

## Variance reduction by importance sampling

**a**

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

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

Proof:

$$\alpha = \int_{\mathbb{R}} g(x) f_W(x) dx \quad (2)$$

$$= \int_{\mathbb{R}} g(x) \frac{f_X(x)}{f_X(x)} f_W(x) dx \quad (3)$$

$$= \int_{\mathbb{R}} g(x) \frac{f_W(x)}{f_X(x)} f_X(x) dx \quad (4)$$

$$= \mathbb{E} \left( g(X) \frac{f_W(x)}{f_X(x)} \right). \quad (5)$$

**b,c**

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

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

The integrand is zero by choosing  $f_X(x) = f_X^*(x) = \frac{f_W(x)g(x)}{\alpha}$ , 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  $\alpha$ , there would be no need for the MC estimator in the first place.

**d**

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

$$\bar{\alpha} = \sum_{m=1}^M \frac{g_{\delta}(X_m)}{M}, \quad (7)$$

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

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

Let

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

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  $\eta = z^* \sqrt{\frac{\bar{\sigma}^2}{M}}$  and obtain the confidence interval  $[\bar{\alpha} - \eta, \bar{\alpha} + \eta]$ .

## 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  $\rho$  as follows

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

with

$$K(x) = 2^{-d} \kappa_{\|x\|_{L_2}} \quad (11)$$

In this estimation, we commit two errors.

### bias error

$$\rho(u) - \mathbb{E}(\hat{\rho}(y)) \quad (12)$$

$$= \rho(y) \int_{\mathbb{R}^d} h^{-d} K\left(\frac{y-z}{h}\right) \rho(z) dz \quad (13)$$

$$= \int_{\mathbb{R}^d} (\rho(y) - \rho(y-hz)) K(z) dz \quad (14)$$

$$\approx \int_{\mathbb{R}^d} \left( \rho(y) - \left( \rho(y) - h z_i \partial_{y_i} \rho(y) + h^2 z_i z_j \partial_{z_i z_j}^2 \rho(y) \right) \right) K(z) dz \quad (15)$$

$$= \frac{h^2}{2} \partial_{y_i y_j} \rho(y) \int_{\mathbb{R}^d} z_i z_j K(z) dz \propto h^2 \quad (16)$$

### statistical error

The variance is bounded by

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

and

$$\mathbb{E}\left(K^2\left(\frac{y-z}{h}\right)\right) \approx \int_{\mathbb{R}^d} K^2(z) h^d \left( \rho(y) - h z_i \partial_{y_i} \rho(y) + \frac{h^2}{2} z_i z_j \partial_{y_i y_j}^2 \rho(y) \right) dz, \quad (18)$$

thus

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

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

### Total error

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

$$\text{TSE} \approx Ah^4 + \frac{B}{Mh^d}. \quad (21)$$

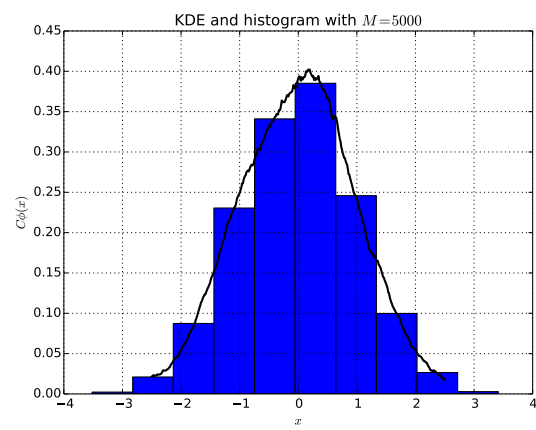
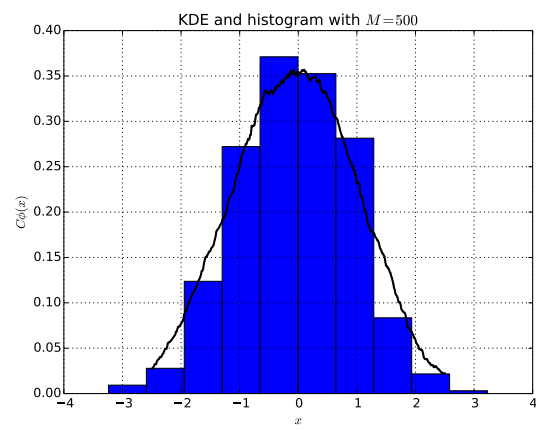
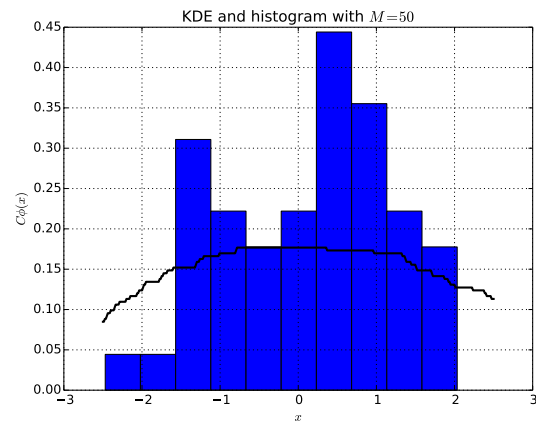
The critical point of the total square error is thus given by

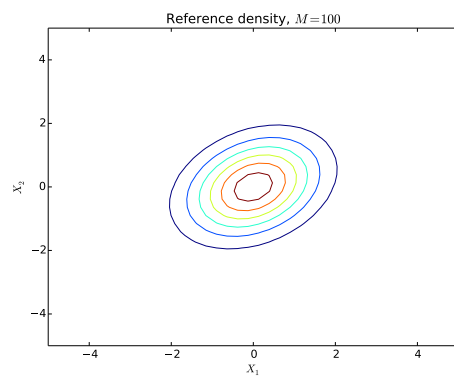
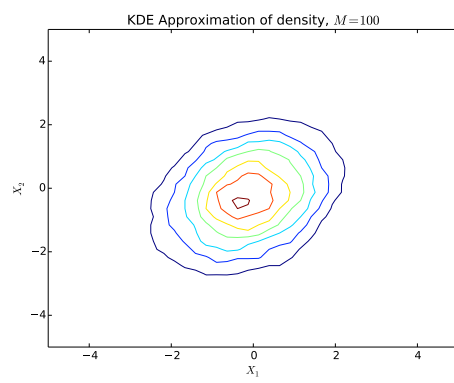
$$h^{4+d} \propto M^{-1}. \quad (22)$$

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

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

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 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

$$V_N = V_0 \prod_{n=1}^N R_n. \quad (24)$$

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

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

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  $\Phi^{-1}$  we can define

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

Then our control variate is

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

We have

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

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

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

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

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

$$\begin{aligned} \bar{C}_4 = \sum_{m=1}^{2M} \frac{\left( \left( \sum_{n=1}^N V_0 \prod_{m=1}^n R_n \right) - K \right)^+ + \beta (G_m - E(Q))}{M} \\ \sum_{m=1}^{2M} \frac{\left( \left( \sum_{n=1}^N V_0 \prod_{m=1}^n R_n \right) - K \right)^+ + \beta (Q_m - E(Q))}{M}, \end{aligned} \quad (32)$$

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  $n$ th time step for realisation  $m$ . The antithetic variables are defined as

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

and

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

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

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

In order to minimise the variance of the estimators, we set  $\beta = \rho \sqrt{\frac{\sigma_C^2}{\sigma_Q^2}}$  with  $\sigma_C^2$  being the variance of the plain vanilla MC realisations and  $\sigma_Q^2$  the

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.

variance of the control variate realisations. With  $M = 10000$ , we obtain the results in table 2: 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  $\bar{C}$  be given as:

$$\bar{C} = \sum_{m=1}^M \frac{X_i}{M}. \quad (37)$$

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

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

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

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

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

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

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

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



$p$	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.