Variance reduction by importance sampling

a

Suppose we want to compute $\alpha \equiv \mathrm{E}\left(g\left(W\right)\right)$ with g non-negative and $W \sim f_W$. Let X be a random variable $X \sim f_X$, then

$$\alpha = \mathrm{E}\left(g\left(X\right)\frac{f_W\left(X\right)}{f_X\left(X\right)}\right). \tag{1}$$

Proof:

$$\begin{aligned}
\alpha =& \int_{\mathbb{R}} g\left(x\right) f_W\left(x\right) dx && (2)\\
=& \int_{\mathbb{R}} g\left(x\right)\frac{f_X\left(x\right)}{f_X\left(x\right)} f_W\left(x\right) dx && (3)\\
=& \int_{\mathbb{R}} g\left(x\right)\frac{f_W\left(x\right)}{f_X\left(x\right)} f_X\left(x\right) dx && (4)\\
=& \mathrm{E}\left(g\left(X\right)\frac{f_W\left(x\right)}{f_X\left(x\right)}\right). && (5)
\end{aligned}$$

b,c

The above result allows solving α by Monte Carlo (MC) methods. When we sample from distribution f_X to estimate α, the mean square error (MSE) of the MC estimator

$$\mathrm{MSE} \propto \int_{\mathbb{R}} \left(\left(g\left(x\right)\frac{f_W}{f_X}\left(x\right)\right)^2 - \alpha^2\right) f_X\left(x\right) dx. \tag{6}$$

The integrand is zero by choosing $f_X\left(x\right) = f_X^*\left(x\right) = \frac{f_W(x)g(x)}{\alpha}$, and thus the variance of the MC estimator is minimised by choosing $f_X = f_X^*$. In practice, this result is of little significance, since if we knew the exact value of α, there would be no need for the MC estimator in the first place.

d

Suppose we wish to evaluate $\mathrm{P}\left(\mathcal{N}\left(0,1\right) > 3.75\right)$. The naive MC estimator would be given by $g\left(x\right) = \mathbf{1}_{x>\frac{15}{4}}\left(x\right)$ and $f_W\left(x\right) = \sqrt{2\pi}\exp\left(-\frac{x^2}{2}\right)$. Alternatively, using an affine change of variables, we may choose for any $\delta \in \mathbf{R}$: $g_\delta\left(x\right) = \mathbf{1}_{x>\frac{15}{4}+\delta}$ and $f_W = \sqrt{2\pi}\exp\left(-\frac{(x-\delta)^2}{2}\right)$ and obtain the correct result.

In order to use the MC estimate, we need to be able to sample numbers, from the distribution $\mathcal{N}\left(\delta,1\right)$. And evaluate g for each of those realisations. An example implementation of this is given in `https://github.com/`

`Virtakuono/SME_HW3_Example/blob/master/examples.py` The MC estimator is given by:

$$\overline{\alpha} = \sum_{m=1}^{M} \frac{g_\delta(X)}{M}, \tag{7}$$

with $X_m \sim N(\delta, 1)$ i.i.d. In the example we set $\delta = 2$. Since $\mathrm{E}(X_m) = \delta$, $\mathrm{E}(\overline{\alpha}) = \alpha$. Central limit theorem guarantees that $\overline{\alpha}$ is approximately normally distributed. To estimate the confidence interval, we compute the sample variance as

$$\overline{\sigma}^2 = \sum_{m=1}^{M} \frac{(g_\delta(x) - \overline{\alpha})}{M-1}. \tag{8}$$

Let

$$\Phi_{\mu,\sigma}(z) = \left(2\pi\sigma^2\right)^{-1} \int_{-\infty}^{z} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{9}$$

Then, there are ready implementations for the inverse of $\Phi_{\mu,\sigma}$, that allow us to compute z^* such that $\Phi_{0,1}(z^*) = 0.95$. Using this together with the fact that our MC estimator is approximately normally distributed, we set $\eta = z^* \sqrt{\frac{\overline{\sigma}^2}{M}}$ and obtain the confidence interval $[\overline{\alpha} - \eta, \overline{\alpha} + \eta]$.

Generate non-uniformly distributed random numbers given uniform

Kernel Density Estimator (KDE)

a

We propose to estimate a density function $\rho(y)$ based on a MC sample of realisations $y_i \in \mathbb{R}^d$ from the density ρ as follows

$$\hat{\rho}(y) = \frac{1}{Mh^{-d}} \sum_{m=1}^{M} K\left(\frac{y - y_i}{h}\right), \tag{10}$$

with

$$K(x) = 2^{-d} \mathbf{1}_{||x||_{L_2}}(x) \tag{11}$$

In this estimation, we commit two errors.

bias error

$$
\begin{aligned}
&\rho\left(u\right)-\mathrm{E}\left(\hat{\rho}\left(y\right)\right) && (12)\\
=&\rho\left(y\right)\int_{\mathbb{R}^d} h^{-d}K\left(\frac{y-z}{h}\right)\rho\left(z\right)dz && (13)\\
=&\int_{\mathbb{R}^d}\left(\rho\left(y\right)-\rho\left(y-hz\right)\right)K\left(z\right)dz && (14)\\
\approx&\int_{\mathbb{R}^d}\left(\rho\left(y\right)-\left(\rho\left(y\right)-hz_i\partial_{y_i}\rho\left(y\right)+h^2z_iz_j\partial^2_{z_i,z_j}\rho\left(y\right)\right)\right)K\left(z\right)dz && (15)\\
=&\frac{h^2}{2}\partial_{y_iy_j}\rho\left(u\right)\int_{\mathbb{R}^d} z_iz_jK\left(z\right)dz\propto h^2 && (16)
\end{aligned}
$$

statistical error

The variance is bounded by

$$
\mathrm{V}\left(K\left(\frac{y-z}{h}\right)\right)\leq\mathrm{E}\left(K^2\left(\frac{y-z}{h}\right)\right) \tag{17}
$$

and

$$
\mathrm{E}\left(K^2\left(\frac{y-z}{h}\right)\right)\approx\int_{\mathbb{R}^d}K^2\left(z\right)h^d\left(\rho\left(y\right)-hz_i\partial_{y_i}\rho\left(y\right)+\frac{h^2}{2}z_iz_j\partial^2_{y_iy_j}\rho\left(y\right)\right)dz, \tag{18}
$$

thus

$$
\mathrm{V}\left(\hat{\rho}\left(y\right)\right)=M^{-1}h^{-2d}\mathrm{V}\left(K\left(\frac{y-z}{h}\right)\right) \tag{19}
$$

$$
\geq M^{-1}h^{-2d}h^d\rho\left(y\right)\int_{\mathbb{R}^d}K^2\left(z\right)dz\propto M^{-1}h^d \tag{20}
$$

Total error

With undefined constants A and B we have that the total error squared (TSE) is given as

$$
\mathrm{TSE}\approx Ah^4+\frac{B}{Mh^d}. \tag{21}
$$

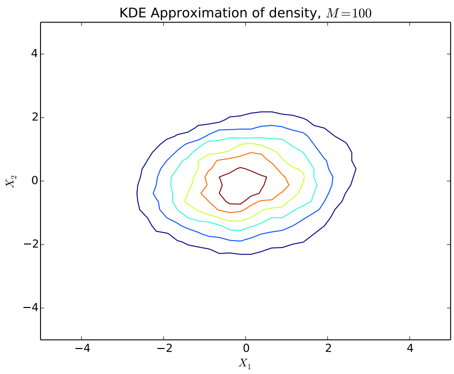
The critical point of the total square error is thus given by

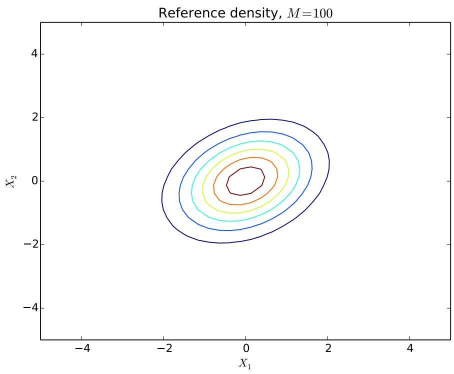
$$
h^{4+d}\propto M^{-1}. \tag{22}
$$

This choice implies that the bias and sampling errors are of the same order

$$
\mathrm{TSE}=\mathcal{O}\left(M^{-\frac{2}{d+4}}\right) \tag{23}
$$

With the above analysis, we produce the set of three images for varying M for a one-dimensional KDE and an example of two-dimensional gaussian example.


KDE Approximation of density, $M = 100$
4
2
0
−2
−4
X_2
−4
−2
0
2
4
X_1


Reference density, $M = 100$
4
2
0
−2
−4
X_2
−4
−2
0
2
4
X_1

KDE for conditional expectation

Monte Carlo and Variance Reduction

The task at hand is to evaluate an option on an underlying that is an arithmetic mean of a log-normal random variable. No closed form solution exists. However, we do have a closed form expression for the arithmetic mean of normal variables. We have

$$V_N = V_0 \prod_n^N R_n. \tag{24}$$

Empirically, we see that the probability P_{tom} of V_N being out of the money is approximately 72 per cent. Define

$$Q = \sum_{n=1}^{N} (50 - n) \left(\ln\, R_n - r\Delta t \right), \tag{25}$$

we note that both Q and V_N are increasing in all the random variables R_n. Furthermore, we know that Q entered and normally distributed with $\tilde{\sigma}^2 = \sum_{n=1}^{N} (50 - n)^2$. With the inverse cumulative distribution Φ^{-1} we can define

$$\tilde{K} = \tilde{\sigma} \Phi^{-1} \left(P_{otm} \right). \tag{26}$$

Then our control variate is

$$G = \left(Q - \tilde{K} \right)^+. \tag{27}$$

We have

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\tilde{\sigma}^2} \right) (x - k)^+ = \frac{\tilde{\sigma}}{\sqrt{2\pi}} \exp\left(-\frac{k^2}{2\tilde{\sigma}^2} \right) + k \left(\Phi\left(k\right) - 1 \right). \tag{28}$$

We have that the correlation coefficient ρ between G and $C = \left(\sum_{n=1}^{N} \frac{V_n}{K} - K, 0 \right)^+$ exceeds 0.99, see fig 1.

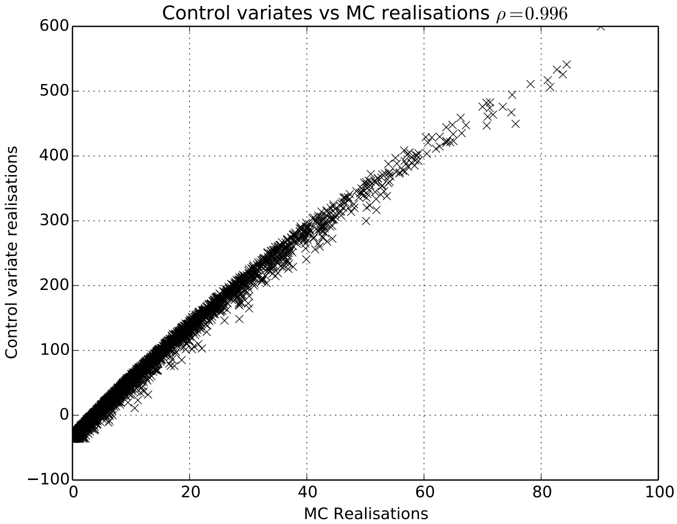


Figure 1: The MC realisations and corresponding control variates show a near-perfect correlation

We can form the estimators

$$\overline{C}_1 = \sum_{m=1}^{M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} R_n\right) - K\right)^+}{M} \tag{29}$$

$$\overline{C}_2 = \sum_{m=1}^{2M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} \tilde{R}_{n,m}\right) - K\right)^+}{M} + \sum_{m=1}^{2M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} \tilde{R}_{n,m}\right) - K\right)^+}{M} \tag{30}$$

$$\overline{C}_3 = \sum_{m=1}^{M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} R_n\right) - K\right)^+ + \beta\left(G_m - \mathrm{E}\left(Q\right)\right)}{M} \tag{31}$$

$$\overline{C}_4 = \sum_{m=1}^{2M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} R_n\right) - K\right)^+ + \beta\left(G_m - \mathrm{E}\left(Q\right)\right)}{M}$$

$$\sum_{m=1}^{2M} \frac{\left(\left(\sum_{n=1}^{N} V_0 \prod_{m=1}^{n} R_n\right) - K\right)^+ + \beta\left(Q_m - \mathrm{E}\left(Q\right)\right)}{M}, \tag{32}$$

Plain vanilla	Antithetic	Control variate	Hybrid
0.15 %	0.10 %	0.012 %	0.0085 %

Table 1: Widths of 95 % confidence bands for different Monte Carlo estimators for $M = 1000$. Modified option price approximately 4.3.

for the plain vanilla, antithetic, control variate, and combination methods, respectively. The $R_{n,m}$ are understood to be independent realisations of the incrementation for nth time step for realisation m The antithetic variables are defined as

$$\tilde{R}_{n,m} = \exp\left(-\log R_{n,m} + \Delta tr\right) \tag{33}$$

and

$$\tilde{Q}_m = \sum_{n=1}^{N} (50 - n) \left(\ln \tilde{R}_{nm} - r\Delta t\right), \tag{34}$$

$$G_m = \left(Q_m - \tilde{K}\right)^+ \tag{35}$$

$$\tilde{G}_m = \left(\tilde{Q}_m - \tilde{K}\right)^+. \tag{36}$$

In order to minimise the variance of the estimators, we set $\beta = \rho\sqrt{\frac{\sigma_C^2}{\sigma_Q^2}}$ with σ_C^2 being the variance of the plain vanilla MC realisations and σ_Q^2 the variance of the control variate realisations. With $M = 10000$, we obtain the results in table 1: The exact computational cost of each estimator depends on how computationally costly it is to generate time steps compared to the cost of generating random numbers. For a rough approximation, we may state that the computational cost of the antithetic and control variate estimators are twice that of the plain vanilla and the hybrid method requires four-fold computational effort.

Confidence interval, Bootstrapping

In order to compute the expected shortfall we reuse the MC sample from the previous exercise. Let our MC estimator $\overline{C}$ be given as:

$$\overline{C} = \sum_{m=1}^{M} \frac{X_i}{M}. \tag{37}$$

Then let us draw $N \times M,\ N, M \in \mathbb{Z}$ random variables $k_i \in \{1, 2, 3, ...M\}$. Then define S_j such that

$$\#A_j = \#\left\{k_i : X_{k_i} > S_j, (j-1)\,M \leq i \leq jM - 1\right\} = \lfloor pM \rfloor \tag{38}$$

p	0.9	0.95	0.99
Confidence interval	$[28.1, 30.2]$	$[36.8, 39.7]$	$[56.1, 62.1]$

Table 2: Confidence intervals for the Expected Shortfall for problem 5 example for various quantiles.

for $j \in 1, 2, 3, 4, ..., N$. and

$$Q_j = \sum_{i \in A_j} \frac{X_i \mathbf{1}_{x > S_j}(X_i)}{\# A_j} \tag{39}$$

then the estimated confidence interval can be noted as $[Q^-, Q^+]$ as

$$\#\left\{j : Q_j < Q^-\right\} = \left\lfloor \frac{qN}{2} \right\rfloor \tag{40}$$

$$\#\left\{j : Q_j < Q^+\right\} = \left\lfloor \frac{qN}{2} \right\rfloor. \tag{41}$$

Setting $N = 1000$, we obtain the results in the table 2