

A REFLECTION-BASED VARIANCE REDUCTION TECHNIQUE FOR SUM OF RANDOM VARIABLES

Guangwu Liu

Department of Management Sciences
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong

ABSTRACT

Monte Carlo simulation has been widely used as a standard tool for estimating expectations. In this paper we develop a variance reduction technique for a particular case when the expectation is taken under a constraint that a sum of random variables is larger than a threshold. The proposed technique is based on a reflection argument on the sample space and requires knowing the joint density of the random variables. It turns out the technique can always guarantee a variance reduction. More importantly, the technique sheds light on how observations violating the constraint can be used more efficiently in estimation, compared to crude Monte Carlo.

1 INTRODUCTION

In this paper we are interested in the problem of estimating

$$\alpha = \mathbb{E} \left[h(\mathbf{X}) 1_{\{X_1 + \dots + X_m > y\}} \right] \quad (1)$$

via Monte Carlo simulation, where $\mathbf{X} = (X_1, \dots, X_m)$ is an m -dimensional random vector, $h(\cdot)$ is a function, and y is a given constant.

The above problem arises in a number of applications such as portfolio risk measurement and option pricing. In the application of risk measurement of a portfolio consisting of m assets whose prices at a future time point is denoted by \mathbf{X} , if we let $h(\cdot) \equiv 1$, then α may represent the probability that the portfolio loss is larger than some threshold, see, e.g., Asmussen et al. (2011). In the application of option pricing, \mathbf{X} may represent the prices of an underlying asset at m discrete time steps. Then the integrand in (1) may denote the discounted payoff of an Asian-type option under risk-neutral measure, and hence α is the value of the option.

In this paper we propose a variance reduction technique for simulating the quantity in (1). In the general setting of simulating expectations, various variance reduction techniques have been extensively studied in the literature. For comprehensive overviews, one is referred to Law and Kelton (2000), Asmussen and Glynn (2007) and Glasserman (2004). This paper contributes to the literature by proposing a new variance reduction technique which makes use of the special structure of the problem in (1). Although the efficiency improvement might not be so marked in some practical examples, the technique always provides variance reduction, which is appealing given that the additional computational effort required is often minor.

The essential of the technique is to cancel the discontinuity introduced by the indicator function in the integrand. To this end, density information of the random variables is employed and a smooth function is derived to replace the indicator function. From such a perspective, the technique shares the same spirit as another variance reduction technique, the so-called conditional Monte Carlo (see, e.g., Asmussen and Glynn (2007) and Fu and Hu (1997)), although the ways of smoothing are different.

Given that the integrand has been smoothed, the technique may also be applied in derivative estimation, which is similar to the idea of smoothed perturbation analysis (SPA), see, e.g., Gong and Ho (1987)

and Fu and Hu (1997). The major argument is that smoothing the integrand enables the use of pathwise method (also called Infinitesimal Perturbation Analysis (IPA), see, e.g., Ho and Cao (1983) and Broadie and Glasserman (1996)), which may not be applicable without smoothing, due to the discontinuity introduced by the indicator function.

The way of smoothing of the proposed technique is based on a reflection argument, and hence we refer to it as a *reflection-based technique*. It has a clear geometric interpretation. Note that a crude Monte Carlo estimator of α is actually a weighted average of the observations of the integrand, with a zero weight for the observations that lie below the hyperplane defined by $\{\mathbf{x} : x_1 + \dots + x_m = y\}$. A key property is that when an observation lies below the hyperplane, its reflection must be above the hyperplane. This makes it possible to somehow replace the observation by its reflection whenever the observation lies below the hyperplane. This replacement is made rigorous mathematically by employing the density of the random vector \mathbf{X} . Although the requirement of knowing the density imposes a restriction on the applicability of the technique, there does exist some practical applications where such density information is available.

The rest of the paper is organized as follows. Section 2 presents the formulation and analysis of the reflection-based variance reduction technique. Section 3 discusses the application of the proposed technique in derivative estimation. Numerical experiments are presented in Section 4, followed by conclusions in Section 5.

2 A REFLECTION BASED VARIANCE REDUCTION TECHNIQUE

2.1 A Motivated Example

Consider a continuous random variable X , with a support on \mathbb{R} . One may want to estimate the probability that X is larger than some threshold value y , i.e.,

$$\alpha = \mathbb{E} [1_{\{X > y\}}].$$

A crude Monte Carlo estimator for α is $1_{\{X > y\}}$. Essentially it assigns a weight 1 to the observations that are larger than y , while a zero weight to the observations that are smaller than or equal to y . Therefore, only the observations larger than y play important roles in estimation. When the threshold value y is large, most of the observations will be smaller than or equal to y and hence are assigned a zero weight. This makes the estimation extremely difficult, because most of the observations contribute very little to the estimation. In such cases, efficiency improvement of the estimation is highly desirable.

A natural idea is to make better use of the observations at hand. To this end, a key question is whether the observations smaller than or equal to y can contribute more in the estimation. As shown in the following discussion, it turns out the answer is yes, given that we know the density of X .

Specifically, if we let $f(\cdot)$ denote the density function of X . Then it can be easily verified that

$$\alpha = \mathbb{E} \left[\frac{f(r(X))}{f(X) + f(2y - X)} \right], \quad (2)$$

where

$$r(X) = \begin{cases} X & \text{if } X > y, \\ 2y - X & \text{if } X \leq y. \end{cases}$$

The representation in (2) delivers an interesting message that even an observation is smaller than or equal to y , it may be assigned a nonzero weight and plays a more important role in the estimation. Compared to the binary weights in crude Monte Carlo, the weights in (2) is smoother and thus has the potential of reducing variance.

2.2 Formulation

We return to the more general setting in (1). Recall that the key message in the motivated example is that even for an observation that does not satisfy $\{X > y\}$, one may still assign a nonzero weight to it. The weight depends on the observation itself (i.e., X) and its reflection (i.e., $2y - X$) with respect to the point y .

To extend the above idea to high-dimensional cases, we first note that $\{x_1 + \dots + x_m = y\}$ defines a hyperplane in the Euclidean \mathbb{R}^m space. Then for any observation, no matter it lies above or below the hyperplane, we are able to find its reflection with respect to the hyperplane. Specifically, for a two-dimensional case, we illustrate a point (a_1, a_2) and its reflection (b_1, b_2) in Figure 1. From the figure we can see that for any observation that lies below the line $x_1 + x_2 = y$, we can find its reflection that lies above the line.

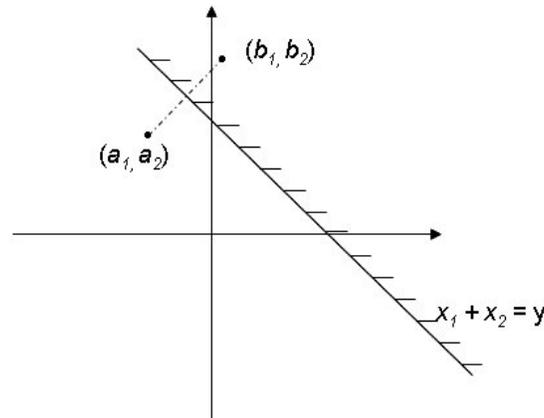


Figure 1: The point (b_1, b_2) is the reflection of (a_1, a_2) with respect to the straight line $x_1 + x_2 = y$.

To express the relation between a point and its reflection in mathematical terms, for any $\mathbf{a} = (a_1, \dots, a_m)^T \in \mathbb{R}^m$, we let $\mathbf{b} = (b_1, \dots, b_m)^T$ denote its reflection with respect to the hyperplane $x_1 + \dots + x_m = y$. We further let $\mathbf{1}_m$ denote an m -dimensional column vector with each entry being one, and $\mathbf{1}_{m \times m}$ an $m \times m$ matrix with each entry being one. Then by elementary algebra, we can see that

$$\mathbf{b} = \mathbf{A}\mathbf{a} + 2y/m\mathbf{1}_m,$$

where $\mathbf{A} = \mathbf{I} - \frac{2}{m}\mathbf{1}_{m \times m}$, with \mathbf{I} denoting the identity $m \times m$ matrix.

It can be easily verified that $\det(\mathbf{A}) = -1$, $\mathbf{A}^{-1} = \mathbf{A}$, and thus $\mathbf{A}^2\mathbf{x} = \mathbf{x}$. Then we have the following theorem on a representation of α .

Theorem 1 Suppose that \mathbf{X} is an m -dimensional continuous random vector with a support \mathbb{S} being an open subset of \mathbb{R}^m . Assume that \mathbf{X} has a density $f(\cdot)$, and $E(|h(\mathbf{X})|) < \infty$. Then,

$$\alpha = E \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} \right],$$

where the random vector $r(\mathbf{X})$ is defined by

$$r(\mathbf{X}) = \begin{cases} \mathbf{X} & \text{if } X_1 + \dots + X_m > y, \\ \mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m & \text{if } X_1 + \dots + X_m \leq y. \end{cases}$$

Proof. For notational ease we let $\mathbb{A}_1 = \{\mathbf{x} \in \mathbb{S} : x_1 + \dots + x_m > y\}$, and $\mathbb{A}_2 = \{\mathbf{x} \in \mathbb{S} : x_1 + \dots + x_m < y\}$. Note that \mathbf{X} is a continuous random vector and thus $\{\mathbf{x} \in \mathbb{S} : x_1 + \dots + x_m = y\}$ is a probability-zero set.

Hence

$$\begin{aligned} & \mathbb{E} \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} \right] \\ &= \mathbb{E} \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} (1_{\mathbb{A}_1} + 1_{\mathbb{A}_2}) \right] \\ &= \mathbb{E} \left[h(\mathbf{X}) \frac{f(\mathbf{X})}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} 1_{\mathbb{A}_1} \right] + \mathbb{E} \left[h(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m) \frac{f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} 1_{\mathbb{A}_2} \right]. \end{aligned} \tag{3}$$

We use a change of variable for the second term in the right-hand-side of Equation (3). Specifically, we let $\mathbf{u} = \mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m$, then it can be verified that the Jacobian of the change of variable is $|\det(A)| = 1$, and $\mathbf{x} = \mathbf{A}\mathbf{u} + 2y/m\mathbf{1}_m$. Therefore,

$$\begin{aligned} & \mathbb{E} \left[h(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m) \frac{f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} 1_{\mathbb{A}_2} \right] \\ &= \int_{\mathbb{A}_2} h(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m) \frac{f(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m)}{f(\mathbf{x}) + f(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m)} f(\mathbf{x}) \, d\mathbf{x} \\ &= \int_{\mathbb{A}_2^r} h(\mathbf{u}) \frac{f(\mathbf{u})}{f(\mathbf{A}\mathbf{u} + 2y/m\mathbf{1}_m) + f(\mathbf{u})} f(\mathbf{A}\mathbf{u} + 2y/m\mathbf{1}_m) \, d\mathbf{u} \\ &= \mathbb{E} \left[h(\mathbf{X}) \frac{f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)}{f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m) + f(\mathbf{X})} 1_{\mathbb{A}_2^r} \right], \end{aligned} \tag{4}$$

where \mathbb{A}_2^r denotes the reflection of the set \mathbb{A}_2 .

Note that $\mathbb{A}_2^r \subset \mathbb{A}_1$. Therefore, by Equations (3) and (4),

$$\begin{aligned} & \mathbb{E} \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} \right] \\ &= \mathbb{E} \left[h(\mathbf{X}) 1_{\mathbb{A}_2^r} + h(\mathbf{X}) \frac{f(\mathbf{X})}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} 1_{\mathbb{A}_1 \setminus \mathbb{A}_2^r} \right]. \end{aligned}$$

We assert that if $\mathbf{x} \in \mathbb{A}_1 \setminus \mathbb{A}_2^r$, then $f(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m) = 0$. The assertion can be shown by contradiction. Suppose that $f(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m) > 0$. Then combining it with $\mathbf{x} \in \mathbb{A}_1$, we have $\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m \in \mathbb{A}_2$, and thus $\mathbf{x} \in \mathbb{A}_2^r$, leading to a contradiction with $\mathbf{x} \in \mathbb{A}_1 \setminus \mathbb{A}_2^r$.

Since $f(\mathbf{A}\mathbf{x} + 2y/m\mathbf{1}_m) = 0$ when $\mathbf{x} \in \mathbb{A}_1 \setminus \mathbb{A}_2^r$, it can be easily seen that

$$\begin{aligned} & \mathbb{E} \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} \right] \\ &= \mathbb{E} \left[h(\mathbf{X}) 1_{\mathbb{A}_2^r} + h(\mathbf{X}) 1_{\mathbb{A}_1 \setminus \mathbb{A}_2^r} \right] = \mathbb{E} [h(\mathbf{X}) 1_{\mathbb{A}_1}] = \alpha, \end{aligned}$$

which completes the proof. □

Theorem 1 provides a new representation the quantity α . It can be seen that under the new representation, the integrand is smoother, thus intuitively leading to an estimator with a smaller variance, compared to the crude Monte Carlo estimator. This intuition will be confirm mathematically in the subsequent sections.

Incorporating into the idea of importance sampling, the theorem can be generalized as follows. Suppose that the random vector \mathbf{X} is generated from another probability distribution with density denoted by $g(\cdot)$ (g is absolutely continuous with respect to f). Then a straightforward extension of Theorem 1 is to represent the quantity being estimated as:

$$\alpha = \mathbb{E}_g \left[h(r(\mathbf{X})) \frac{f(r(\mathbf{X}))}{g(\mathbf{X}) + g(\mathbf{A}\mathbf{X} + 2y/m\mathbf{1}_m)} \right], \tag{5}$$

where $r(\mathbf{X})$ is defined in Theorem 1, and E_g denotes the operator of taking expectation under the probability measure g .

2.2.1 A Pathological Example

We consider a simple but pathological example to illustrate that the variance reduction could be substantial. Let $\mathbf{X} = (X_1, \dots, X_m)$, where X_i 's are m independent standard normal random variables. We want to estimate

$$\alpha = E [1_{\{X_1 + \dots + X_m > 0\}}].$$

It can be easily seen that the exact value of α is 0.5, and the variance of the crude Monte Carlo estimator is 0.25. Applying the result of Theorem 1, an estimator of α is

$$W(\mathbf{X}) = \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(A\mathbf{X})},$$

where $f(\mathbf{x}) = \prod_{k=1}^m \frac{1}{\sqrt{2\pi}} e^{-x_k^2/2}$, and $r(\mathbf{X})$ is defined in Theorem 1 by setting $y = 0$.

It can be easily verified that $W(\mathbf{X}) \equiv 1/2$, and thus the proposed technique provides an estimator with zero variance.

2.2.2 A More Practical Example

Suppose that X_1, \dots, X_m are the values of a geometric Brownian motion evaluated at discrete time points t_1, \dots, t_m respectively, where $t_i = iT/m$ with T being the time horizon of the process. In particular, the geometric Brownian motion is specified as follows:

$$dX_t = rX_t dt + \sigma X_t dB_t,$$

where r and σ are the drift and volatility respectively, and B_t a standard Brownian motion.

Consider the estimation of the quantity

$$\alpha = E [1_{\{(X_1 + \dots + X_m)/m > y\}}].$$

Here α can be viewed as the price of an Asian-type digital option where the underlying asset price is modeled as a geometric Brownian motion.

Using the result of Theorem 1, an estimator of α is

$$W(\mathbf{X}) = \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(A\mathbf{X} + 2y\mathbf{1}_m)},$$

where $r(\mathbf{X})$ is defined in Theorem 1 by replacing y with my .

Note that in this example the density $f(\mathbf{x})$ is known and can be represented as a product of the transition densities from X_{k-1} to X_k for $k = 1, \dots, m$.

2.3 Variance Reduction

For notational ease, from now on we let $W(\mathbf{X}, y)$ denote the weight function, i.e.,

$$W(\mathbf{X}, y) = \frac{f(r(\mathbf{X}))}{f(\mathbf{X}) + f(A\mathbf{X} + 2y/m\mathbf{1}_m)}.$$

Theorem 1 shows that

$$E [h(\mathbf{X}) 1_{\{X_1 + \dots + X_m > y\}}] = E [h(r(\mathbf{X})) W(\mathbf{X}, y)], \quad (6)$$

where $r(\mathbf{X})$ is defined in Theorem 1.

Hence, to estimate the quantity α , we can use either the crude estimator $h(\mathbf{X})1_{\{X_1+\dots+X_m>y\}}$, or the reflection-based estimator $h(r(\mathbf{X}))W(\mathbf{X},y)$. Intuitively, the latter is smoother than the former, and thus may have a smaller variance. It turns out that this is indeed the case.

Theorem 2 Suppose that the conditions in Theorem 1 are satisfied, and $E[h^2(\mathbf{X})] < \infty$. Then we have

$$\text{Var}[h(r(\mathbf{X}))W(\mathbf{X},y)] \leq \text{Var}[h(\mathbf{X})1_{\{X_1+\dots+X_m>y\}}].$$

Proof. Note that by Theorem 1, we have Equation (6). Then, to prove Theorem 2, it suffices to show that

$$E[h^2(r(\mathbf{X}))W^2(\mathbf{X},y)] \leq E[h^2(\mathbf{X})1_{\{X_1+\dots+X_m>y\}}].$$

It can be easily seen that $0 \leq W(\mathbf{X},y) \leq 1$ with probability 1. Therefore,

$$E[h^2(r(\mathbf{X}))W^2(\mathbf{X},y)] \leq E[h^2(r(\mathbf{X}))W(\mathbf{X},y)] = E[h^2(\mathbf{X})1_{\{X_1+\dots+X_m>y\}}],$$

where the equality follows from directly applying Theorem 1 on $E[h^2(\mathbf{X})1_{\{X_1+\dots+X_m \geq y\}}]$. \square

Theorem 2 shows that the proposed technique always provides a variance reduction, compared to the crude Monte Carlo estimator.

In a more general setting involving change of measure as in (5), an importance-sampling (IS) estimator of α is $h(\mathbf{X})f(\mathbf{X})/g(\mathbf{X})1_{\{X_1+\dots+X_m>y\}}$, while an estimator obtained by combining importance sampling with the reflection-based technique is $h(r(\mathbf{X}))W_g(\mathbf{X},y)$, where \mathbf{X} is simulated under the probability density g , and

$$W_g(\mathbf{X},y) = \frac{f(r(\mathbf{X}))}{g(\mathbf{X}) + g(A\mathbf{X} + 2y/m\mathbf{1}_m)}.$$

Similar to the result in Theorem 2, we can show that

$$\text{Var}_g[h(r(\mathbf{X}))W_g(\mathbf{X},y)] \leq \text{Var}_g\left[h(\mathbf{X})\frac{f(\mathbf{X})}{g(\mathbf{X})}1_{\{X_1+\dots+X_m>y\}}\right].$$

In other words, the new estimator $h(r(\mathbf{X}))W_g(\mathbf{X},y)$ has a smaller variance than the IS estimator.

3 APPLICATION IN DERIVATIVE ESTIMATION

As shown in the previous subsection, the reflection-based technique can be employed as a variance-reduction tool in estimating α . But this is not the only use of the technique. It turns out the technique might also be useful in the application of derivative estimation.

Specifically, let θ denote a parameter that the random vector \mathbf{X} may depend on. We use the notation $\mathbf{X}(\theta)$ to replace \mathbf{X} in such a context. Then the quantity of interest, α , is a function of θ and y . In particular,

$$\alpha(\theta, y) = E[h(\mathbf{X}(\theta))1_{\{X_1(\theta)+\dots+X_m(\theta)>y\}}]. \quad (7)$$

One may be interested in estimating the derivatives of $\alpha(\theta, y)$ with respect to θ or y . These derivatives are of importance in many applications such as risk management. For instance, they play crucial roles in hedging of financial options. To specify the problem mathematically, we are interested in estimating $\partial\alpha(\theta, y)/\partial\theta$ and $\partial\alpha(\theta, y)/\partial y$. Here without loss of generality we assume that θ is one-dimensional. If θ is multidimensional, we may treat each dimension as a one-dimensional parameter while fixing other dimensions constants.

Various derivative estimation methods have been proposed in the literature, including finite-difference approximations, the likelihood ratio (LR) method, and the pathwise method. For comprehensive overviews, interested readers are referred to Glasserman (2004) and Fu (2006).

Among the various derivative estimation methods, the pathwise method is often the one with minimum variance when applicable. However, in the form of (7), the pathwise method cannot be applied due to the discontinuity of the indicator function. Recall that the reflection-based technique may cancel the discontinuity and replace the indicator function with a smooth function. Hence it may enable the use of the pathwise method. Specifically, using the reflection-based technique, we may rewrite the expectation as:

$$\alpha(\theta, y) = E[h(r(\mathbf{X}(\theta)))W(\mathbf{X}(\theta), y)].$$

Typically, to apply the pathwise method, we require that the integrand $h(r(\mathbf{X}(\theta)))W(\mathbf{X}(\theta), y)$ is Lipschitz continuous in (θ, y) . Here the Lipschitz continuity of a random variable $Y(\theta)$ means that there exists a random variable K such that $|Y(\theta + \Delta\theta) - Y(\theta)| \leq K|\Delta\theta|$ when $\Delta\theta$ is small enough. For similar assumptions, one is referred to Broadie and Glasserman (1996) and Liu and Hong (2011). Typically, if $\mathbf{X}(\theta)$ is Lipschitz continuous in θ , $h(\cdot)$ and $f(\cdot)$ are Lipschitz continuous functions, then it can be verified that the above assumption is satisfied. Under this assumption, we have

$$\begin{aligned} \frac{\partial \alpha(\theta, y)}{\partial \theta} &= E[\partial_{\theta}\{h(r(\mathbf{X}(\theta)))W(\mathbf{X}(\theta), y)\}], \quad \text{and} \\ \frac{\partial \alpha(\theta, y)}{\partial y} &= E[h(r(\mathbf{X}(\theta)))\partial_y W(\mathbf{X}(\theta), y)], \end{aligned}$$

where ∂_{θ} and ∂_y denote the notations of taking partial derivatives in θ and y , respectively.

To close this section, we remark the applicability domains of other derivative estimation methods in our setting. For a general function $h(\cdot)$, it is not clear how smoothed perturbation analysis can be applied; the likelihood ratio method can be applied for estimating $\partial \alpha(\theta, y)/\partial \theta$, but fails in estimating $\partial \alpha(\theta, y)/\partial y$, which may have potential applications in density estimation.

4 NUMERICAL EXPERIMENTS

To numerically illustrate the performance of the reflection-based technique, we consider the example in Section 2.2.2. We apply the proposed technique for two different purposes: variance reduction and derivative estimation.

For variance reduction, we estimate α using both the reflection-based technique and crude Monte Carlo and compare their variances. For derivative estimation, we are interested in estimating two derivatives, $\partial \alpha(X_0, y)/\partial X_0$ and $\partial \alpha(X_0, y)/\partial y$. For the former one, we compare the variance of the reflection-based estimator to that of the LR estimator, while for the latter one we compare the variance of the reflection-based estimator to that of an SPA estimator. In particular, the SPA estimator is obtained by conditioning on X_1, \dots, X_{m-1} . Because the derivations of the LR and SPA estimators are standard, they are omitted here.

During the implementation, we let $r = 5\%$, $\sigma = 20\%$, $X_0 = 100$, $T = 1$. We vary m and y to illustrate the numerical results under different circumstances. The results reported are based on a sample size of 10^5 .

Numerical results for estimators of α are presented in Table 1. The estimated standard deviations of the crude Monte Carlo (CMC) estimator and the reflection-based (Ref) estimator, and the estimated variance ratio (VR) of CMC to Ref are shown in the table. We can see that VR is always larger than 1, indicating that the reflection-based technique always provides a variance reduction. Although the reduction may not be so marked in this example, the technique may still be useful, given that the additional computational effort required is very minor.

Numerical results of the estimators of $\partial \alpha(X_0, y)/\partial X_0$ and $\partial \alpha(X_0, y)/\partial y$ are presented in Table 2, where for notational ease we let μ_1 and μ_2 denote the above two derivatives respectively. For estimating μ_1 , we show the estimated variance ratio of the LR estimator to the reflection-based estimator, while for estimating μ_2 , we show the estimated variance ratio of the SPA estimator to the reflection-based estimator. The results show that the reflection-based estimator is uniformly better than the LR and SPA estimators. When m is large and y is small, the reflection-based estimator may be significantly better.

Table 1: Performance of the estimators of α .

	y	α	standard deviation		
			CMC	Ref	VR
$m = 5$	70	0.998	0.041	0.022	3.4
	90	0.968	0.175	0.094	3.5
	100	0.561	0.496	0.326	2.3
	120	0.113	0.316	0.216	2.2
	140	0.010	0.100	0.064	2.4
$m = 10$	70	0.999	0.028	0.015	3.3
	90	0.978	0.154	0.093	2.7
	100	0.558	0.497	0.380	1.7
	120	0.096	0.295	0.228	1.7
	140	0.006	0.077	0.058	1.7
$m = 20$	70	0.999	0.025	0.015	2.5
	90	0.980	0.141	0.098	2.1
	100	0.559	0.497	0.416	1.4
	120	0.088	0.284	0.237	1.4
	140	0.005	0.069	0.056	1.5
$m = 50$	70	1.000	0.021	0.015	1.9
	90	0.981	0.135	0.110	1.5
	100	0.557	0.497	0.447	1.2
	120	0.084	0.277	0.249	1.2
	140	0.004	0.062	0.054	1.3

Table 2: Performance of the estimators of α .

	y	μ_1	μ_2	standard deviation				Variance Ratio	
				LR(μ_1)	Ref(μ_1)	SPA(μ_2)	Ref(μ_2)	μ_1	μ_2
$m = 5$	70	5×10^{-4}	7×10^{-4}	0.111	0.007	0.013	0.010	271.9	1.8
	90	0.006	0.007	0.107	0.020	0.038	0.024	29.8	2.4
	100	0.030	0.030	0.076	0.034	0.062	0.034	4.9	3.3
	120	0.014	0.012	0.050	0.031	0.036	0.026	2.5	2.0
	140	0.002	0.001	0.021	0.014	0.012	0.010	2.3	1.4
$m = 10$	70	3×10^{-4}	4×10^{-4}	0.158	0.006	0.017	0.008	834.6	4.4
	90	0.005	0.006	0.154	0.021	0.059	0.026	53.7	5.2
	100	0.032	0.032	0.112	0.046	0.115	0.046	5.9	6.2
	120	0.014	0.011	0.060	0.038	0.062	0.032	2.6	3.9
	140	0.001	9×10^{-4}	0.020	0.014	0.016	0.010	1.9	2.6
$m = 20$	70	2×10^{-4}	3×10^{-4}	0.223	0.005	0.029	0.008	1687.2	13.6
	90	0.004	0.005	0.219	0.023	0.099	0.029	87.4	11.5
	100	0.033	0.033	0.163	0.060	0.201	0.060	7.4	11.3
	120	0.013	0.011	0.076	0.045	0.100	0.037	2.8	7.1
	140	0.001	8×10^{-4}	0.022	0.015	0.025	0.011	2.0	5.2
$m = 50$	70	2×10^{-4}	2×10^{-4}	0.352	0.006	0.044	0.008	3766.5	28.0
	90	0.004	0.005	0.350	0.028	0.197	0.034	161.5	32.8
	100	0.034	0.034	0.261	0.081	0.420	0.081	10.5	27.2
	120	0.013	0.011	0.108	0.057	0.195	0.047	3.6	17.1
	140	0.001	7×10^{-4}	0.026	0.017	0.043	0.013	2.4	10.7

5 CONCLUDING REMARKS

This paper proposes a reflection-based technique for estimating an expectation which is taken under a constraint that a sum of random variables is larger than a threshold. The proposed technique always provides a variance reduction. Furthermore, it may also be useful in derivative estimation.

A straightforward extension is to replace $X_1 + \dots + X_m$ by a compound sum $X_1 + \dots + X_M$ in the formulation, where M is an integer-valued random variable independent of X_i 's. The compound sum may have applications in actuarial science and financial risk management, and is an issue of current investigation.

ACKNOWLEDGMENTS

This research is partially supported by City University of Hong Kong under start-up grant 7200172 and by the Hong Kong Research Grants Council under grant GRF 145610.

REFERENCES

- Asmussen, S., J. Blanchet, S. Juneja, and L. Rojas-Nandayapa. 2011. "Efficient simulation of tail probabilities of sums of correlated lognormals". *Annals of Operations Research* Forthcoming.
- Asmussen, S., and P. W. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York: Springer.
- Broadie, M., and P. Glasserman. 1996. "Estimating security price derivatives using simulation". *Management Science* 42:269–285.
- Fu, M. C. 2006. "Gradient estimation". In *Simulation: Handbooks in Operations Research and Management Science*, edited by S. G. Henderson and B. L. Nelson, 575–616. Amsterdam, The Netherlands: Elsevier.
- Fu, M. C., and J.-Q. Hu. 1997. *Conditional Monte Carlo, Gradience Estimation and Optimization Applications*. Boston: Kluwer Academic Publishers.
- Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. New York: Springer.
- Gong, W.-B., and Y. C. Ho. 1987. "Smoothed (conditional) perturbation analysis of discrete dynamical systems". *IEEE Transactions on Automatic Control* 32:858–866.
- Ho, Y. C., and X. R. Cao. 1983. "Perturbation analysis and optimization of queueing networks". *Journal of Optimization Theory and Applications* 40:559–582.
- Law, A. M., and W. D. Kelton. 2000. *Simulation Modeling & Analysis*. 3rd ed. New York: McGraw-Hill, Inc.
- Liu, G., and L. J. Hong. 2011. "Kernel estimation of the Greeks for options with discontinuous payoffs". *Operations Research* 59:96–108.

AUTHOR BIOGRAPHY

GUANGWU LIU is an assistant professor in the Department of Management Sciences at City University of Hong Kong. His research interests include financial risk management and stochastic simulation. His e-mail address is msgw.liu@cityu.edu.hk.