Journal of the Chinese
Statistical Association
Vol. 51, (2013) 253–273

EFFICIENT OPTION PRICING WITH IMPORTANCE SAMPLING

Chun-Chieh Chen, Cheng-Der Fuh and Huei-Wen Teng
Graduate Institute of Statisticsa, National Central University

ABSTRACT

With the rapid development of financial instruments, pricing options correctly and efficiently is a challenging task. It is known that an option price is an integral, where the integrand is a product of the payoff function of an option and a probability density function under the risk-neutral probability measure. Closed-form formulas for exotic or complicated options price rarely exist even under the standard Black-Scholes assumptions, and consequently additional numerical techniques are required. Among them, Monte Carlo approaches are flexible and easy to be adjusted for complicated payoff functions. Although Monte Carlo estimators are usually unbiased, they suffer from large variances. To tackle this problem, we first propose an importance sampling procedure, to which it is an exponential tilting measure minimizing the variance of Monte Carlo estimators. Next we apply our method to calculate the option prices, such as digital and European options.

Key words and phrases: financial options pricing, variance reduction, Monte Carlo simulation, importance sampling.
JEL classification: C15, C53, C63, G13, G17.

1. Introduction

Financial options have been traded in the exchange since 1970's, and are more and more frequently used by market participants for hedging or speculating purposes. As a result, pricing an option correctly and efficiently is a challenging task both in practice and in theory.

In this paper, we assume that the underlying asset follows the geometric Brownian motions. By option pricing theory, the option price is the discounted expected value of the payoff, and is an integral where the integrand is a product of the payoff and the probability density function under the risk-neutral probability measure. However, this expected value of the payoff may not yield a closed-form formula, particularly when the payoff is of higher dimensions or the dynamics of the underlying is complicated. Numerical methods, such as numerical integration and Monte Carlo method, have been considered as solutions.

Monte Carlo method has been regarded as the most flexible and easy to implement approach, particularly in high-dimensional cases. However, a disadvantage of Monte Carlo method is the need of a large number of trials to obtain an accurate estimate. As a result, additional variance reduction techniques are required (Boyle, Broadie, and Glasserman, 1997). Typical variance reduction methods for Monte Carlo simulation include control variate approach, antithetic variates, moment matching method, stratified sampling, Latin hypercube sampling, importance sampling, conditional Monte Carlo method, and quasi-Monte Carlo methods. See Lyuu (2002), Glasserman (2004), and references therein. In practice, it is common that a specific variance reduction technique is tailored for a certain type of options.

Importance sampling is one popular variance reduction technique. Briefly speaking, importance sampling generates samples from a sampling distribution (instead of the original distribution), and adjust the estimator by an importance sampling weight. Nonparametric importance sampling has been applied for derivative pricing in low dimensional cases for example in Givens and Raftery (1996), Zhang (1996), and Kim, Roh, and Lee (2000). The basic idea is to employ a nonparametric estimate of the optimal sampling distribution. However, nonparametric importance sampling is less applicable for high-dimensional cases, and additional techniques, such as principal component analysis, are required (Neddermeyer, 2011).

In contrast to the nonparametric importance sampling, the large deviations theory has been applied to choose a new measure. See Siegmund (1976), Glasserman and Wang (1997), Dupuis and Wang (2004), Dupuis and Wang (2005), and Glasserman, Heidelberger, and Shahabuddin (1999), for example. Others choose a new measure to minimize the variance of the importance sampling estimator. For example, Vazquez-Abad and Dufresne (1998) use a gradient estimation and stochastic approximation to minimize the variance of estimator under the sampling probability measure. Instead, Su and Fu (2000) minimize the variance under the original probability measure. For moderate deviation rare event simulations, efficient importance sampling has been studied by Do and Hall (1991), Fuh and Hu (2004), Fuh and Hu (2007) and Fuh et al. (2011). A more general account can be found in Fuh et al. (2013).

The challenging step of importance sampling is to select a proper sampling distribution. In this paper, we focus on the exponential tilting measure with a tilting parameter θ. We first propose an importance sampling considering the upper bound minimization criterion as motivated by Ross (2006). Briefly speaking, we choose a tilting parameter to minimize the upper bound of an integrand, so that the range of the estimator becomes narrower and the variance is expected to be reduced. This method reflects the basic idea of the large deviation approximations in Glasserman and Wang (1997). Although this method seems to be able to reduce variance, it is difficult to implement in general, and it does not minimize the variance directly. Therefore, we further propose an importance sampling considering the variance minimization criterion, that directly minimizes the variance of the estimator. We characterize its optimal solution for the tilting parameter using the first-order-condition. Our method does not require a gradient estimation as in Vazquez-Abad and Dufresne (1998) and Su and Fu (2000), and are hence more applicable in practice.

The rest of this paper is organized as follows. We present our importance sampling method in Section 2, and apply it for financial options pricing in Section 3. Numerical examples for European styled options are illustrated in Section 4. Section 5 concludes. Detailed mathematical treatments are deferred to the Appendix.

2. Importance sampling algorithms

Let $X = (X_1, \ldots, X_d)^t$ be a d-dimensional random vector having a probability density function (pdf) $f(x) = f(x_1, \ldots, x_d)$, where the superscript t denote the matrix transpose. Let $h(\cdot)$ be a real-valued function from $\Re^d$ to $\Re$. We are interested in the expectation of $h(X)$, namely,

$$\mu = E^P[h(X)], \tag{1}$$

where $E^P[\cdot]$ is the expectation operator, under which X has a pdf $f(x)$. Eq. (1) can be rewritten as a d-dimensional integral:

$$\mu = \int_{\Re^d} h(x)f(x)dx.$$

When the expectation in Eq. (1) does not have a closed-form formula, Monte Carlo method can be used to approximate the value. The procedures of implementing the naive Monte Carlo method for Eq. (1) is given in Algorithm 1.

Algorithm 1. *The naive Monte Carlo method to estimate μ in Eq. (1) is implemented as follows.*

(1) Generate independent samples $X^{(i)}$ having pdf $f(x)$ for $i = 1, \ldots, n$.

(2) Calculate $h^{(i)} = h(X^{(i)})$ for $i = 1, \ldots, n$.

(3) The Monte Carlo estimator for μ is $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} h^{(i)}$.

Although the Monte Carlo method is easy to implement, the Monte Carlo estimator usually suffers from large variance. The importance sampling has been known as a popular method for variance reduction. Its idea is to select a new measure Q, under which X have a pdf $q(x) = q(x_1, \ldots, x_d)$, to obtain an alternative importance sampling estimator by multiplying the original Monte Carlo estimator with an importance sampling weight. Mathematically, Eq. (1) can be rewritten as

$$\int_{\Re^d} h(x)f(x)dx = \int_{\Re^d} h(x)\frac{f(x)}{q(x)}q(x)dx = \mathrm{E}^Q\left[h(X)\frac{f(X)}{q(X)}\right], \tag{2}$$

where $\mathrm{E}^Q[\cdot]$ is the expectation operator, under which X has pdf $q(x)$, and $f(X)/q(X)$ is called the importance sampling weight (also known as the likelihood ratio or the

Radon-Nykodym derivative). The procedure of implementing an importance sampling to estimate μ in Eq. (1) based on the Eq. (2) is given in Algorithm 2.

Algorithm 2. *The importance sampling method to estimate μ in Eq. (1) based on the Eq. (2) is implemented as follows.*

(1) Generate independent samples $X^{(i)}$ having pdf $q(x)$ for $i = 1, \ldots, n$.

(2) Calculate $h^{(i)} = h(X^{(i)})\frac{f(X^{(i)})}{q(X^{(i)})}$ for $i = 1, \cdots, n$.

(3) The importance sampling estimator is $\hat{\mu}^Q = \frac{1}{n}\sum_{i=1}^{n} h^{(i)}$.

The mathematical formulation for the importance sampling is a change of measure and is conceptually easy. However, the key question is the selection of a criterion $\mathcal{C}(Q)$ and a new measure Q to optimize such criterion. This optimization problem may be conducted over an infinite dimension of functions $q(x)$ (the pdf of X under the measure Q).

Here, we consider specifically the exponential tilting measure Q_θ of P, where $\theta = (\theta_1, \ldots, \theta_d)^t$ is called the tilting parameter. Assume the moment generating function of X exists and is denoted by $\Psi(\theta)$. Let $f_\theta(x)$ be the pdf of X under the exponential tilting measure Q_θ. $f_\theta(x)$ is defined based on $f(x)$,

$$f_\theta(x) \propto \mathrm{e}^{-\theta^t x} f(x),$$

and has an explicit form as

$$f_\theta(x) = f(x)\frac{\mathrm{e}^{\theta^t x}}{M(\theta)} = f(x)\mathrm{e}^{\theta^t x - \psi(\theta)},$$

where $\psi(\theta) = \ln \Psi(\theta)$ is the cumulant function. Therefore, Eq. (1) can be rewritten as

$$\int_{\Re^d} h(x)f(x)dx = \int_{\Re^d} h(x)\frac{f(x)}{f_\theta(x)}f_\theta(x)dx = E^{Q_\theta}\left[h(X)\mathrm{e}^{-\theta^t X + \psi(\theta)}\right]. \tag{3}$$

As a result, the optimization over an infinite space of functions ($q(x)$) is reduced to the optimization over a finite-dimensional space (θ). One major advantage of using the exponential tilting measure Q_θ instead of Q is the reduction of mathematical and numerical burden. The procedure of estimating μ based on in Eq. (3) is given in Algorithm 3.

Algorithm 3. *The importance sampling method to estimate μ in Eq. (1) based on the Eq. (3) is implemented as follows.*

(1) Generate independent samples $X^{(i)}$ having the pdf $f_\theta(x)$ for $i = 1, \ldots, n$.

(2) Calculate $h^{(i)} = h(X^{(i)})e^{-\theta^t X^{(i)}+\psi(\theta)}$ for $i = 1, \cdots, n$.

(3) The estimator is $\hat{\mu}^{Q_\theta} = \frac{1}{n}\sum_{i=1}^{n} h^{(i)}$.

2.1 The upper bound minimization criterion

Motivated by Ross (2006), we first consider the upper bound minimization criterion (U criterion). It decides its optimal tilting parameter, denoted by θ^+, as presented in the following algorithm.

Algorithm 4. *We obtain θ^+ under U criterion by using the following three steps.*

1. Given θ, identify the maximum point of the estimator,

$$\chi(\theta) = \arg\max_x h(x)e^{-\theta^t x+\psi(\theta)}.$$

2. Evaluate the upper bound of $h(x)e^{-\theta^t x+\psi(\theta)}$,

$$U(\theta) = h(x)e^{-\theta^t x+\psi(\theta)}|x = \chi(\theta).$$

3. Find θ^+ to minimize this upper bound,

$$\theta^+ = \arg\min_\theta U(\theta).$$

Although it seems promising that minimizing the upper bound is a reasonable approximation to minimize the variance of the estimator, such approximation is however not always ideal. The approximation idea behind the U method is similar to the large-deviation approximation, and is particularly popular for rare events estimation. However, when the event is less rare (in the case of financial options pricing), this approximation is questionable. Second, it is very possible that the upper bound of $h(x)\mathrm{e}^{-\theta^t x+\psi(\theta)}$ is difficult to find. As illustrated in Section 3, we will show that an upper bound of the importance sampling estimator is difficult to identify even for vanilla European options.

2.2 The variance minimization criterion

Unlike the U criterion that minimizes the upper bound of the estimator, we further propose the variance minimization criterion (V criterion) that directly minimizes the variance of the importance sampling estimator. Note that the variance of the importance sampling estimator under Q_θ is

$$\begin{aligned} & Var^{Q_\theta}\left[h(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right] \\ = \ & E^{Q_\theta}\left[\left(h(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right)^2\right] - \left(E^{Q}\left[h(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right]\right)^2 \\ = \ & E^{Q_\theta}\left[\left(h(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right)^2\right] - \mu^2, \end{aligned} \tag{4}$$

where the second equality comes from the unbiasedness of the importance sampling estimator. For ease of notation, we define

$$G(\theta) = E^{Q_\theta}\left[\left(h(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right)^2\right],$$

which can be simplified under the P measure,

$$G(\theta) = E^{P}\left[h^2(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right].$$

Thus, minimizing the variance of the importance sampling estimator under Q_θ equals minimizing $G(\theta)$.

To find an optimal tilting parameter under the V criterion, denoted by θ^*, the first-order condition requires θ^* to satisfy $\frac{d}{d\theta}G(\theta)\mid_{\theta=\theta^*}= 0$. Standard calculation gives

$$\begin{aligned} \frac{d}{d\theta}G(\theta) &= \frac{d}{d\theta}E^{P}\left[h^2(X)\mathrm{e}^{-\theta^t X+\psi(\theta)}\right] \\ &= E^{P}\left[h^2(X)(-X+\psi'(\theta))\mathrm{e}^{-\theta^t X+\psi(\theta)}\right]. \end{aligned}$$

As a result, θ^* is the root of the following simplified equation,

$$\psi'(\theta) = \frac{E^{P}\left[h^2(X)X\mathrm{e}^{-\theta^t X}\right]}{E^{P}\left[h^2(X)\mathrm{e}^{-\theta^t X}\right]}.$$

We conclude the necessary condition for θ^* in Theorem 1.

Theorem 1. *The optimal tilting parameter under the variance minimization criterion, θ^*, that minimizes the variance of the importance sampling estimator in Eq. (4), satisfies the following equation.*

$$\psi'(\theta) = \frac{E^P\left[h^2(X)X\,e^{-\theta^t X}\right]}{E^P\left[h^2(X)e^{-\theta^t X}\right]}. \tag{5}$$

The root of Eq. (5) can be found by standard root finding algorithms. Note that both the numerator and the denominator in the right hand side in Eq. (5) are high-dimensional integrals and their closed-form formulas may not exist. When this is the case, we can use the Monte Carlo method to approximate these two terms. Thus, Eq. (5) is approximated by

$$\psi'(\theta) = \frac{\sum_{i=1}^n h^2(X^{(i)})X^{(i)}\mathrm{e}^{-\theta^t X^{(i)}}}{\sum_{i=1}^n h^2(X^{(i)})\mathrm{e}^{-\theta^t X^{(i)}}}. \tag{6}$$

where $X^{(i)}$ for $i = 1, \ldots, n$ are independent samples having the pdf $f(x)$. Similarly, we can apply standard root finding algorithms to obtain the root of Eq. (6).

3. Financial options pricing

Section 2 provides a general framework of importance sampling for calculating the expectation (or integral). The price of a financial option, from a mathematical aspect, equals the discounted expected payoff function, and hence is an expectation. In this section, we will apply our importance sampling method to the field of financial option pricing. In particular, we consider the case of one dimensional option pricing. Multi-dimensional case will be published in a separate paper.

Let X be a 1-dimensional normal distribution with mean η and variance σ^2, written as $X \sim N(\eta, \sigma^2)$. Then X has the pdf,

$$\phi(x; \eta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{(x-\eta)^2}{2\sigma^2}}.$$

Assume that the stock price follows the geometric Brownian motion as in the Black-Scholes assumptions. Let σ be the volatility of the stock, r be the risk-free interest rate, and S_0 be the current stock price. The stock price at time t, denoted by $S_t(X)$, can be generated by

$$S_t(X) = S_0 \exp\{(r - \frac{\sigma^2}{2})t + \sigma\sqrt{t}X\}, \tag{7}$$

where $X \sim N(0,1)$.

Let the time to maturity be T and the strike price be K. The option pricing theory gives the option price equal to the discounted expected value of the payoff function,

$$p = \mathrm{e}^{-rT} E^P[\wp(S_T(X))], \tag{8}$$

where $\wp(\cdot)$ is the payoff function.

To begin with, we derive the distribution of X under Q_θ assuming that $X \sim N(0,1)$ under the original risk-neutral measure P. Note that, under the measure P, X has the pdf

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\},$$

and its cumulant is $\psi(\theta) = \frac{1}{2}\theta^2$. Therefore, the pdf $f_\theta(x)$ of X under Q_θ is

$$\begin{aligned} f_\theta(x) &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \exp\left\{\theta x - \frac{1}{2}\theta^2\right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2 - 2x\theta + \theta^2}{2}\right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x-\theta)^2}{2}\right\}. \end{aligned}$$

As a result, we recognize that $X \sim N(\theta, 1)$ under Q_θ.

To find θ^+ in the U criterion, we simply implement Algorithm 4 with $h(x) = \mathrm{e}^{-rT}\wp(x)$ (or it does not cause any problems if the discounted factor e^{-rT} is ignored). To find θ^* in the V criterion, applying Theorem 1 and by the fact that $\psi'(\theta) = \theta$, the optimal tilting parameter θ^* satisfies

$$\theta = \frac{E^P\left[h^2(X)X\mathrm{e}^{-\theta X}\right]}{E^P\left[h^2(X)\mathrm{e}^{-\theta X}\right]} = \frac{E^P\left[\wp^2(X)X\mathrm{e}^{-\theta X}\right]}{E^P\left[\wp^2(X)\mathrm{e}^{-\theta X}\right]}$$

where $h(X) = \mathrm{e}^{-rT}\wp(X)$ with $\wp(\cdot)$ being the payoff function of the option of interest. Now, we are ready to present the solution for the optimal tilting parameter θ^* for option pricing in the following theorem.

Theorem 2. *Under the Black-Scholes assumptions, the stock price can be generated using Eq. (7). Given an option with a payoff function $\wp(\cdot)$, the optimal tilting parameter θ^* under the V criterion is the root of the following equation,*

$$\theta = \frac{E^P\left[\wp^2(X)X\,e^{-\theta X}\right]}{E^P\left[\wp^2(X)e^{-\theta X}\right]}. \tag{9}$$

Furthermore, when the numerator and the denominator in the right-hand-side (RHS) of Eq. (9) do not have closed-formed formulas, we use the Monte Carlo simulation with common random numbers for approximation. Therefore, the optimal tilting parameter θ^* in this case is the root of the following equation,

$$\theta = \frac{\sum_{i=1}^{n}\wp^2(X^{(i)})X^{(i)}\mathrm{e}^{-\theta X^{(i)}}}{\sum_{i=1}^{n}\wp^2(X^{(i)})\mathrm{e}^{-\theta X^{(i)}}}. \tag{10}$$

where $X^{(i)}$ are independent samples from $N(0,1)$ for $i = 1,\dots,n$.

In the following subsections, we illustrate how to find θ^+ in the U criterion and θ^* in the V criterion for digital and vanilla options. We define d_1 and d_2 for further derivations by

$$d_1 = \frac{\log S_0/K + (r+\sigma^2/2)T}{\sigma\sqrt{T}}, \tag{11}$$

$$d_2 = \frac{\log S_0/K + (r-\sigma^2/2)T}{\sigma\sqrt{T}}. \tag{12}$$

Note that $d_2 = d_1 - \sigma\sqrt{T}$.

3.1 Digital options

A digital call option has the payoff function,

$$\wp(x) = \mathbf{1}_{\{S_T(X)\geq K\}}(X), \tag{13}$$

where $\mathbf{1}_A(\cdot)$ is the indicator function to indicate the support set A. The call price of the digital option has a closed-form formula,

$$p = \mathrm{e}^{-rT}E^P\left[\mathbf{1}_{\{S_T(X)\geq K\}}(X)\right] = \mathrm{e}^{-rT}\Phi(d_2), \tag{14}$$

where $\Phi(\cdot)$ is the cumulative probability density function of the standard normal distribution, and d_2 is in Eq. (12).

Because $\{x : S_T(x) \geq K\} = \{x : x \geq -d_2\}$, where d_2 is defined in Eq. (12), we have

$$E^P\left[\mathbf{1}_{\{S_T(X)\geq K\}}(X)\right] = E^{Q_\theta}\left[\mathbf{1}_{\{X\geq -d_2\}}(X)\mathrm{e}^{-\theta X+\frac{1}{2}\theta^2}\right].$$

The importance sampling estimator becomes $\mathbf{1}_{\{X\geq -d_2\}}(X)\mathrm{e}^{-\theta X+\frac{1}{2}\theta^2}$ under Q_θ, and its upper bound is simply

$$U(\theta) = \max_x \mathbf{1}_{\{x\geq -d_2\}}\mathrm{e}^{-\theta x+\theta^2/2} = \mathrm{e}^{\theta d_2+\frac{1}{2}\theta^2},$$

conditional on $\theta > 0$. Therefore, θ^+ in the U criterion is

$$\theta^+ = \arg\min_\theta U(\theta) = -d_2.$$

We proceed to find θ^* in the V criterion. Theorem 2 has shown that the optimal titling parameter is the root of Eq. (9). First of all, the numerator in the RHS of Eq. (9) is

$$E^P\left[\wp^2(X)X\mathrm{e}^{-\theta X}\right] = E^P\left[\mathbf{1}^2_{\{S_T(X)\geq K\}}(X)X\mathrm{e}^{-\theta X}\right] = \mathrm{e}^{\theta^2/2}\left(\phi(d_2-\theta) - \theta\Phi(d_2-\theta)\right),$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ is the pdf and the cumulative probability density function of the standard normal distribution. Similarly, the denominator in the RHS of Eq. (9) is

$$E^P\left[\wp^2(X)\mathrm{e}^{-\theta X}\right] = E^P\left[\mathbf{1}^2_{\{S_T(X)\geq K\}}(X)\mathrm{e}^{-\theta X}\right] = \mathrm{e}^{\theta^2/2}\Phi(d_2-\theta).$$

As a result, θ^* is the root of

$$\theta = \frac{\phi(d_2-\theta) - \theta\Phi(d_2-\theta)}{\Phi(d_2-\theta)}.$$

3.2 European options

A European call option has the payoff function

$$\wp(X) = (S_T(X) - K)^+, \tag{15}$$

where $(\cdot)^+$ denotes the positive function. The price of European call option has a closed-form formula,

$$p = \mathrm{e}^{-rT}E^P[(S_T(X)-K)^+] = S_0\Phi(d_1) - K\mathrm{e}^{-rT}\Phi(d_2), \tag{16}$$

where d_1 and d_2 are in Eqs. (11) and (12), respectively.

Under the U criterion, standard calculation yields that

$$\chi(\theta) = \frac{ln[\frac{-\theta K}{S_0(\sigma\sqrt{T}-\theta)}] - (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}},$$

conditional on $\theta > \sigma\sqrt{T}$ and θ^* is the root of the following equation,

$$\theta = -d_2 + \frac{1}{\sigma\sqrt{T}} ln(\frac{-\theta}{\sigma\sqrt{T}-\theta}).$$

We solve this equation using standard root finding algorithms conditional on that θ is larger than $\sigma\sqrt{T}$. Detailed mathematical treatments are deferred to Appendix **??**. Although the payoff of the European option is simple, the detailed mathematical treatment is rather tedious. We remark that in general θ^+ under the U criterion is difficult to obtain, either when the payoff or the dynamics of the underlying is complicated.

As for the V criterion, θ^* is the root of Eq. (9) in Theorem 2, where numerator in the RHS of Eq. (9) equals

$$\begin{aligned}
& \frac{1}{\sqrt{2\pi}} S_0^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T + \frac{1}{2}(\theta - 2\sigma\sqrt{T})^2} \\
& \times \left[\mathrm{e}^{-\frac{1}{2}[-d_2+(\theta-2\sigma\sqrt{T})]^2} - (\theta - 2\sigma\sqrt{T})\sqrt{2\pi}(1 - \Phi(-d_2 + (\theta - 2\sigma\sqrt{T})))\right] \\
- & \frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T + \frac{1}{2}(\theta - \sigma\sqrt{T})^2} \\
& \times \left[\mathrm{e}^{-\frac{1}{2}[-d_2+(\theta-\sigma\sqrt{T})]^2} - (\theta - \sigma\sqrt{T})\sqrt{2\pi}(1 - \Phi(-d_2 + (\theta - \sigma\sqrt{T})))\right] \\
+ & \frac{1}{\sqrt{2\pi}} K^2 \mathrm{e}^{\frac{1}{2}\theta^2} \left[\mathrm{e}^{-\frac{1}{2}(-d_2+\theta)^2} - \theta\sqrt{2\pi}(1 - \Phi(-d_2 + \theta))\right],
\end{aligned}$$

and the denominator in the RHS of Eq. (9) equals

$$\begin{aligned}
& S_0^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T + \frac{1}{2}(\theta - 2\sigma\sqrt{T})^2} [1 - \Phi(-d_2 + (\theta - 2\sigma\sqrt{T}))] \\
- \; & 2S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T + \frac{1}{2}(\theta - \sigma\sqrt{T})^2} [1 - \Phi(-d_2 + (\theta - \sigma\sqrt{T}))] \\
+ \; & K^2 \mathrm{e}^{\frac{1}{2}\theta^2} [1 - \Phi(-d_2 + \theta)].
\end{aligned}$$

Detailed calculation is deferred to Appendix.

4. Numerical results

We here consider options involved with just one underlying asset. The benchmark value is obtained by the naive Monte Carlo simulation. Each estimate is obtained by using 10,000 trials. Tables 1 and 2 report numerical results for digital and European options, respectively. The relative efficiency (RE) is the ratio of the variance of the naive Monte Carlo method over that of corresponding importance sampling method, and hence can be used to measure the efficiency of a proposed variance reduction method. It is shown from these two examples that RE is larger than one in all cases when applicable, and RE is higher for deeper out-of-the-money options. It is not surprising that RE using the V criterion is higher than that using the U criterion, because the importance sampling using the U criterion minimizes the variance of the importance sampling estimators.

Table 1 Digital options

We compare the mean and the standard error (SE) of the naive Monte Carlo method (MC), and the importance sampling using the upper bound minimization criterion (U) and the variance minimization criterion (V) for a variety of strike prices K. RE is the relative efficiency, which the ratio of the variance of the naive Monte Carlo method over that of corresponding the importance sampling method. Each estimate is obtained using a sample size of 10,000. The benchmark value is calculated using Eq. (14). When the strike price is less than or equal to the current stock price, d_2 is equal to or larger than zero, and the importance sampling estimator using the U criterion does not produce further variance reduction. Parameters: initial stock price $S_0 = 42$, risk-free rate $r = 0.1$, volatility $\sigma = 0.2$, and time to maturity $T = 0.5$.

		MC		U				V			
K	Benchmark	Mean	SE	Mean	SE	θ^+	RE	Mean	SE	θ^*	RE
34	0.915	0.914	0.002					0.915	0.002	0.047	1.053
36	0.870	0.869	0.003					0.875	0.002	0.099	1.087
38	0.798	0.798	0.004					0.799	0.003	0.182	1.095
40	0.699	0.701	0.004					0.699	0.004	0.301	1.150
42	0.582	0.586	0.005					0.581	0.004	0.455	1.222
44	0.458	0.467	0.005	0.449	0.005	0.046	1.037	0.448	0.004	0.640	1.347
46	0.342	0.336	0.005	0.342	0.004	0.360	1.283	0.349	0.003	0.847	1.474
48	0.242	0.246	0.004	0.240	0.003	0.661	1.549	0.243	0.002	1.068	1.713
50	0.163	0.160	0.004	0.162	0.002	0.950	1.829	0.162	0.002	1.297	1.950
52	0.104	0.099	0.003	0.105	0.001	1.227	2.171	0.105	0.001	1.529	2.253

Table 2 European options

We compare the mean and the standard error (SE) of the naive Monte Carlo method (MC), and the importance sampling using the upper bound minimization criterion (U) and the variance minimization criterion (V) for a variety of strike prices K. RE is the relative efficiency, which the ratio of the variance of the naive Monte Carlo method over that of corresponding the importance sampling method. Each estimate is obtained using a sample size of 10,000. The benchmark value is calculated using Eq. (16). Parameters: initial stock price $S_0 = 42$, risk-free rate $r = 0.1$, volatility $\sigma = 0.2$, and time to maturity $T = 0.5$.

		MC		U				V			
K	Benchmark	Mean	SE	Mean	SE	θ^+	RE	Mean	SE	θ^*	RE
34	9.724	9.676	0.059	9.732	0.029	0.316	2.005	9.711	0.017	0.573	3.409
36	7.934	7.915	0.057	7.876	0.030	0.364	1.910	7.938	0.019	0.666	3.005
38	6.261	6.208	0.053	6.321	0.027	0.428	1.956	6.283	0.019	0.778	2.887
40	4.759	4.733	0.050	4.744	0.025	0.513	1.993	4.781	0.017	0.909	2.855
42	3.477	3.503	0.045	3.475	0.021	0.625	2.168	3.483	0.015	1.057	2.937
44	2.437	2.449	0.039	2.436	0.016	0.767	2.378	2.433	0.012	1.220	3.089
46	1.639	1.635	0.032	1.623	0.012	0.938	2.674	1.638	0.010	1.397	3.353
48	1.059	1.073	0.026	1.053	0.008	1.133	3.141	1.067	0.007	1.583	3.782
50	0.658	0.639	0.020	0.656	0.005	1.343	3.646	0.657	0.005	1.777	4.228
52	0.394	0.402	0.015	0.397	0.004	1.563	4.337	0.394	0.003	1.975	4.895

5. Conclusion

Pricing options correctly and efficiently is a challenging task in the growing financial derivatives market. Closed-form formulas for option prices rarely exist, even under the simplest Black-Scholes assumptions. Consequently, additional numerical methods to calculate an expectation (or an integral) are required. Among them, Monte Carlo methods are flexible, easy to implement, and do not suffer from the curse of dimensionality. To reduce the variance of the Monte Carlo estimator, we consider the importance sampling procedure by employing an exponential titling measure, and determine the optimal titling parameter by minimizing the upper bound and the variance of the Monte Carlo estimator.

To minimize the variance of the Monte Carlo estimator, we characterize the optimal parameters using the first-order-condition of the original optimization problem. As a result, our method does not need to calculate higher-order derivatives, and are more applicable than the method proposed by Su and Fu (2000). For example, our method can be directly applied to calculate multi-variate options and price sensitivities once unbiased formulas for Greeks exist, cf. Lyuu and Teng (2011). Numerical experiments confirm all these above claims, which will be published in a separate paper.

It is worth mentioning that our proposed importance sampling method is not just limited to the field of financial options pricing, but can be applied to other fields, as long as the quantity of interest is an expectation. As for the future work, we are interested in generalizing our method to non-Gaussian cases, for example, either the marginal distributions or the dependence structure are non-Gaussian.

Acknowledgements

The research of Cheng-Der Fuh is supported in part by grants NSC 100-2118- M-008-002-MY3 and NSC 101-3113-P-008-005, and the research of Huei-Wen Teng is supported in part by grants NSC 101-2118-M-008-004 and NSC 101-3113-P-008-005.

Appendix A Importance sampling for a European option using the U criterion

First, notice that

$$E^P[(S(T)-K)^+] = E^P[I_{\{S_T(X)\geq K\}}(S_T(X)-K)] = E^Q[I_{\{X\geq -d_2\}}(S_T(X)-K)\mathrm{e}^{-\theta X+\frac{1}{2}\theta^2}].$$

To find the upper bound of the estimator, $I_{\{X\geq -d_2\}}(S_T(X)-K)\mathrm{e}^{-\theta X+\frac{1}{2}\theta^2}$, we only need to consider $(S_T(X)-K)\mathrm{e}^{-\theta X+\frac{1}{2}\theta^2}$ because $I_{\{X\geq -d_2\}}$ is an indicator. Note that the first order partial derivative with respective to x is

$$\begin{aligned}
& \frac{d}{dx}(S_T(x)-K)\mathrm{e}^{-\theta x+\frac{1}{2}\theta^2} \\
= \; & \frac{d}{dx}\left[S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}x-\theta x+\frac{1}{2}\theta^2} - K\mathrm{e}^{-\theta x+\frac{1}{2}\theta^2}\right] \\
= \; & \frac{d}{dx}\left[S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}\theta^2+(\sigma\sqrt{T}-\theta)x} - K\mathrm{e}^{-\theta x+\frac{1}{2}\theta^2}\right] \\
= \; & S_0(\sigma\sqrt{T}-\theta)\mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}\theta^2+(\sigma\sqrt{T}-\theta)x} + \theta K\mathrm{e}^{-\theta x+\frac{1}{2}\theta^2}.
\end{aligned} \tag{17}$$

If θ needs to be less than $\sigma\sqrt{T}$, the first-order derivative with respect to x is always positive and a maximum would not exist. Therefore, the maximum point must satisfy that the first-order-condition in Eq. (17), and equals

$$\chi(\theta) = \frac{\ln[\frac{-\theta K}{S_0(\sigma\sqrt{T}-\theta)}] - (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = -d_2 + \frac{1}{\sigma\sqrt{T}}\ln(\frac{-\theta}{\sigma\sqrt{T}-\theta}), \tag{18}$$

conditional on $\theta > \sigma\sqrt{T}$. Note that

$$\begin{aligned}
\frac{d^2}{dx^2}(S_T(x) - K)\mathrm{e}^{-\theta x + \frac{1}{2}\theta^2} &= S_0(\sigma\sqrt{T}-\theta)^2\mathrm{e}^{(r-\frac{\sigma^2}{2})T + \frac{1}{2}\theta^2 + (\sigma\sqrt{T}-\theta)x} - \theta^2 K\mathrm{e}^{-\theta x + \frac{1}{2}\theta^2} \\
&= [S_0(\sigma\sqrt{T}-\theta)^2\mathrm{e}^{(r-\frac{\sigma^2}{2})T + \sigma\sqrt{T}x} - \theta^2 K]\mathrm{e}^{-\theta x + \frac{1}{2}\theta^2}.
\end{aligned}$$

Plugging $\chi(\theta)$ into the above formula, the above formula becomes

$$-\theta K\sigma\sqrt{T}\mathrm{e}^{-\theta x + \theta^2/2} < 0.$$

Thus, $\chi(\theta)$ in Eq. (18) is indeed the maximum point. The upper bound of the estimator is

$$U(\theta) = [S_T(x) - K]\mathrm{e}^{-\theta x + \frac{1}{2}\theta^2}|_{x=\chi(\theta)}, \tag{19}$$

for $\theta > \sigma\sqrt{T}$.

To find θ minimize Eq. (19), the first-order derivative of $U(\theta)$ with respect to θ equals

$$\begin{aligned}
&\frac{d}{d\theta}\left[(S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T + \sigma\sqrt{T}\chi(\theta)} - K)\mathrm{e}^{-\theta\chi(\theta) + \frac{1}{2}\theta^2}\right] \\
=\ &\frac{d}{d\theta}\left[S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T + \frac{1}{2}\theta^2 + (\sigma\sqrt{T}-\theta)\chi(\theta)} - K\mathrm{e}^{-\theta\chi(\theta) + \frac{1}{2}\theta^2}\right] \\
=\ &S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T + \frac{1}{2}\theta^2 + (\sigma\sqrt{T}-\text{ theta})\chi(\theta)}[\theta - \chi(\theta) + (\sigma\sqrt{T}-\theta)\chi'(\theta)] \\
&- K\mathrm{e}^{-\theta\chi(\theta) + \frac{1}{2}\theta^2}[-\chi(\theta) - \theta\chi'(\theta) + \theta],
\end{aligned}$$

where

$$\chi'(\theta) = \frac{d\chi(\theta)}{d\theta} = \frac{1}{\theta(\sigma\sqrt{T}-\theta)}.$$

To satisfy the first-order-condition, standard calculation shows that θ is the root of

$$\chi(\theta) = \theta,$$

or, equivalently,

$$-d_2 + \frac{1}{\sigma\sqrt{T}} \ln(\frac{-\theta}{\sigma\sqrt{T}-\theta}) = \theta.$$

Appendix B Importance sampling for a European option using the V criterion

We present detailed calculation for Eq. (9) in the following. The numerator in the RHS of Eq. (9) equals

$$\begin{aligned}
& E^P[((S_T(X)-K)^+)^2 X\mathrm{e}^{-\theta X}] \\
= & \int_{-\infty}^{\infty} ((S_T(x)-K)^+)^2 x\mathrm{e}^{-\theta x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} dx \\
= & \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} (S_T(x)-K)^2 x\mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} (S_0\mathrm{e}^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}x} - K)^2 x\mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T} \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}x^2-\theta x+2\sigma\sqrt{T}x} dx \\
& - \frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T} \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}x^2-\theta x+\sigma\sqrt{T}x} dx \\
& + \frac{1}{\sqrt{2\pi}} K^2 \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-2\sigma\sqrt{T})^2} \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}[x+(\theta-2\sigma\sqrt{T})]^2} dx \\
& - \frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-\sigma\sqrt{T})^2} \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}[x+(\theta-\sigma\sqrt{T})]^2} dx \\
& + \frac{1}{\sqrt{2\pi}} K^2 \mathrm{e}^{\frac{1}{2}\theta^2} \int_{-d_2}^{\infty} x\mathrm{e}^{-\frac{1}{2}(x+\theta)^2} dx \\
= & \frac{1}{\sqrt{2\pi}} (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-2\sigma\sqrt{T})^2} \\
& \times \left[\mathrm{e}^{-\frac{1}{2}[-d_2+(\theta-2\sigma\sqrt{T})]^2} - (\theta-2\sigma\sqrt{T})\sqrt{2\pi}(1-\Phi(-d_2+(\theta-2\sigma\sqrt{T})))\right] \\
& - \frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-\sigma\sqrt{T})^2} \\
& \times \left[\mathrm{e}^{-\frac{1}{2}[-d_2+(\theta-\sigma\sqrt{T})]^2} - (\theta-\sigma\sqrt{T})\sqrt{2\pi}(1-\Phi(-d_2+(\theta-\sigma\sqrt{T})))\right] \\
& + \frac{1}{\sqrt{2\pi}} K^2 \mathrm{e}^{\frac{1}{2}\theta^2} \left[\mathrm{e}^{-\frac{1}{2}(-d_2+\theta)^2} - \theta\sqrt{2\pi}(1-\Phi(-d_2+\theta))\right].
\end{aligned}$$

The denominator in the RHS of Eq. (9) equals

$$
\begin{aligned}
& E_P[((S_T(X)-K)^+)^2 \mathrm{e}^{-\theta X}] \\
= & \int_{-\infty}^{\infty} ((S_T(x)-K)^+)^2 \mathrm{e}^{-\theta x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} dx \\
= & \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} (S_T(x)-K)^2 \mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} (S_0 \mathrm{e}^{(r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}x} - K)^2 \mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T} \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2-\theta x+2\sigma\sqrt{T}x} dx \\
& -\frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T} \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2-\theta x+\sigma\sqrt{T}x} dx \\
& +\frac{1}{\sqrt{2\pi}} K^2 \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2-\theta x} dx \\
= & \frac{1}{\sqrt{2\pi}} (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-2\sigma\sqrt{T})^2} \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}[x+(\theta-2\sigma\sqrt{T})]^2} dx \\
& -\frac{2}{\sqrt{2\pi}} S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-\sigma\sqrt{T})^2} \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}[x+(\theta-\sigma\sqrt{T})]^2} dx \\
& +\frac{1}{\sqrt{2\pi}} K^2 \mathrm{e}^{\frac{1}{2}\theta^2} \int_{-d_2}^{\infty} \mathrm{e}^{-\frac{1}{2}(x+\theta)^2} dx \\
= & (S_0)^2 \mathrm{e}^{2(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-2\sigma\sqrt{T})^2} [1-\Phi(-d_2+(\theta-2\sigma\sqrt{T}))] \\
& -2S_0 K \mathrm{e}^{(r-\frac{\sigma^2}{2})T+\frac{1}{2}(\theta-\sigma\sqrt{T})^2} [1-\Phi(-d_2+(\theta-\sigma\sqrt{T}))] \\
& +K^2 \mathrm{e}^{\frac{1}{2}\theta^2} [1-\Phi(-d_2+\theta)].
\end{aligned}
$$

References

Boyle, P., Broadie, M. and Glasserman P. (1997). Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, **21**, 1267-1321.

Do, K.-A. and Hall, P. (1991). On importance resampling for the bootstrap. *Biometrika*, **78**, 161-167.

Dupuis, P. and Wang, H. (2004). Importance sampling, large deviations, and differential games. *Stochastics and Stochastics Reports*, **76**, 481-508.

Dupuis, P. and Wang, H. (2005). Dynamic importance sampling for uniformly recurrent markov chains. *The Annals of Applied Probability*, **15**, 1-38.

Fuh, C.-D. and Hu, I. (2004). Efficient importance sampling for events of moderate deviations with applications. *Biometrika*, **91**, 471-490.

Fuh, C.-D. and Hu, I. (2007). Estimation in hidden markov models via efficient importance sampling. *Bernoulli*, **13**, 492-513.

Fuh, C.-D., Hu, I., Hsu, Y.-H., and Wang, R.-H. (2011). Efficient simulation of value at risk with heavy-tailed risk factors. *Operations Research*, **59**, 1395-1406.

Fuh, C.-D., Teng, H.-W. and Wang R.-H. (2013). Efficient importance sampling for rare event simulation with applications. Technical Report.

Givens, G. H. and Raftery, A. E. (1996). Local adaptive importance sampling for multivariate densities with strong nonlinear relationships. *Journal of the American Statistical Association*, **91**, 132-141.

Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering. New York: Springer.

Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999). Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance*, **9**, 117-152.

Glasserman, P. and Wang, Y. (1997). Counterexamples in importance sampling for large deviations probabilities. *The Annals of Statistics*, **7**, 731-746.

Kim, Y. B., Roh, D. S., and Lee, M. Y. (2000). Nonparametric adaptive importance sampling for rare event simulation. *Proceedings of the 2000 Winter Simulation Conference*, **1**,767-772.

Lyuu, Y.-D. (2002). Financial Engineering and Computation: Principles, Mathematics, Algorithms. Cambridge: Cambridge University Press.

Lyuu, Y.-D. and Teng, H.-W. (2011). Unbiased and efficient Greeks of financial options. *Finance and Stochastics*, **15**, 141-181.

Neddermeyer, J. (2011) Non-parametric partial importance sampling for financial derivative pricing. *Quantitative Finance*, **11**, 1193-1206.

Ross, S. M. (2006). Simulation. New York: Academic Press.

Siegmund, D. (1976). Importance sampling in the Monte Carlo study of sequential test. *The Annals of Statistics*, **4**, 673-684.

Su, Y. and Fu, M. C. (2000). Importance sampling in derivative securities pricing. *Proceedings of the 2000 Winter Simulation Conference*, **1**, 587-596.

Vazquez-Abad, F. J. and Dufresne, D. (1998). Accelerated simulation for pricing Asian options. *Proceedings of the 1998 Winter Simulation Conference*, **2**, 1493 - 1500.

Zhang, P. (1996). Nonparametric Importance Sampling. *Journal of the American Statistical Association*, **91**, 1245-1253.

[Received January 2013; accepted February 2013.]

Journal of the Chinese Statistical Association
Vol. 51, (2013) 253–273

利用重點抽樣的有效率選擇權訂價

陳醇潔　　傅承德　　鄧惠文

國立中央大學統計研究所

摘　要

隨著金融商品迅速發展, 如何正確且有效率的計算選擇權價格是一個具有挑戰性的問題。我們已知選擇權價格是選擇權連結商品的收益函數 (payoff functions) 的期望值, 其中機率密度函數是風險中立測度。奇異選擇權 (exotic options) 或複雜的選擇權通常沒有封閉解, 就算在布萊克-肖爾斯 (Black-Scholes) 假設下也是如此, 因此需要使用數值方法。其中, 蒙地卡羅近似是一個合適的方法, 且對於複雜的收益函數也很容易做調整。雖然蒙地卡羅估計量通常是不偏的, 但卻有較大的變異數。爲了解決這個問題, 我們提出了一個重點抽樣的方法, 用指數平移測度來極小化蒙地卡羅估計量的變異數。接著我們利用這個方法計算數位選擇權 (digital options) 和歐式選擇權價格作爲例子。

關鍵詞: 財務選擇權定價, 變異數縮減, 蒙地卡羅模擬, 重點抽樣。

JEL classification: C15, C53, C63, G13, G17.