Peter Jäckel[*]

Weighted sampling for variance reduction

First version: 1st October 2004
This version: 12th July 2008

Assume we wish to compute the expectation of a binary option by means of a Monte Carlo simulation with n samples:

$$\mathsf{E}\left[\mathbf{1}_{\{z>K\}}\right] \approx x_n \tag{1}$$

with

$$x_n := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{z_i>K\}}\,. \tag{2}$$

Assume further that we want z to have the law of a standard normal distribution, and that we wish to enhance the convergence of our Monte Carlo simulation by sampling from a normal distribution that is shifted by μ, adjusting for the likelihood ratio:

$$\tilde{x}_n := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{z_i+\mu>K\}} \frac{\varphi(z_i+\mu)}{\varphi(z_i)} \qquad \text{with} \qquad z_i \sim \mathcal{N}(\mu,1)\ . \tag{3}$$

The sum $\tilde{x}_n$ is a random number with expectation

$$\mathsf{E}[\tilde{x}_n] \;=\; \Phi(-K)\ . \tag{4}$$

Its second moment is

$$\begin{aligned}
\mathsf{E}\left[\tilde{x}_n^2\right] \;&=\; \frac{1}{n^2}\mathsf{E}\left[\left(\sum_{i=1}^{n} \mathbf{1}_{\{z_i+\mu>K\}} \frac{\varphi(z_i+\mu)}{\varphi(z_i)}\right)^2\right] &(5)\\
&=\; \frac{n}{n^2}\int\limits_{z=K-\mu}^{\infty} \left(\frac{\varphi(z+\mu)}{\varphi(z)}\right)^2 \varphi(z)\,\mathrm{d}z + \frac{n^2-n}{n^2}\left(\int\limits_{z=K-\mu}^{\infty}\left(\frac{\varphi(z+\mu)}{\varphi(z)}\right)\varphi(z)\,\mathrm{d}z\right)^2 &(6)\\
&=\; \frac{1}{n}\int\limits_{z=K-\mu}^{\infty} \frac{\mathrm{e}^{-\frac{1}{2}(z+\mu)^2\cdot 2+\frac{1}{2}z^2}}{\sqrt{2\pi}}\,\mathrm{d}z + \left(1-\frac{1}{n}\right)\Phi^2(-K) &(7)\\
&=\; \frac{1}{n}\mathrm{e}^{\mu^2}\cdot\int\limits_{z=K-\mu}^{\infty} \frac{\mathrm{e}^{-\frac{1}{2}(z+2\mu)^2}}{\sqrt{2\pi}}\,\mathrm{d}z + \left(1-\frac{1}{n}\right)\Phi^2(-K) &(8)\\
&=\; \frac{1}{n}\mathrm{e}^{\mu^2}\Phi(-K-\mu) + \left(1-\frac{1}{n}\right)\Phi^2(-K) &(9)
\end{aligned}$$

The variance of $\tilde{x}_n$, i.e.

$$\mathsf{V}[\tilde{x}_n] \;=\; \frac{1}{n}\left(\mathrm{e}^{\mu^2}\Phi(-K-\mu) - \Phi^2(-K)\right) \tag{10}$$

is therefore a function of both μ and K, and we show this in figure 1. In order to minimise the

[*]OTC Analytics

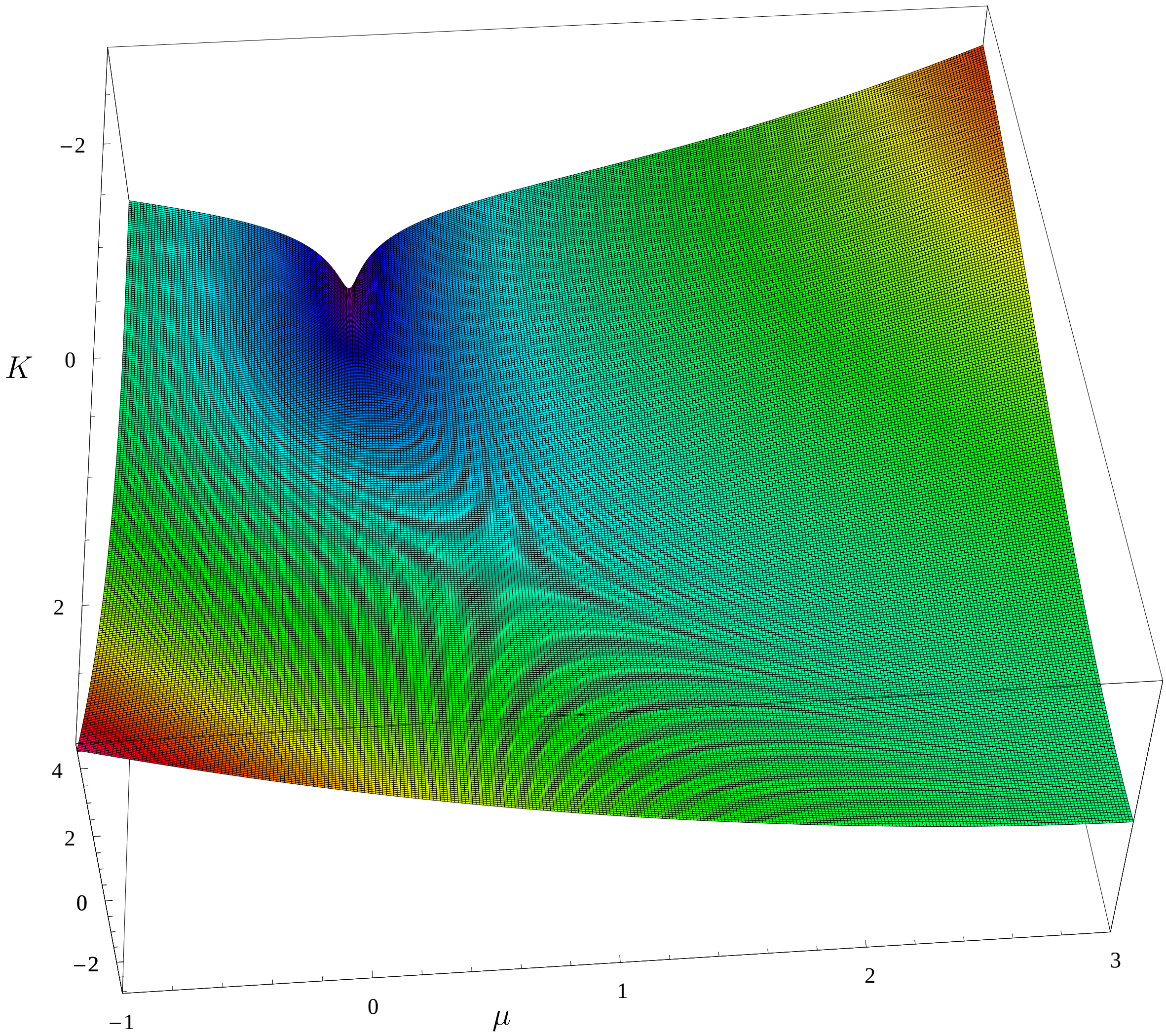

Figure 1: The decadic logarithm of normalised variance, i.e. $\log_{10}\left(n \cdot \mathsf{V}[\tilde{x}_n] / \mathsf{E}[\tilde{x}_n]^2\right)$.

variance of the simulation result as a function of the shift parameter μ, we set

$$\partial_\mu \mathsf{V}[\tilde{x}_n] \;=\; 0 \tag{11}$$

which gives us

$$2\mu\Phi(-K-\mu) = \varphi(K+\mu) \; . \tag{12}$$

The value $\mu = \mu^*(K)$ that solves (12) minimises the variance of $\tilde{x}_n$.

By direct differentiation of equation (12), we can derive that the slope of the optimal curve $\mu^*(K)$ can be expressed as the comparatively simple function

$$\mu^{*\prime}(K) \;=\; \frac{\mu^*(K)(\mu^*(K)-K)}{1-\mu^*(K)(\mu^*(K)-K)} \; . \tag{13}$$

This result can be used for the construction of an approximation. Choosing α and β such that the hyperbolic approximation

$$\beta K + \sqrt{(1-\beta)^2K^2+\alpha^2} \tag{14}$$

matches $\mu^*(K)$ both in value and in slope at $K = 0$ leads to

$$\alpha = \mu^*(0) = 0.6120031809624\ldots \tag{15}$$
$$\beta = \frac{\alpha^2}{1-\alpha^2} = 0.5988434439993\ldots \tag{16}$$

and this approximation is fairly accurate for $K \geq 0$. For $K < 0$, however, the hyperbolic form is not the most suitable. Instead, the exponential form

$$\alpha e^{\beta/\alpha K + \gamma K^2} \tag{17}$$

with

$$\gamma = \frac{1-2\beta}{2\alpha^2} = -0.2639006805605\ldots \tag{18}$$

is more appropriate. In total, a reasonable approximation for $\mu^*(K)$ is given by:

$$\mu^*_{\text{approximation}}(K) = \begin{cases} \alpha e^{\beta/\alpha K + \gamma K^2} & \text{for} \quad K < 0 \\ \beta K + \sqrt{(1-\beta)^2 K^2 + \alpha^2} & \text{for} \quad K \geq 0 \end{cases} \tag{19}$$

The curves $\mu^*(K)$ and $\mu^*_{\text{approximation}}(K)$ are shown in figure 2. The variance reduction factor

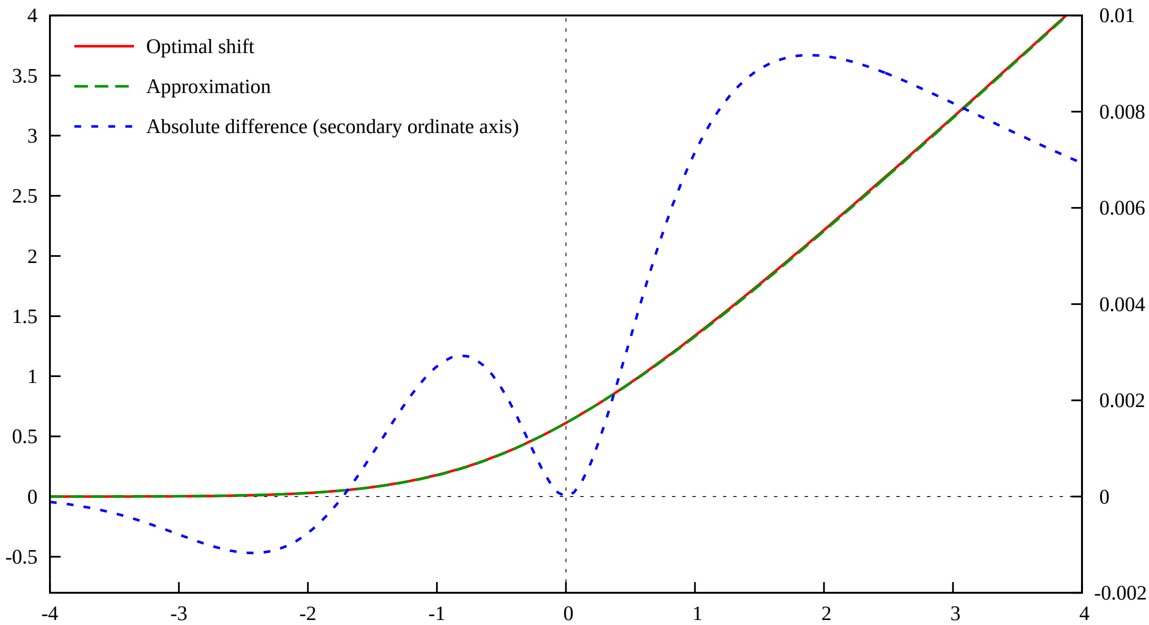

Figure 2: The variance minimising shift $\mu^*(K)$ (red solid line) and its hyperbolic approximation $\mu^*_{\text{approximation}}(K)$ (green dashed line) and their difference (blue short dashed line).

achieved by the use of the shift $\mu^*(K)$ and its approximation $\mu^*_{\text{approximation}}(K)$ is shown in figure 3. Note that for $K \approx 3^1\!/_2$, the reduction in variance already exceeds a factor of 10^3.

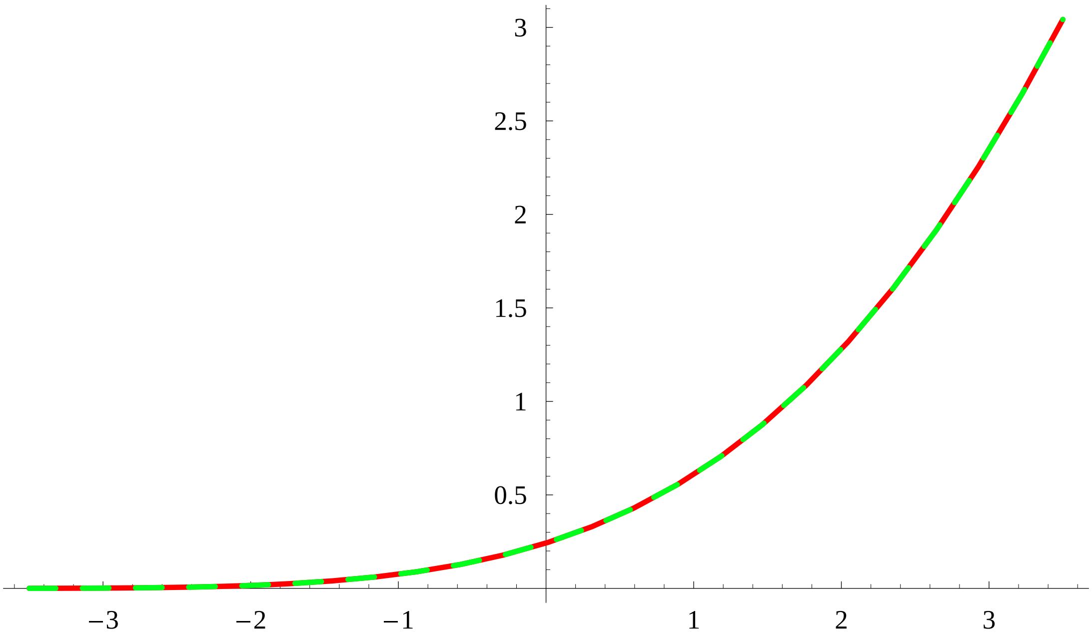

Figure 3: The decadic logarithm of the reduction in variance resulting from the use of the Gaussian shift $\mu^*(K)$ (red solid line) and its approximation $\mu^*_{\text{approximation}}(K)$ (green dashed line) as a function of the strike K.