

# A general method for debiasing a Monte Carlo estimator

Don McLeish

Department of Statistics and Actuarial Science,  
University of Waterloo  
Tel.: +519-8884567 Ext. 35534  
dlmcleis@uwaterloo.ca

## Abstract

Consider a stochastic process  $X_n$ ,  $n = 0, 1, 2, \dots$  such that  $EX_n \rightarrow x_\infty$  as  $n \rightarrow \infty$ . The sequence  $\{X_n\}$  may be a deterministic one, obtained by using a numerical integration scheme, or obtained from Monte-Carlo methods involving an approximation to an integral, or a Newton-Raphson iteration to approximate the root of an equation but we will assume that we can sample from the distribution of  $X_1, X_2, \dots, X_m$  for finite  $m$ . We propose a scheme for unbiased estimation of the limiting value  $x_\infty$ , together with estimates of standard error and apply this to examples including numerical integrals, root-finding and option pricing in a Heston Stochastic Volatility model.

**Keywords and phrases:** Monte Carlo simulation, unbiased estimates, numerical integration, finance, stochastic volatility model

## 1 Introduction.

Suppose  $X_n$ ,  $n = 0, 1, 2, \dots$  is a stochastic process such that  $E(X_n) \rightarrow x_\infty$  as  $n \rightarrow \infty$ . Typically  $X_0 = x_0$  is a deterministic seed or arbitrary value initiating the iteration and we are interested in the limiting value  $x_\infty$ . The sequence  $\{X_n\}$  may be deterministic, obtained by using a numerical integration scheme to approximate an integral, or a Newton-Raphson scheme to approximate the root of an equation. It may be a ratio estimator estimating a population ratio or the result of a stochastic or deterministic approximation to a root or maximum. In general we will only assume that it is possible to sample from the distribution of the stochastic process for a finite period, i.e. sample  $X_1, X_2, \dots, X_m$  for fixed  $m$ .

A common argument advanced in favour of the use Monte Carlo (MC) methods as an alternative to numerical ones is that the MC estimator is usually unbiased with estimable variance. By increasing the sample size we are assured by unbiasedness that the estimator is consistent and we can produce, for any sample size, a standard error of the estimator. The statistical argument is advanced against the use of numerical methods that they do not offer easily obtained estimates of error. The purpose of this brief note is to show that this argument is flawed; generally any consistent sequence of estimators can be easily rendered unbiased and an error estimate is easily achieved. We do not attempt to merely reduce the bias, but by introducing randomization into the sequence, to completely eliminate it. The price we pay is an additional randomization inserted into the sequence and a possible increase in the mean squared error (MSE).

## 2 The debiased sequence and its variance

Suppose  $N$  is a random variable, independent of the sequence  $\{X_n, n = 0, 1, 2, \dots\}$  taking finite *non-negative* integer values. Suppose  $Q_n = P(N \geq n) > 0$  for all  $n = 1, 2, \dots$ . For a given sequence  $X_n, n = 0, 1, 2, \dots$  we define the first backward difference as  $\nabla X_n = X_n - X_{n-1}$ . Define the random variable

$$\begin{aligned} Y &= x_0 + \sum_{n=1}^N \frac{\nabla X_n}{Q_n} \\ &= x_0 + \sum_{n=1}^{\infty} \nabla X_n \frac{I(n \leq N)}{Q_n} \end{aligned}$$

This can be written in the more general form

$$Y = X_0 + \sum_{n=1}^{\infty} \nabla X_n F_n \tag{1}$$

where  $F_n, n = 1, 2, \dots$  are random variables with  $E[F_n | X_n, X_{n-1}] = 1$  and for some value of  $N < \infty$  we have  $F_i = 0$  for  $i > N$ . We will show that

$Y$  is an unbiased estimator of the limit  $x_\infty$  with easily estimated standard error. It is obviously unbiased provided that it is integrable and one can interchange the sum and the expected value since with  $\mathcal{F} = \sigma(X_1, X_2, \dots)$ ,  $E\left(x_0 + \sum_{n=1}^{\infty} \nabla X_n \frac{I(n \leq N)}{Q_n}\right) = E\left[x_0 + \sum_{n=1}^{\infty} \nabla X_n E\left(\frac{I(n \leq N)}{Q_n} | \mathcal{F}\right)\right] = x_0 + \sum_{n=1}^{\infty} \nabla X_n = x_\infty$ .

Assume for the calculation of the variance that  $E(X_n - x_\infty)^2 \rightarrow 0$  as  $n \rightarrow \infty$ .

Then

$$\begin{aligned}
\sigma_Y^2 &= \text{var}(Y) = E[\text{var}(Y|\mathcal{F})] + \text{var}(E(Y|\mathcal{F})) \\
&= E[\text{var}(Y|\mathcal{F})] + \text{var}(x_\infty) \\
&= E\left(\sum_{n=1}^{\infty} \frac{(\nabla X_n)^2}{Q_n^2} Q_n(1 - Q_n) + 2 \sum_{n=2}^{\infty} \sum_{j=1}^{n-1} \frac{\nabla X_n \nabla X_j}{Q_n Q_j} Q_n(1 - Q_j)\right) \\
&= E\left(\sum_{n=1}^{\infty} \frac{(\nabla X_n)^2}{Q_n} (1 - Q_n) + 2 \sum_{j=1}^{\infty} \sum_{n=j+1}^{\infty} \frac{\nabla X_n \nabla X_j}{Q_j} (1 - Q_j)\right) \quad (2) \\
&= E\left(\sum_{n=1}^{\infty} \frac{(\nabla X_n)^2}{Q_n} (1 - Q_n) + 2 \sum_{j=1}^{\infty} \frac{(x_\infty - X_j) \nabla X_j}{Q_j} (1 - Q_j)\right) \\
&= E\left(\sum_{n=1}^{\infty} [(\nabla X_n)^2 + 2(x_\infty - X_n) \nabla X_n] \left(\frac{1 - Q_n}{Q_n}\right)\right) \\
&= \sum_{n=1}^{\infty} E\left(\frac{\nabla X_n [2x_\infty - X_n - X_{n-1}]}{Q_n} - \sum_{n=1}^{\infty} \nabla X_n [2x_\infty - X_n - X_{n-1}]\right) \\
&= \sum_{n=1}^{\infty} E\left(\frac{2x_\infty \nabla X_n - \nabla X_n^2}{Q_n} - \sum_{n=1}^{\infty} (2x_\infty \nabla X_n - \nabla X_n^2)\right) \\
&= \sum_{n=1}^{\infty} E\left(\frac{2x_\infty \nabla X_n - \nabla X_n^2}{Q_n}\right) - (2x_\infty(x_\infty - x_0) - (x_\infty^2 - x_0^2)) \\
&= \sum_{n=1}^{\infty} E\left(\frac{2x_\infty \nabla X_n - \nabla X_n^2}{Q_n}\right) - (x_\infty - x_0)^2 \\
&= \sum_{n=1}^{\infty} \frac{2x_\infty \nabla \mu_n - \nabla(\sigma_n^2 + \mu_n^2)}{Q_n} - (x_\infty - x_0)^2 \\
&= \sum_{n=1}^{\infty} \frac{2(x_\infty - \xi_j) \nabla \mu_j - \nabla \sigma_j^2}{Q_n} - (x_\infty - x_0)^2 \quad (3)
\end{aligned}$$

where  $E[\nabla X_n^2] = EX_n^2 - EX_{n-1}^2 = \sigma_n^2 + \mu_n^2 - \sigma_{n-1}^2 - \mu_{n-1}^2 = \nabla(\sigma_n^2 + \mu_n^2)$  and  $\xi_n = \frac{\mu_n + \mu_{n-1}}{2}$ . Then  $\sigma_Y^2$ , as given in (2) can be unbiasedly estimated using

$$\hat{\sigma}_Y^2 = \sum_{n=1}^N \frac{(\nabla X_n)^2}{Q_n^2} (1 - Q_n) + 2 \sum_{j=1}^N \sum_{n=j+1}^N \frac{\nabla X_n \nabla X_j}{Q_j^2 Q_n} (1 - Q_j).$$

Suppose  $N \geq n_s$  with probability one. Then the average over a large number of values of  $Y$ , i.e. a large number of values of  $N$ ,  $\{X_i\}$  takes the form

$$\begin{aligned}\bar{Y} &= \bar{X}_0 + \sum_{n=1}^{N_{\max}} \overline{\nabla X_n} \frac{\text{freq}(n \leq N)}{Q_n}, \\ &= \bar{X}_{n_s} + \sum_{n=n_s+1}^{N_{\max}} \overline{\nabla X_n} \frac{\text{freq}(n \leq N)}{Q_n}\end{aligned}$$

where  $N_{\max}$  is the largest observed value of  $N$  and  $\overline{\nabla X_n}$  denotes the average of the observed values of  $\nabla X_n$  for which the corresponding  $N \geq n$ . This takes the form of (1) with term  $F_n = \frac{\text{freq}(n \leq N)}{Q_n}$  obtained from simulating values of  $N$  alone. Suppose we wish to minimize the variance subject to a constraint on the expected value of  $N$ , i.e.

$$\begin{aligned}\min_Q \left\{ \sum_{n=1}^{\infty} \frac{2x_{\infty} \nabla \mu_n - \nabla(\sigma_n^2 + \mu_n^2)}{Q_n} - (x_{\infty} - x_0)^2 \right\} \\ \text{subject to } \sum_n Q_n = \mu.\end{aligned}$$

Of course we also require that  $Q_n$  is non-increasing and positive but for the moment we will ignore these additional constraints. Then we obtain, on differentiating the Lagrangian with respect to  $Q_n$ , with  $\xi_n = \frac{\mu_n + \mu_{n-1}}{2}$ ,

$$Q_n \sim c \sqrt{|2(x_{\infty} - \xi_n) \nabla \mu_n - \nabla \sigma_n^2|} \quad (4)$$

where  $c$  is determined by the constraint  $\sum_n Q_n = \mu$  and the minimum variance is finite provided that

$$\sum_{j=1}^{\infty} \sqrt{|2(x_{\infty} - \xi_j) \nabla \mu_j - \nabla \sigma_j^2|} < \infty.$$

While this is not entirely practical because it requires  $x_{\infty}$ , it is common to have some information on the rate of convergence of the sequence that can be used to design an asymptotically appropriate sequence  $Q_n$ . For example if we believe  $x_{\infty} - X_n \sim ar^n$  for some  $r < 1$  and  $a$ , then we might choose  $Q_n \sim cr^{n/2}$  or a random variable  $N$  which has a geometric distribution, at least in the tails. Suppose the sequence  $X_n$  is deterministic and we use  $Q_n \sim c\sqrt{|\nabla x_n|}$ . Then the variance is finite provided the series  $\sum_n \sqrt{|\nabla x_n|}$  is convergent.

Let us consider a simple example before we look at more complex ones. Suppose  $X_n = b + ar^n$  for  $n = 1, \dots, |r| < 1$  and  $x_{\infty} = b$ . Then  $\nabla X_n = ar^{n-1}(r-1)$ ,  $n \geq 2$  and  $\nabla X_n^2 = 2abr^{n-1}(r-1) + a^2r^{2n-2}(r^2-1)$  and, as we already verified more generally,  $E(X) = E\left[X_0 + \sum_{n=1}^{\infty} \nabla X_n \frac{I(n \leq N)}{Q_n}\right] = b$  whatever the distribution of  $N$ . Suppose we use a shifted geometric distribution

for  $N$  so that  $P(N \geq n) = q^{n-s}$ ,  $n = s, s+1, \dots$  for  $|q| < 1$ . Evidently to minimize the variance we should choose

$$Q_n \sim c|r|^n, n = 1, \dots \quad (5)$$

so that  $q = |r|$ . The variance for general  $q$  is

$$\begin{aligned} \sum_{n=s+1}^{\infty} \frac{2(x_{\infty} - \xi_n) \nabla X_n}{Q_n} - (x_{\infty} - X_s)^2 &= \sum_{n=s+1}^{\infty} \frac{-2a \left( \frac{r^n + r^{n-1}}{2} \right) ar^{n-1} (r-1)}{q^{n-s}} - a^2 r^{2s} \\ &= a^2 r^{2s} \left[ \frac{1-q}{q-r^2} \right] \quad \text{where } 1 > q > r^2. \end{aligned}$$

Suppose we wish to minimize this over the values of  $s$  and  $q$  subject to the constraint that  $E(N) = \frac{q}{1-q} + s = \mu_N$  is constant (ignoring the integer constraint on  $s$ ). Then with  $z = \frac{1}{q}$ ,

$$\begin{aligned} \min_{s,z} r^{2s} \left[ \frac{z-1}{1-r^2 z} \right] \quad \text{subject to } \frac{1}{z-1} + s &= \mu_N \text{ or} \\ \min_z r^{-2/(z-1)} \left[ \frac{z-1}{1-r^2 z} \right], \quad \text{for } \frac{1}{r^2} > z > 1 \end{aligned}$$

which minimum occurs when  $z = \frac{1}{r}$ , or  $q = |r|$  and  $s \simeq \mu_N - \frac{|r|}{1-|r|}$  and then the minimum variance is

$$a^2 |r|^{2s-1} \simeq a^2 |r|^{2\mu_N - \frac{2|r|}{1-|r|}}$$

Notice that the mean squared error, if we were to stop after  $\mu_N$  iterations, is  $a^2 |r|^{2\mu_N}$  so we have purchased unbiasedness of the estimator at a cost of increasing the MSE by a factor of approximately  $|r|^{-\frac{2|r|}{1-|r|}}$ . This factor is plotted in Figure 1. It can be interpreted as follows: in the worst case scenario when  $p$  is around .4, we will need about 3 times the sample size for the debiased estimator to achieve the same mean squared error as a conventional iteration using deterministic  $N$ . However when  $|r|$  is close to 1 indicating a slower rate of convergence, there is very little relative increase in the MSE.

Note: the optimisation problem above tacitly assumed that the computation time required to generate the sequence is  $O(n)$ . This is not the case with some applications; for example in the numerical integral below the computation time is  $O(2^n)$  since there are  $2^n$  intervals and  $2^n + 1$  function evaluations and in this case a more appropriate minimization problem, having budget constraint  $E(2^N) = 2^s \frac{1-q}{1-2q} = c \geq 2^s$ , is, with  $0 < q < \frac{1}{2}$ ,

$$\min_{s,q} r^{2s} \left[ \frac{1-q}{q-r^2} \right] \quad \text{subject to } \frac{1}{2} > q > r^2 \text{ and } 2^s \frac{1-q}{1-2q} = c > 2^s, s = 0, 1, 2, \dots$$

or, putting  $q = \frac{c-2^s}{2c-2^s}$ ,

$$\min_{s \leq \log_2(c)} r^{2s} \left[ \frac{c}{c-2^s - r^2(2c-2^s)} \right]$$

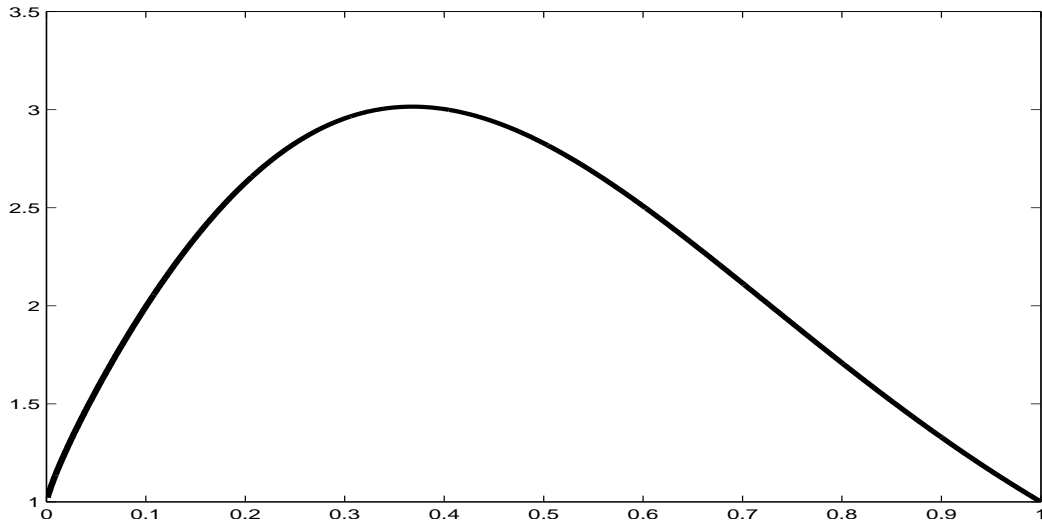


Figure 1: Relative increase in MSE due to debiasing the sequence.

which minimum appears to occur for  $s = \lceil \log_2(c) \rceil$  and  $q = \frac{c-2^s}{2c-2^s}$  when  $r < 0.53$  and otherwise  $s$  may be somewhat smaller. Intuitively, when the rate of convergence is reasonably fast (so  $r$  is small) then the minimum variance is achieved by a large guarantee on the value of  $N$  ( $s$  large) and then the residual budget ( $c - 2^s$ ) used to produce unbiasedness by appropriate choice of  $q$ .

### 3 Examples

#### Example 1 *Unbiased Estimation of a root*

Suppose we wish to find the root of a nonlinear function  $h$ . For a toy example, suppose we wish to solve for  $x$  the equation

$$h(x) - \alpha = 0.$$

We might wish to use (modified) Newton's method with a random starting value to solve this problem, requiring randomly generating the initial value  $X_0$  and then iterating

$$X_{n+1} = X_n - \delta_n \quad \text{where } \delta_n = \min\left(1, \max\left(-1, \frac{h(X_n) - \alpha}{h'(X_n)}\right)\right)$$

but of course after a finite number of steps, the current estimate  $X_n$  is likely a biased estimator of the true value of the root. We implemented the debiasing procedure above with  $h(x) = x^3$  and  $\alpha = 1$ . We generated  $X_0$  from a  $U(-2, 3)$  distribution, chose  $P(N = n) = (1 - p)p^{n-s}$ ,  $n = s, s + 1, \dots$  and for simplicity

used  $s = 4, p = \frac{3}{4}$  and repeated for 100,000 simulations. The sample mean of the estimates was 1.0023 and the sample variance 0.011. Although the procedure works well in this case when we start sufficiently close to the root, it should be noted that this example argues for an adaptive choice of  $Q_n$ , one which permits a larger number of iterations (larger values of  $N$ ) when the sequence seems to indicate that we have not yet converged. This is discussed below.

### Stopping times for $N$ .

In view of the last example, particularly if  $X_0$  is far from  $x_\infty$ , it would appear desirable to allow  $N$  to be a stopping time. In order to retain unbiasedness, it is sufficient that

$$EY = E \left[ x_0 + \sum_{n=1}^{\infty} \nabla X_n \frac{I(n \leq N)}{Q_n} \right] = E \left[ x_0 + \sum_{n=1}^{\infty} \nabla X_n \right]$$

or

$$E \left[ \frac{I(n \leq N)}{Q_n} | X_1, X_2, \dots, X_n \right] = 1.$$

Therefore it is sufficient that  $Q_n = P(N \geq n | X_1, X_2, \dots, X_n)$  and one simple rule for an adaptive construction of  $N$  is:

$$P(N \geq n | X_1, X_2, \dots, X_n) = \begin{cases} P(N \geq n-1 | X_1, X_2, \dots, X_{n-1}) & \text{if } \nabla X_n > \varepsilon \\ pP(N \geq n-1 | X_1, X_2, \dots, X_{n-1}) & \text{if } \nabla X_n \leq \varepsilon \end{cases}$$

There are, of course, many potential more powerful rules for determining the shift in the distribution of  $N$  but we concentrate here on establishing the properties of the simplest version of this procedure.

### Example 2 *Simpson's rule.*

Consider using a trapezoidal rule for estimating the integral

$$I_\infty = \int_0^1 f(x) dx$$

using  $2^n + 1$  function evaluations which evaluate the function on the grid  $0, \Delta x, 2x, \dots, 2^n \Delta x = 1$ . Denote the estimate of the integral  $I_n$ . Here  $\Delta x = 2^{-n}$  and the error in Simpson's rule assuming that the function has bounded fourth derivative is  $O((\Delta x)^4) = O((2^{-n})^4) = O(16^{-n})$ . This suggests a random  $N$  such that  $Q_n \sim \frac{1}{4^n}$  or a (possibly shifted) geometric distribution with  $p = \frac{3}{4}$ . Suppose  $Q_n = 4^{-n+2}, n = 2, 3, \dots$ . This means  $E(N) = \frac{7}{3}$  which is quite small. In general, the estimator has finite variance since

$$\sum_{n=1}^{\infty} \frac{\nabla I_n}{Q_n} < \infty.$$

More generally, if  $N$  has a shifted geometric distribution with probability function  $P(N = n) = p(1-p)^{n-s}$ ,  $n = s, s+1, \dots$  parameter  $p$ , the expected number of function evaluations in the quadrature rule is

$$\begin{aligned} \sum_{n=1}^{\infty} (2^n + 1)P(N = n) &= \sum_{n=s}^{\infty} (2^n + 1)(1-p)^{n-s}p \\ &= 1 + 2^s \frac{p}{2p-1} \end{aligned}$$

and this is 7, for example, when  $p = \frac{3}{4}$ ,  $s = 2$ . How well does this perform? This provides an unbiased estimator of the integral with variance

$$\sigma_X^2 = \sum_{n=2}^{\infty} \frac{\nabla I_n (1 - Q_n)}{Q_n} [2I_{\infty} - I_n - I_{n-1}]$$

which can be evaluated or estimated in particular examples and compared with the variance of the corresponding crude Monte Carlo estimator. For a reasonable comparison, the latter should have the same (expected) number of function evaluations, i.e. 7 and therefore has variance

$$\frac{1}{7} \left( \int_0^1 f^2(x) dx - I_{\infty}^2 \right)$$

Take, for example, the function  $f(x) = \sin(\pi x)$  so that  $I_{\infty} = \int_0^1 \sin(\pi x) dx = \frac{2}{\pi} = 0.63662$  and  $\int_0^1 (\sin(\pi x))^2 dx = \frac{1}{2}$ . In this case the variance of the MC estimator with seven function evaluations is  $\frac{1}{7} \left( 0.5 - \left(\frac{2}{\pi}\right)^2 \right) \simeq 0.013531$ . We compare this with the estimator obtained by randomizing the number of points in a Simpson's rule. Here it is easy to check that

$n$	1	2	3	4	5	6
$I_n$	0.3047	0.2149	0.2124	0.2122	0.2122	0.2122
$\nabla I_n$	-0.089821	-0.002562	-0.000139	-0.000008	-0.000001	-0.00000003

Table 1: Values of the numerical integral  $I_n$  and  $\nabla I_n$  with  $2^n$  intervals

and in this case the variance of the debiased Simpson's rule estimate is  $\sigma_X^2 \simeq 6.41 \times 10^{-6}$  indicating more than a two thousand-fold gain in efficiency over crude Monte Carlo.

**Note:** We have chosen the grid size  $2^{-n}$  in view of the fact that when  $N = n$ , we need the integrals  $I_j$  for all  $j \leq n$ . In this case, we can simply augment the function evaluations we used for  $I_j$  in order to obtain  $I_{j+1}$ .

The major advantage of this debiasing procedure however is not as a replacement for Crude Monte Carlo in cases where unbiased estimators exist, but as a device for creating unbiased estimators when their construction is not at all obvious. This is the case whenever the function of interest is a nonlinear function of variables that can be easily and unbiasedly estimated as in the following example.



**Example 3** *Heston Stochastic Volatility model*

In the Heston stochastic volatility model, under the risk neutral measure  $Q$ , the price of an asset  $S_t$  and the volatility process  $V_t$  are governed by the pair of stochastic differential equations

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t \rho dW_1(t) + \sqrt{1 - \rho^2} dW_2(t), & S_0 &= s_0 \\ dV_t &= \kappa(\theta - V_t) dt + \sigma_V \sqrt{V_t} dW_1(t), & V_0 &= v_0 \end{aligned}$$

where  $W_1(t)$  and  $W_2(t)$  are independent Brownian motion processes,  $r$  is the interest rate,  $\rho$  is the correlation between the Brownian motions driving the asset price and the volatility process,  $\theta$  is the long-run level of volatility and  $\kappa$  is a parameter governing the reversion. Denote by  $BS(S_0, K, r, T, \sigma)$  the Black-Scholes price of a call option having initial stock value  $S_0$ , volatility  $\sigma$ , interest rate  $r$ , expiration time  $T$ , option strike price  $K$  and 0 dividend yield. The price of a call option in the Heston model can be written as an expected value under the risk-neutral measure  $Q$  of a function of two variables  $g(V_T, I(T))$  (see for example Willard (1997) and Broadie and Kaya (2006))

$$\begin{aligned} e^{-rT} E^Q(S_T - K)^+ &= E[BS(S_0 \xi, K, r, T, \tilde{\sigma} \sqrt{1 - \rho^2})] \text{ where} \\ \xi &= \xi(V_T, I(T)) = \exp\left(-\frac{\rho^2}{2} I(T) + \rho \int_0^T \sqrt{V_s} dW_1(s)\right) \\ &= \exp\left(-\frac{\rho^2}{2} I(T) + \frac{\rho}{\sigma} (V_T - V_0 + \kappa I(T) - \kappa \theta T)\right) \text{ and} \\ \tilde{\sigma} &= \tilde{\sigma}(I(T)) = \sqrt{\frac{I(T)}{T}} \text{ where } I(T) = \int_0^T V_s ds. \end{aligned}$$

This can be valued conditionally on  $V_T, I(T)$  with the usual Black-Scholes formula. In particular with  $g(V_T, I(T)) = BS(S_0 \xi, K, r, T, \tilde{\sigma} \sqrt{1 - \rho^2})$ , the option price is  $E^Q g(V_T, I(T))$ .

Note that  $g$  is clearly a highly nonlinear function of  $V_T$  and  $I(T)$  and so, even if exact simulations of the latter were available, it is not clear how to obtain an unbiased simulation of  $g$ . In the Heston model, and indeed various other stochastic volatility models, it is possible to obtain an exact simulation of the value of the process  $V_t$  at finitely many values of  $t$  so it is possible to approximate the integral  $I(T)$  using  $I_n(T)$  obtained from a trapezoidal rule with  $1 + 2^n$  points. This raises the question of what we should choose as a distribution for  $N$ . Under conditions on the continuity of the functional of the process whose expected value is sought, Kloeden and Platen (1995, Theorem 14.1.5, page 460) show that the Euler approximation to the process with interval size  $2^{-n}$  results in an error in the expected value of order  $2^{-n\chi}$  where  $\chi = 1$  for sufficiently smooth (four times continuously differentiable) drift and diffusion coefficients so for simplicity consider this case. This implies that

$$|Eg(V_T, I_n(T)) - Eg(V_T, I(T))| < \text{constant} \times 2^{-n}.$$

which suggests we choose  $Q_n \sim 2^{-n}$ .

As before we randomly generate  $N$  from a (possibly shifted) geometric( $p$ ) distribution with  $p = \frac{1}{2}$ . The function to be integrated  $V_s$  is not twice differentiable so we need to determine empirically the amount of the shift (and we experimented with reasonable values of  $p$ ). We chose parameters  $p = 0.5$  and shifted the geometric random variable by  $s = 2$  so that  $P(N = n) = p(1-p)^{n-s}$  for  $n = s, s + 1, \dots$ . The parameters used in our simulation were taken from Broadie and Kaya(2004):  $s_0 = 100; K = 100; V_0 = 0.09; \kappa = 2.00; \theta = 0.09; r = .05; \sigma = 1; \rho = -.3; T = 5$ ; for which, according to Broadie and Kaya, the true option price is around 34.9998. 1,000,000 simulations with  $p = 0.75$  and  $s = 4$  provided an estimate of this option price of 34.9846 with a standard error of 0.0107 so there is no evidence of bias in these simulations. With parameter values  $\theta = 0.019; \kappa = 6.21; \sigma = 0.61; v_0 = 0.010201; r = 0.0319; s_0 = 100; K = 100; T = 1; \rho = -0.7$  and  $p = 0.75$  and  $s = 4$ , we conducted  $10^6$  simulations leading to an estimate of 6.8115 with a standard error of 0.0048998. This is in agreement with the Broadie and Kaya "true option price" of 6.801. Note that the Feller condition for positivity requires  $2\kappa\theta > \sigma^2$  which fails in the above cases. This means that the volatility process hits zero with probability one, and for some parameter values, it does so frequently which may call into question the value of this model with these parameter values. 100,000 simulations from these models used about 10-13 minutes running Matlab 5.0 on an intel Core 2 Quad CPU @2.5 GHz.

## 4 Conclusion

When numerical methods such as quadrature or numerical solutions to equations may result in a biased estimator, a procedure is suggested which eliminates this bias and provides statistical estimates of error. This procedure is successfully implemented both in simple root finding problems and in more complicated problems in finance and has enormous potential for providing Monte Carlo extensions of numerical procedures which allow unbiased estimates and error estimates.

## References

- [1] Willard, G. A. Calculating prices and sensitivities for path independent derivative securities in multifactor models. *J.Derivatives* 5(1) 45–61 (1997)
- [2] Broadie, M. and Kaya, O. Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes *Operations Research* 54 (2), 217–231 (2006)
- [3] Glasserman, P.: *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York (2003)

- [4] Kloeden, P.E. and Platen, E.: *Numerical Solution of Stochastic Differential Equations (Stochastic Modelling and Applied Probability)*, Springer, New York (1995)