Numerical Methods II
M. Giles

Problem sheet 1

1. A change of variables from $x \equiv (x_1, x_2, \ldots x_n)$ to $y \equiv (y_1, y_2, \ldots y_n)$ leads to

$$\int f \; p_x(x) \; \mathrm{d}x = \int f \; p_x(x) \; J^{-1}(x) \; \mathrm{d}y$$

where J is the determinant of the Jacobian of the mapping $\dfrac{\partial(y_1, y_2, \ldots y_n)}{\partial(x_1, x_2, \ldots x_n)}$.
Hence, if $p_x(x)$ represents the probability density function for x then
$p_y(y) = p_x(x) \; J^{-1}(x)$ is the probability density function for y.

Use this to confirm that the Box-Muller method does convert two independent uniformly distributed random variables, into two independent Normally distributed random variables as stated in lecture 1.

2. If X is a random variable with zero mean, prove that

$$\widehat{\sigma}^2 = \frac{N}{N-1} \left(N^{-1} \sum_{n=1}^{N} X_n^2 - \left(N^{-1} \sum_{n=1}^{N} X_n \right)^2 \right)$$

is an unbiased estimator for $\mathbb{V}[X]$, i.e. $\mathbb{E}[\widehat{\sigma}^2] = \mathbb{V}[X]$.

It can be proved (see
`http://mathworld.wolfram.com/SampleVarianceDistribution.html`) that

$$\mathbb{V}[\widehat{\sigma}^2] = N^{-1} \mu_4 - \frac{(N-3)}{N\,(N-1)} \mu_2^2$$

where

$$\mu_2 \equiv \mathbb{E}[X^2], \quad \mu_4 \equiv \mathbb{E}[X^4]$$

Given this, if you assume that the distribution of $\widehat{\sigma}^2$ is approximately Normal, determine how big N must be to get the correct value for σ^2 to within $\pm 10\%$ in each of the following two cases:

(a) X is a $N(0,1)$ random variable

(b) X takes value $+1$ with probability $p \ll 1$, value -1 with probability p, and value 0 with probability $1-2p$.

3. Suppose there are two control variates g and h with known expectations, and they are to be used by computing the average of

$$f - \lambda\,(g - \mathbb{E}[g]) - \gamma\,(h - \mathbb{E}[h])$$

for N independent samples to get an estimate for $\mathbb{E}[f]$.

How would you choose the values for λ and γ to minimise the variance of this estimator?

4. (Exam question from 2007-8)

Suppose one wishes to use Monte Carlo simulation to estimate the value of

$$\mathbb{E}[f(X)]$$

where $f(X)$ is a twice-differentiable function (with a continuous second derivative) and X is Normally distributed with zero mean and variance $\varepsilon^2 \ll 1$.

(a) Show that the standard Monte Carlo estimator with N samples has variance which is $O(\varepsilon^2/N)$, whereas the use of antithetic variables reduces the variance to $O(\varepsilon^4/N)$.

(b) By considering

$$g(X) = \tfrac{1}{2}\,(f(X) + f(-X)) - f(0)$$

explain how the variance can be further reduced through the use of importance sampling.

5. Suppose that X is a $N(0,1)$ random variable and we wish to improve the variance of a Monte Carlo estimate for $\mathbb{E}[f(X)]$ by using importance sampling, either by changing the mean so that X is taken from $N(\mu,1)$ or by changing the variance so that X is taken from $N(0,\sigma^2)$.

(a) In the case $f(X) = e^X$, explain why it is better to change the mean, and find the optimal choice for μ

(b) In the case $f(X) = X^4$, explain why it is better to change the variance, and find the optimal choice for σ^2

6. Slide 13 in Lecture 4 suggested a combination of stratified sampling and antithetic variables for a 1D integration. By modifying the analysis in slides 11-12, determine the asymptotic behaviour of the variance in this case, assuming just 2 samples per stratum (and that $f(U)$ has a bounded second derivative).