Numerical Methods II
M. Giles

Problem sheet 2: solutions

1. It is unbiased because

$$\mathbb{E}[f(x_j)] = N \int_{j/N}^{(j+1)/N} f(x)\ \mathrm{d}x,$$

so

$$\mathbb{E}\left[N^{-1}\sum_{j=0}^{N-1} f(x_j)\right] = \int_0^1 f(x), \mathrm{d}x.$$

Since

$$f(x_j) = f\left(\frac{j+\frac{1}{2}}{N}\right) + f'\left(\frac{j+\frac{1}{2}}{N}\right)\ \frac{U-\frac{1}{2}}{N}\ +\ O\left(N^{-2}\right),$$

the variance is approximately

$$\left(N^{-1}\sum_j f'((j+\tfrac{1}{2})/N)\right)^2 N^{-2}\ \mathbb{V}[U] \approx \frac{1}{12\,N^2}\left(f(1)-f(0)\right)^2$$

(Note: this is poorer than the $O(N^{-3})$ variance of stratified sampling.)

2. Choosing $d = a/b$ we get

$$\mathrm{d}Y_t = c\ \exp(bt)\,\mathrm{d}W_t$$

and hence Y_T is Normally distributed with mean $Y_0 = x_0 - d$ and variance

$$\int_0^T c^2 \exp(2bt)\,\mathrm{d}t = \frac{c^2}{2b}\left(\exp(2bT)-1\right).$$

Therefore, X_T is also Normally distributed, with mean

$$\mu = d + \exp(-bT)(x_0 - d)$$

and variance

$$\sigma^2 = \frac{c^2}{2b}\left(1-\exp(-2bT)\right).$$

The p.d.f. is

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

so

$$\log p = -\tfrac{1}{2}\log(2\pi) - \tfrac{1}{2}\log(\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}.$$

We then get

$$\frac{\partial \log p}{\partial x_0} = \frac{\partial \log p}{\partial \mu} \frac{\partial \mu}{\partial x_0} = \frac{(x-\mu)}{\sigma^2} \; \exp(-bT)$$

and

$$\frac{\partial \log p}{\partial c} = \frac{\partial \log p}{\partial(\sigma^2)} \frac{\partial(\sigma^2)}{\partial c} = \left(-\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^4} \right) \frac{c}{b} \left(1 - \exp(-2bT)\right).$$

3. (i) $X = \Phi^{-1}(U)$ gives a vector of uncorrelated Normal random variables.

   If $Y = L\,X$ then Y is a vector of Normal random variables with covariance

$$\mathbb{E}[Y\,Y^T] = \mathbb{E}[L\,X\,X^T L^T] = L\,\mathbb{E}[X\,X^T]\,L^T = L\,L^T$$

   Hence we need L to be such that $L\,L^T = \Sigma$. The simplest definition is to use a Cholesky factorisation of Σ in which L is lower-triangular.

   (ii) In quasi-Monte Carlo simulation we generate quasi-uniform random points $U^{(n)}$ in the 5D unit hypercube in a way which ensures a much more uniform coverage than is achieved with pseudo-random points. Examples include Sobol points and rank-1 lattice rules.

   In the best case, the error from an estimate based on N quasi-random points is $O((\log N)^5 N^{-1})$, (but I wouldn't worry too much about the log term) whereas the Central Limit Theorem gives an $O(N^{-1/2})$ error for the standard Monte Carlo estimate using the pseudo-random points.

   In randomised QMC, the set of QMC points is randomised as a set (by a random offset for rank-1 lattice points, or by a random vector bit-wise exclusive-or operation for Sobol points, but I don't expect them to give the details). For each randomisation k, you compute an average P_k, and then you view the P_k as a collection of random samples in the usual way to obtain a mean and confidence interval.

   (iii) We know that $\mathbb{E}[S_i(T)] = S_i(0)e^{rT}$ so the idea with control variates is to use the unbiased estimator

$$N^{-1} \sum_n \left(P(U^{(n)}) - \sum_{i=1}^{5} \lambda_i (S_i(T) - \mathbb{E}[S_i(T)]) \right)$$

   The variance of this estimator is

$$N^{-1} \left(\mathbb{V}[P] - 2 \sum_{i=1}^{5} \lambda_i \operatorname{cov}(P, S_i(T)) + \sum_{i,j=1}^{5} \lambda_i\,\lambda_j \operatorname{cov}(S_i(T), S_j(T)) \right)$$

   which is minimised by choosing the λ_i so that

$$\sum_j \operatorname{cov}(S_i(T), S_j(T))\,\lambda_j = \operatorname{cov}(P, S_i(T)).$$

   This gives a set of 5 equations in 5 unknowns to be solved to give the λ_i.

(iv) The standard Monte Carlo estimator uses

$$N^{-1} \sum_n P(U^{(n)})$$

where the $U^{(n)}$ are uniformly distributed over the 5D space $(0,1)^5$.
If we restrict the components of U to lie within the given ranges, then the uniform pdf is

$$\prod_{i=1}^{5} (u_i - l_i)^{-1}$$

and so the unbiased Monte Carlo estimator is

$$N^{-1} \prod_{i=1}^{5} (u_i - l_i) \sum_n P(U^{(n)})$$

The variance of the standard MC estimator is

$$\mathbb{E}[P^2] - (\mathbb{E}[P])^2$$

where the expectation is over the full $(0,1)^5$ hypercube.
The variance of the new estimator is

$$\mathbb{E}_Q[R^2\, P^2] - (\mathbb{E}_Q[R\, P])^2$$

where $R \equiv \prod_{i=1}^{5} (u_i - l_i)$ is the Radon-Nikodym derivative and the expectation is now over the restricted hyper-rectangle.
Since

$$\mathbb{E}_Q[R\, P] = \mathbb{E}[P]$$

and

$$\mathbb{E}_Q[R^2\, P^2] = R\; \mathbb{E}[P^2]$$

the new variance is

$$R\; \mathbb{E}[P^2] - (\mathbb{E}[P])^2$$

which is less than the original variance.
We can generate U_i by defining it as

$$U_i = l_i + V_i (u_i - l_i)$$

where V_i is a $(0,1)$ uniform random variable.

4. (i) CDF is

$$C(x) = \begin{cases} \frac{1}{2}\exp(\lambda\, x), & x < 0 \\ 1 - \frac{1}{2}\exp(-\lambda\, x), & x \geq 0 \end{cases}$$

Given uniformly distributed sample $U^{(n)}$, set

$$X^{(n)} = C^{-1}(U^{(n)}) = \begin{cases} \lambda^{-1}\log(2\, U^{(n)}), & U^{(n)} < 1/2 \\ -\lambda^{-1}\log(2 - 2\, U^{(n)}) & U^{(n)} \geq 1/2 \end{cases}$$

(ii) Using N independent samples $X^{(n)}$, an estimate of the expected payoff is given by the sample mean

$$\hat{\mu} = N^{-1}\sum_n P^{(n)},$$

where $P^{(n)} \equiv \max(S_0 \exp(X^{(n)}) - K, 0)$.
When N is large, this estimate is asymptotically Normally distributed with the correct mean, and variance σ^2/N where σ is asymptotically equal to the empirical variance of the sample,

$$\hat{\sigma}^2 = N^{-1}\sum_n \left(P^{(n)} - \hat{\mu}\right)^2.$$

(Comment: $\hat{\sigma}^2 N/(N-1)$ is an unbiased estimator for σ^2 but they're not required to state this, just that $\hat{\sigma} \to \sigma$ as $N \to \infty$. In practice, N is so large that the slight bias is irrelevant.)
Hence, a 99% confidence interval is given by $\hat{\mu} \pm c\,\hat{\sigma}/\sqrt{N}$, and $c > 0$ is a constant such that

$$\Phi(c) - \Phi(-c) = 0.99,$$

or equivalently (since $\Phi(c) = 1 - \Phi(-c)$) $\Phi(c) = 0.995$ or $\Phi(-c) = 0.005$, where $\Phi(x)$ is the Normal CDF.

(iii) The pathwise sensitivity with respect to S_0 is given by

$$N^{-1}\sum_n \frac{\partial P}{\partial S}\,\frac{\partial S}{\partial S_0}$$

where

$$\frac{\partial P}{\partial S} = \mathbf{1}_{S>K}$$

and

$$\frac{\partial S}{\partial S_0} = \frac{S}{S_0}.$$

Similarly, the pathwise sensitivity with respect to λ is given by

$$N^{-1}\sum_n \frac{\partial P}{\partial S}\,\frac{\partial S}{\partial \lambda}$$

where

$$\frac{\partial S}{\partial \lambda} = \frac{\partial S}{\partial X}\,\frac{\partial X}{\partial \lambda} = S\,\left.\frac{\partial X}{\partial \lambda}\right|_U = -\,S\,X\,/\,\lambda$$

(iv) The probability density functions for X and S are related through

$$p_X \, dX = p_S \, dS$$

so

$$p_S = p_X \frac{dX}{dS} = p_X / S = \frac{\lambda}{2} \exp(-\lambda \, |X|) / S.$$

Since $S/S_0 = \exp(X)$ it follows that

$$p_S = \begin{cases} \frac{\lambda}{2} S^{\lambda-1} S_0^{-\lambda}, & S < S_0 \\ \frac{\lambda}{2} S^{-\lambda-1} S_0^{\lambda}, & S \geq S_0 \end{cases}$$

(v) Pathwise sensitivity analysis could not be used in this case because of the discontinuity in the payoff. Since $\partial P / \partial S = 0$, pathwise sensitivity would give a value of zero, for any initial asset price S_0, which is clearly wrong. (This is the argument I give in lectures.)

Using LRM, we have the estimate

$$N^{-1} \sum_n \mathbf{1}_{S>K} \frac{\partial (\log p_S)}{\partial S_0}$$

where

$$\frac{\partial (\log p_S)}{\partial S_0} = \begin{cases} -\lambda / S_0, & S < S_0 \\ +\lambda / S_0, & S \geq S_0 \end{cases}$$