmatrix} \frac{\partial G_1(x;\theta)}{\partial x_1} & \frac{\partial G_2(x;\theta)}{\partial x_1} & \cdots & \frac{\partial G_m(x;\theta)}{\partial x_1} \\ \frac{\partial G_1(x;\theta)}{\partial x_2} & \frac{\partial G_2(x;\theta)}{\partial x_2} & \cdots & \frac{\partial G_m(x;\theta)}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_1(x;\theta)}{\partial x_m} & \frac{\partial G_2(x;\theta)}{\partial x_m} & \cdots & \frac{\partial G_m(x;\theta)}{\partial x_m} \end{pmatrix}^{-1} .$$

Uniform Integrability: For $i = 1,..,m$, suppose $G_i(x;\theta)$ is twice continuously differentiable with respect to x,θ; $f(x;\theta)$ is continuously differentiable with respect to x,θ; there exists an integrable function $U_1(x)$ such that $\forall\ x \in \mathbb{R}^n$:

$$\max\left\{\sup_{\theta\in\Theta}\left|\Upsilon_{ij}(x;\theta)\ \frac{\partial G_i(x;\theta)}{\partial\theta}\right|, \sup_{\theta\in\Theta}|\varphi(G_{1:m}(x;\theta))|\right\} f(x;\theta) \le U_1(x),$$

$$\sup_{\theta\in\Theta}\left|\frac{\partial \ln f(x;\theta)}{\partial\theta} - \sum_{i=1}^m\sum_{j=1}^m\left(\frac{\partial\Upsilon_{ij}(x;\theta)}{\partial x_j}\frac{\partial G_i(x;\theta)}{\partial\theta} + \Upsilon_{ij}(x;\theta)\frac{\partial^2 G_i(x;\theta)}{\partial x_j\partial\theta}\right.\right.$$
$$\left.\left. + \Upsilon_{ij}(x;\theta)\frac{\partial G_i(x;\theta)}{\partial\theta}\frac{\partial \ln f(x;\theta)}{\partial x_j}\right)\right| f(x;\theta) \le U_1(x),$$

where Θ is any set containing a neighborhood of θ; there exists $U_2(y)$ such that $\forall\ y \in \mathbb{R}^n$:

$$\sup_{\theta\in\Theta}\left|\left\{\frac{\partial \ln f}{\partial\theta} - \sum_{i=1}^m\sum_{j=1}^m\left(\frac{\partial\Upsilon_{ij}}{\partial x_j}\ \frac{\partial G_i}{\partial\theta} + \Upsilon_{ij}\ \frac{\partial^2 G_i}{\partial x_j\partial\theta} + \Upsilon_{ij}\ \frac{\partial G_i}{\partial\theta}\ \frac{\partial \ln f}{\partial x_j}\right)\right\}(G^{-1}(y;\theta);\theta)\right|$$
$$\times f(G^{-1}(y;\theta);\theta)\ |\det(J(y;\theta))| \le U_2(y),$$

(A.3) Topology Condition: Denote a sub-matrix of $\frac{\partial}{\partial x_{1:m}} \otimes G_{1:m}(x;\theta)$ without the i-th row as $\Gamma_i(x)$. In addition, there exist $\varepsilon, M > 0$ and an $(m-1)\times(m-1)$ sub-matrix of $\Gamma_i(x)$ without the j-th column, denoted as $\Gamma_{ij}(x)$, such that

$$\inf_{x\in\mathbb{R}^n}\left|\det\left(\frac{\partial}{\partial x_{1:m}} \otimes G_{1:m}(x;\theta)\right)\right| > \varepsilon, \quad \sup_{x\in\mathbb{R}^n}\left|\det(\Gamma_{ij}(x))\right| < M\ .$$

(A.4) Dual Integrability Condition: for $q \in (1,\infty]$,

$$\sup_{\theta\in\Theta}\int_{\mathbb{R}^{n-m}} dy_{m+1:n}\left\{\int_{\mathbb{R}^m}|f(G^{-1}(y;\theta))\det(\Sigma_s^{-1}(G^{-1}(y;\theta));\theta)|^q dy_{1:m}\right\}^{\frac{1}{q}} < \infty,$$

and

$$\int_{\mathbb{R}^{n-m}} dy_{m+1:n}\left\{\int_{\mathbb{R}^m} U_2^q(y)\ dy_{1:m}\right\}^{\frac{1}{q}} < \infty\ .$$

The matrix invertibility condition in (A.2) guarantees inversion (2) exists by the implicit function theorem (Rudin 1964), so whenever push-out LRM that requires an explicit form transformation applies, condition (A.2) is satisfied. Uniform integrability in (A.2) is a typical condition to apply dominated convergence theorem that justifies the interchange of differentiation and expectation. (A.1) guarantees the existence of a smooth function converging to φ a.e., while for a measurable function φ in $\mathbb{L}^p$ space, there exists a smooth function converging to φ in $\mathbb{L}^p$ space. For general measurable function φ, the conclusion in the following Theorem 1 can first be established for a truncated sequence, and then extended to the general case using uniform integrability in (A.1). Condition (A.3) guarantees the image of transformation (1) is the whole space.

Theorem 1 (i) Under Conditions (A.1) and (A.2), the following Equation (5) holds;
(ii) Under Conditions (A.2) and (A.4) , the following Equation (5) holds;
(iii) Under Conditions (A.2) and (A.3), the following Equation (5) holds.

$$\begin{aligned}\frac{\partial}{\partial\theta}E[V(X;\theta)] = E\Bigg[V(X;\theta)\Bigg\{\sum_{j=1}^{m}\Bigg[\left(\frac{\partial}{\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)^{-1}\left(\frac{\partial^2}{\partial x_j\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)\\ \left(\frac{\partial}{\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)^{-1}e_j\Bigg]^T\left(G_{1:m}(x;\theta)\otimes\frac{\partial}{\partial\theta}\right) - tr\Bigg[\left(\frac{\partial}{\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)^{-1}\\ \left(\frac{\partial^2}{\partial\theta\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)\Bigg] - \left(\frac{\partial}{\partial\theta}\otimes G_{1:m}(x;\theta)\right)\left(\frac{\partial}{\partial x_{1:m}}\otimes G_{1:m}(x;\theta)\right)^{-1}\\ \left(\frac{\partial}{\partial x_{1:m}}\otimes \ln f(x;\theta)\right) + \frac{\partial \ln f(x;\theta)}{\partial\theta}\Bigg\}\Bigg|_{x=X}\Bigg],\end{aligned} \tag{5}$$

where $e_1 \doteq (1,0,0,..)^T$, $e_2 \doteq (0,1,0,..)^T$, ..., $e_m \doteq (0,0,..,1)^T$.

Notice that the formula of GLRM on the right hand side of Equation (5) only involves differentiation and elementary operations on $G(x;\theta)$ and density $f(x;\theta)$ of input random variables, which are basic building blocks of output random variables. The proof of the theorem can be found in Peng et al. (2016a). Proofs of (i) and (ii) use function smoothing and integration by parts. The difference is that (i) requires some continuity condition (A.1), while (ii) does not but requires stronger integrability condition (A.4). Basically, the less smoothness φ has, the stronger integrability condition is required. (iii) is obtained by change of variables and differentiation of implicit function. Therefore, conclusion (iii) shows GLRM is a generalization of push-out LRM when inversion (2) exists but not in closed form. Topology condition (A.3) is stronger than the invertibility condition in (A.2), but is satisfied by the invertible linear transformation (3) that has been applied most frequently in the application of push-out LRM.

In the rest of the section, we discuss the applicability of regularity conditions (A.1)-(A.4) in practice. Condition (A.1) is easy to satisfy. For example, it covers the case $\varphi(y_1,..,y_m) = \prod_{i=1}^m h_i(y_i)$, where $h_i(y_i)$ has countably many discontinuity points, which covers two examples in Section 4 and all examples in Peng et al. (2016b), Peng et al. (2016a), Peng et al. (2016c), and Peng et al. (2016d). The linear transformation given in Section 2 satisfies the condition (A.2). For a Markovian model, there exists a transition function $Y_i = \psi_i(Y_{i-1},X_i)$, $i>1$, and $Y_1 = X_1$. With the condition that $\prod_{i=2}^n \partial\psi_i(y_i,x_i)/\partial x_i \neq 0$ a.e., matrix invertibility condition in (A.2) can be guaranteed by the Markov model. We provide a simple example that satisfies the matrix invervibility condition in (A.2), without explicit inversion. Let $v_i(\theta)$, $i=2,..,m$, be linear independent row vectors in $\mathbb{R}^m$, let $v_1(\theta) = \sum_{i=2}^m v_i(\theta)$, and $A_{11}(\theta) = (v_2(\theta),..,v_m(\theta))^T$. Define the following transform:

$$y_1 = h(x) + v_1(\theta)\ x_{2:m}^T,$$
$$y_{2:m}^T = A_{11}(\theta)\ x_{2:m}^T\ .$$

For $h(x) = c_1(\theta)e^{c_2(\theta)x_1} + c_3(\theta)x_1 + c_4(\theta)$ and $h(x) = c_1(\theta)x_1^3 + c_2(\theta)x_1 + c_4(\theta)$, where $c_1(\theta), c_2(\theta), c_3(\theta) > 0$, the matrix invertibility condition in (A.2) can be easily proved to be satisfied for the transformation above, even without explicit inversion.

The Topology condition (A.3) required by (iii) could be violated in many practical problems. In Peng et al. (2016a), an application for the sensitivity estimates of a compound option, an option on an option, is provided. The payoff function of an European compound option is given by

$$Q(X_1;\theta) = e^{-rt_1}V(X_1;\theta),$$

where $X_1 = B_{t_1}/\sqrt{t_1}$,

$$V(X_1;\theta) = G_1(X_1;\theta)\ \mathbf{1}\{G_1(X_1;\theta) > 0\},$$

and

$$G_1(X_1;\theta) = C(S_{t_1};t_2-t_1;K_2) - K_1,$$

where $S_t = S_0 \exp\left((r-\sigma^2/2)t + \sigma\sqrt{t}X_1\right)$ and

$$C(S_t;t_2-t_1;K_2) = N(d_1)S_t - N(d_2)K_2e^{-r(t_2-t_1)},$$

where $N(\cdot)$ is the CDF of the standard normal distribution, and $C(\cdot)$ is the price of an option with expiration date at time t_2 and strike price K_1, of the underlying asset S_t at time t_1, given by the following Black-Scholes formula:

$$d_1 = \frac{1}{\sigma\sqrt{t_2-t_1}}\left[\log\left(\frac{S_t}{K_2}\right) + (r+\frac{\sigma^2}{2})(t_2-t_1)\right],$$
$$d_2 = d_1 - \sigma\sqrt{t_2-t_1}\ .$$

In addition, we have

$$\frac{\partial C(S_{t_1};t_2-t_1;K_2)}{\partial X_1} = \frac{\partial C(S_{t_1};t_2-t_1;K_2)}{\partial S}S_{t_1}\sigma\sqrt{t_1},$$

where $\frac{\partial C}{\partial S} = N(d_1)$. Conditions in (i) (or (ii)) can be easily shown to be satisfied. However, we can see $G_1(\cdot;\theta)$: $\mathbb{R} \to [-K_1,\infty)$ and

$$\lim_{X_1\to\infty}\left(\frac{\partial C(S_{t_1};t_2-t_1;K_2)}{\partial X_1}\right)^{-1} = 0,$$

therefore Condition (A.3) is not satisfied by this example.

4 APPLICATIONS

4.1 Asian Option

An Asian option is a path-dependent financial derivative with its payoff depending on an average of prices of underlying assets throughout the path. The sensitivity of Asian option is considered to be a difficult problem (Boyle and Potapchik 2008). We illustrate how to apply GLRM to address this problem. The payoff function of an arithmetic mean Asian call option is given by

$$Q(X_1,..,X_n;\theta) = e^{-rn\Delta}V(X_1,..,X_n;\theta),$$

where for $i = 1,..,n$, $X_i = (B_{i\Delta} - B_{(i-1)\Delta})/\sqrt{\Delta}$, where Δ is the duration between two adjacent price ticks,

$$V(X_1,..,X_n;\theta) = G_1(X_1,..,X_n;\theta)\ \mathbf{1}\{G_1(X_1,..,X_n;\theta) > 0\},$$

and

$$G_1(X_1,..,X_n;\theta) = \frac{1}{n}\sum_{i=1}^{n} S_0 \exp\left\{\sigma\sqrt{\Delta}\sum_{j=1}^{i} X_j + i\left(r - \frac{\sigma^2}{2}\right)\Delta\right\} - K\ .$$

The first-order sensitivity estimate is not a issue for an Asian option, since the payoff is still continuous, although it has a kink. However, the discontinuity is an issue for estimating the second-order sensitivity (as well as the first-order sensitivity of a digital Asian option). For example, the payoff function of the first-order derivative with respect to K is

$$Q(X_1,..,X_n;\theta) = -e^{-rn\Delta}V(X_1,..,X_n;\theta),$$

where

$$V(X_1,..,X_n;\theta) = -\mathbf{1}\{G_1(X_1,..,X_n;\theta) > 0\} \ .$$

The payoff function for an Asian call option with geometric mean is given by

$$Q(X_1,..,X_n;\theta) = e^{-rn\Delta}KV(X_1,..,X_n;\theta),$$

where

$$V(X_1,..,X_n;\theta) = \prod_{i=1}^{n} \exp\{G_i(X_1,..,X_i;\theta)\}^{\frac{1}{n}} \ \mathbf{1}\left\{\prod_{i=1}^{n} \exp\{G_i(X_1,..,X_i;\theta)\}^{\frac{1}{n}} > 1\right\},$$

and for $i = 1,..,n$,

$$G_i(X_1,..,X_i;\theta) = \log S_0 - \log K + \sigma\sqrt{\Delta}\sum_{j=1}^{i} X_j + i\left(r - \frac{\sigma^2}{2}\right)\Delta \ .$$

4.2 Inventory Processes

We use the formulation of a single-commodity, discrete-time, multiperiod inventory model with backlogging in Pflug and Rubinstein (2002), which applied push-out LRM to address the same problem. As discussed in the last section, GLRM is a generalization of push-out LRM, so in this example where the latter applies, they would have the same estimator. Please refer to Pflug and Rubinstein (2002) for numerical results. We illustrate how to put this problem into the framework given in Section 2. After fitting the problem with the framework, there is no need to decide how to push the parameter into distribution, and the GLRM formula in Theorem 1 has implicitly pushed the parameter out of the discontinuous sample performance.

We focus on the scenario that has stochastic lead times under no-overtaking rule. Applying the quasi-regenerative theorem, Pflug and Rubinstein (2002) show the steady-state cost per time can be written as the following performance measure:

$$\ell(S,L) = \frac{E\left[\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\bar{\mathscr{Y}}_m)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2} > L\right\}\right]+K}{E\left[\sum_{n=0}^{\infty}\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n < L\right\}+R_1\right]} + \frac{E\left[\sum_{n=0}^{\infty}\left(\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n)+\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n-\bar{\mathscr{Y}}_m)\right)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n < L\right\}\right]}{E\left[\sum_{n=0}^{\infty}\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n < L\right\}+R_1\right]},$$

where $L = S - s$, s and S are the parameters in the (s,S) policy, $\varphi(x) = c_1x^+ + c_2x^-$, $x^+ = \max\{x,0\}$ and $x^- = \max\{x,0\}$, R_1 and R_2 are independent copies of the stochastic lead time, $\mathscr{Y}_n = \sum_{i=1}^{n} Y_i$, $\hat{\mathscr{Y}}_m = \sum_{j=1}^{m}\hat{Y}_j$, $\bar{\mathscr{Y}}_l = \sum_{k=1}^{l}\bar{Y}_k$, Y_i, $\hat{Y}_j$, $\bar{Y}_k$ are independent copies of the stochastic demands, $\sum_{i=1}^{j} \doteq 0$ if $j \le 0$, and K is the fixed cost for each order.

For the sensitivity with respect to the parameter θ, we need to estimate

$$\frac{\partial}{\partial\theta}E\left[\sum_{n=0}^{\infty}\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n<L\right\}+R_1\right]$$
$$=E\left[\left(\sum_{n=0}^{\infty}\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n<L\right\}+R_1\right)\left(\frac{\partial\ln f_R(R_1;\theta)}{\partial\theta}+\frac{\partial\ln f_R(R_2;\theta)}{\partial\theta}\right)\right]$$
$$+\sum_{n=0}^{\infty}E\left[\frac{\partial}{\partial\theta}E\left[V_n(X_1,,..,X_{n+R_2};\theta)|R_1,R_2\right]\right],$$

where $X_i\doteq Y_i,\ i=1,..,n,\ X_{j+n}\doteq\hat{Y}_j,\ j=1,..,R_2$, and

$$V_n(X_1,,..,X_{n+R_2};\theta)\doteq\mathbf{1}\left\{G_n(X_1,,..,X_{n+R_2};\theta)<0\right\},$$
$$G_n(X_1,,..,X_{n+R_2};\theta)\doteq\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n-L;$$

$$\frac{\partial}{\partial\theta}E\left[\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\bar{\mathscr{Y}}_m)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}>L\right\}\right]$$
$$=E\left[\left(\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\bar{\mathscr{Y}}_m)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}>L\right\}\right)\left(\frac{\partial\ln f_R(R_1;\theta)}{\partial\theta}+\frac{\partial\ln f_R(R_2;\theta)}{\partial\theta}\right)\right]$$
$$+E\left[\frac{\partial}{\partial\theta}E\left[Q(X_1,,..,X_{R_1+R_2-1};\theta)|R_1,R_2\right]\right],$$

where $X_i\doteq\hat{Y}_i,\ i=1,..,R_2,\ X_{R_2+j}\doteq\bar{Y}_j,\ j=1,..,R_1-1$, and

$$Q(X_1,,..,X_{R_1+R_2-1};\theta)\doteq(S-L)\mathbf{1}\left\{G_1(X_1,..,X_{R_2};\theta)>1\right\}$$
$$\times\left[\sum_{m=1}^{R_1-1}\varphi\left(G_1(X_1,..,X_{R_2};\theta)+G_m(X_{R_2},..,X_{m+R_2};\theta)\right)\right],$$
$$G_1(X_1,,..,X_{R_2};\theta)\doteq\frac{S-\hat{\mathscr{Y}}_{R_2}}{S-L},\quad G_m(X_{R_2+1},..,X_{m+R_2};\theta)\doteq-\frac{\bar{\mathscr{Y}}_m}{S-L},\ m=1,..,R_1-1;$$

$$\frac{\partial}{\partial\theta}E\left[\sum_{n=0}^{\infty}\left(\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n)+\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n-\bar{\mathscr{Y}}_m)\right)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n<L\right\}\right]$$
$$=E\left[\left\{\sum_{n=0}^{\infty}\left(\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n)+\sum_{m=1}^{R_1-1}\varphi(S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n-\bar{\mathscr{Y}}_m)\right)\mathbf{1}\left\{\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n<L\right\}\right\}\right.$$
$$\left.\times\left(\frac{\partial\ln f_R(R_1;\theta)}{\partial\theta}+\frac{\partial\ln f_R(R_2;\theta)}{\partial\theta}\right)\right]$$
$$+\sum_{n=0}^{\infty}E\left[\frac{\partial}{\partial\theta}E\left[Q_n(X_1,,..,X_{n+R_1+R_2-1};\theta)|R_1,R_2\right]\right],$$

where $X_i\doteq Y_i,\ i=1,..,n,\ X_{n+j}\doteq\hat{Y}_j,\ j=1,..,R_2,\ X_{n+R_2+k}\doteq\bar{Y}_k,\ j=1,..,R_1-1$, and

$$Q_n(X_1,,..,X_{R_1+R_2-1};\theta)\doteq(S-L)\mathbf{1}\left\{G_1(X_1,..,X_{n+R_2};\theta)>1\right\}\times$$
$$\left[\varphi\left(G_1(X_1,..,X_{n+R_2};\theta)\right)+\sum_{m=1}^{R_1-1}\varphi\left(G_1(X_1,..,X_{n+R_2};\theta)+G_m(X_{n+R_2},..,X_{n+R_2+m};\theta)\right)\right],$$
$$G_1(X_1,,..,X_{n+R_2};\theta)\doteq\frac{S-\hat{\mathscr{Y}}_{R_2}-\mathscr{Y}_n}{S-L},\quad G_m(X_{n+R_2+1},..,X_{n+R_2+m};\theta)\doteq-\frac{\bar{\mathscr{Y}}_m}{S-L},\ m=1,..,R_1-1.$$

5 NUMERICAL EXPERIMENT

In this section, we test the numerical performance of GLRM on sensitivity estimates of the arithmetic mean Asian call option in Section 4, and compare it with a finite difference estimate with common random number (FDC), and IPA. We first test the sensitivity estimate with respect to K. In this example, $m = 1$ and we have $\frac{\partial G_1(x;\theta)}{\partial x_1} = \sigma\sqrt{\Delta}(G_1(x;\theta)+K)$, which satisfies Condition (A.2), $\frac{\partial^2 G_1(x;\theta)}{\partial x_1^2} = \sigma^2\Delta(G_1(x;\theta)+K)$, $\frac{\partial G_1(x;\theta)}{\partial K} = -1$, $\frac{\ln f(x)}{\partial x_1} = -x_1$, and $\frac{\ln f(x)}{\partial K} = 0$. Therefore, a GLRM estimator for the first derivative sensitivity can be explicitly given by

$$-e^{-rn\Delta}\mathbf{1}\{G_1(X;\theta) > 0\}\frac{G_1(X;\theta)}{G_1(X;\theta)+K}\left(1+\frac{X_1}{\sigma\sqrt{\Delta}}\right),$$

and a GLRM estimator for the second derivative sensitivity is

$$e^{-rn\Delta}\mathbf{1}\{G_1(X;\theta) > 0\}\frac{G_1(X;\theta)}{(G_1(X;\theta)+K)^2}\left[\frac{1}{\sigma\sqrt{\Delta}}+\left(1+\frac{X_1}{\sigma\sqrt{\Delta}}\right)^2\right] .$$

The IPA estimator for the first derivative sensitivity is

$$-e^{-rn\Delta}\mathbf{1}\{G_1(X;\theta) > 0\},$$

but the IPA estimator for the second derivative sensitivity fails to be unbiased. The sensitivity is tested at $\sigma = 0.1$, $r = 0.005$, $S_0 = K = 100$, $\Delta = 1$, and $n = 5$.

Table 1: First and second derivatives estimates of arithmetic mean Asian call option with respect to K, based on 10^6 independent replications (mean$\pm$ standard error).

	FDC	IPA	GLRM
1st derivative	$-0.395 \pm 4\times10^{-4}$	$-0.396 \pm 4\times10^{-4}$	-0.396 ± 10^{-3}
2nd derivative	$0.020 \pm 3\times10^{-3}$	N/A	0.018 ± 10^{-4}

In general, GLRM estimator for the first derivative sensitivity can be explicitly given by

$$-Q(x;\theta)\left\{n\Delta\frac{\partial r}{\partial\theta}+\left[\frac{\partial^2 G_1(x;\theta)}{\partial\theta\partial x_1}+\frac{\partial G_1(x;\theta)}{\partial\theta}\left(\frac{\partial ln f(x)}{\partial x_1}-\frac{\partial^2 G_1(x;\theta)}{\partial x_1^2}\Big/\frac{\partial G_1(x;\theta)}{\partial x_1}\right)\right]\Big/\frac{\partial G_1(x;\theta)}{\partial x_1}\right\} .$$

Although GLRM can be applied to the second-order sensitivity in general form, the computation is tedious and we omit the discussion. Besides the arithmetic mean Asian call option, we can also consider a digital arithmetic mean Asian call option with a payoff:

$$Q(x;\theta) = e^{-rn\Delta}\mathbf{1}\{G_1(x;\theta) > 0\},$$

which is discontinuous. For σ, we have

$$\frac{\partial G_1(x;\theta)}{\partial\sigma} = \frac{1}{n}\sum_{i=1}^{n} S_0\left(\sqrt{\Delta}\sum_{j=1}^{i} x_j - i\sigma\Delta\right)\exp\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i} x_j + i\left(r-\frac{\sigma^2}{2}\right)\Delta\right),$$

and

$$\frac{\partial^2 G_1(x;\theta)}{\partial\sigma\partial x_1} = \frac{1}{n}\sum_{i=1}^{n} S_0\sqrt{\Delta}\left[\sigma\left(\sqrt{\Delta}\sum_{j=1}^{i} x_j - i\sigma\Delta\right)+1\right]\exp\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i} x_j + i\left(r-\frac{\sigma^2}{2}\Delta\right)\right) .$$

Table 2: Derivatives estimates of arithmetic mean Asian and digital arithmetic mean Asian call options with respect to σ, based on 10^6 independent replications (mean$\pm$ standard error).

	FDC	IPA	GLRM
Asian option	57.7 ± 0.1	57.5 ± 0.1	57.5 ± 0.4
digital Asian option	-0.68 ± 10^{-2}	N/A	$-0.71 \pm 8\times10^{-3}$

For r, we have

$$\frac{\partial G_1(x;\theta)}{\partial r} = \frac{1}{n}\sum_{i=1}^{n} S_0 i\Delta \exp\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i} x_j + i\left(r-\frac{\sigma^2}{2}\right)\Delta\right),$$

and

$$\frac{\partial^2 G_1(x;\theta)}{\partial r\partial x_1} = \frac{1}{n}\sum_{i=1}^{n} S_0 i\sigma\Delta^{3/2} \exp\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i} x_j + i\left(r-\frac{\sigma^2}{2}\right)\Delta\right) .$$

Table 3: Derivatives estimates of arithmetic mean Asian and digital arithmetic mean Asian call options with respect to r, based on 10^6 independent replications (mean$\pm$ standard error).

	FDC	IPA	GLRM
Asian option	139.2 ± 0.1	138.6 ± 0.1	138.4 ± 0.4
digital Asian option	$5.4 \pm 8\times10^{-2}$	N/A	$5.4 \pm 2\times10^{-2}$

For the first derivative estimates in Table 1, 2 and 3, we can see FDC has comparable performance with IPA, while both of them have variances lower than GLRM. IPA differentiates the sample performance, which generally leads to low variance for the sample performance with semi-linear form $(x-x_0)^+$. Notice that the sample performance goes to infinity as x goes to infinity while its derivative which is $\mathbf{1}\{x > x_0\}$ is bounded by one. However, for the second derivative estimates in Table 1 and the sensitivities of digital arithmetic mean Asian option in Table 2 and 3, we can see the variance of GLRM is lower than that of FDC. Unlike the semi-linear sample performance in the first derivative, the sample performances of the second derivative and digital option are an indicator bounded by one. This observations are consistent with those for other numerical examples in Peng et al. (2016a).

6 CONCLUSIONS

In this paper, we compare different sets of regularity conditions under which GLRM is derived in different ways. We find that the regularity conditions required in the derivation using function smoothing and integration by parts are easier to satisfy than those required in the derivation using change of variables and differentiation of implicit functions in practice. However, the latter derivation establishes a direct connection with the push-out LRM that is well-known in stochastic derivative estimation, which makes GLRM more easily understandable to theorists and practitioners in this field. In addition, we apply GLRM to more applications, which substantiates the capability of GLRM to solve broad problems in a unified framework. In the numerical experiment where the sensitivity of Asian option is estimated, although GLRM is not always the best alternative in all cases, it applies to more general scenarios than other stochastic derivative estimators and is also not dominated by any method in terms of performance.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) under Grants CMMI-1362303, CMMI-1434419, by the National Natural Science Foundation of China (NSFC) under Grants 71571048, 71071040, by the Air Force of Scientific Research (AFOSR) under Grant FA9550-15-10050, by the Science and Technology Agency of Sichuan Province under Grant 2014GZX0002, by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institution of Higher Learning, and by the China Postdoctoral Science Foundation under Grant 2015M571495.

REFERENCES

Boyle, P. P., and A. Potapchik. 2008. "Prices and sensitivities of Asian options: a survey". *Insurance Mathematics and Economics* 42 (1): 189–211.

Chen, N., and P. Glasserman. 2007. "Malliavin Greeks without Malliavin calculus". *Stochastic Processes and their Applications* 117 (11): 1689–1723.

Fournié, E., J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi. 1999. "Applications of Malliavin calculus to Monte Carlo methods in finance". *Finance and Stochastics* 3 (4): 391–412.

Fu, M. C. 2006. "Gradient Estimation". In *Henderson S.G. and Nelson B.L.(ed.), Chapter 19 in Handbooks in Operations Research and Management Science*, Volume 13, 575–616: Elsevier.

Fu, M. C. 2008. "What you should know about simulation and derivatives". *Naval Research Logistics* 55 (8): 723–736.

Fu, M. C. 2015. "Stochastic Gradient Estimation". In *Fu, Michael C. (ed.), Chapter 5 in Handbooks of Simulation Optimization*, Volume 216, 105–147: Springer.

Glynn, P. W. 1990. "Likelihood ratio gradient estimation for stochastic systems". *Communications of the ACM* 33 (10): 75–84.

L'Ecuyer, P. 1990. "A unified view of the IPA, SF, and LR gradient estimation techniques". *Management Science* 36 (11): 1364–1383.

Peng, Y., M. C. Fu, and J. Hu. 2016a. "A general framework for single-run stochastic derivative estimation with discontinuities". *Submitted to Operations Research*.

Peng, Y., M. C. Fu, and J. Hu. 2016b. "Quantile sensitivity estimation for stochastic models with discontinuities". *Submitted to Operations Research*.

Peng, Y., M. C. Fu, and J.-Q. Hu. 2016c. "Estimating distribution sensitivity using generalized likelihood ratio method". *WODES*:123–128.

Peng, Y., M. C. Fu, and J.-Q. Hu. 2016d. "Sensitivity analysis of distorted risk measure by Monte Carlo simulation". *working paper*.

Pflug, G., and R. Y. Rubinstein. 2002. "Inventory processes: Quasi-regenerative property, performance evaluation, and sensitivity estimation via simulation". *Stochastic Models* 18 (3): 469–496.

Rubinstein, R. Y. 1992. "Sensitivity analysis of discrete event systems by the "push out" method". *Annals of Operations Research* 39 (1): 229–250.

Rubinstein, R. Y. 1997. "Optimization of computer simulation models with rare events". *European Journal of Operational Research* 99 (1): 89–112.

Rubinstein, R. Y., and A. Shapiro. 1993. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, Volume 346. Wiley New York.

Rudin, W. 1964. *Principles of Mathematical Analysis*, Volume 3. McGraw-Hill New York.

Wang, Y., M. C. Fu, and S. I. Marcus. 2012. "A new stochastic derivative estimator for discontinuous payoff functions with application to financial derivatives". *Operations Research* 60 (2): 447–460.

AUTHOR BIOGRAPHIES

YIJIE PENG is a postdoctoral scholar in the Department of Management Science, School of Management, Fudan University. He received his B.S. degree in applied mathematics from Wuhan University, China, and Ph.D. degree in management science from Fudan University, China. His research interests lie in simulation methodology, modeling, analysis, and optimization. His email address is pengy10@fudan.edu.cn.

MICHAEL C. FU is Ralph J. Tyser Professor of Management Science in the Robert H. Smith School of Business, with a joint appointment in the Institute for Systems Research and affiliate faculty appointment in the Department of Electrical and Computer Engineering, all at the University of Maryland. His research interests include simulation optimization and applied probability, with applications in supply chain management and financial engineering. He has a Ph.D. in applied math from Harvard and degrees in math and EECS from MIT. He served as WSC2011 Program Chair, NSF Operations Research Program Director, *Management Science* Stochastic Models and Simulation Department Editor, and *Operations Research* Simulation Area Editor. He is a Fellow of INFORMS and IEEE. His email address is mfu@umd.edu.

JIAN-QIANG HU is a Professor with the Department of Management Science, School of Management, Fudan University. He received his B.S. degree in applied mathematics from Fudan University, China, and M.S. and Ph.D. degrees in applied mathematics from Harvard University. He was an Associate Professor with the Department of Mechanical Engineering and the Division of Systems Engineering at Boston University before joining Fudan University. His research interests include discrete-event stochastic systems, simulation, queueing network theory, stochastic control theory, with applications towards supply chain management, risk management in financial markets and derivatives, and communication networks. His e-mail addresses is hujq@fudan.edu.cn.