Monte Carlo Methods, Fall 2007, Courant Institute, NYU
Homework 1, Due September 25

1. Suppose $f(x)$, $g(x)$, and $h(x)$ are three probability densities with

$$\frac{f(x)}{g(x)} \leq c \ , \quad \text{and} \ \frac{g(x)}{h(x)} \leq c \ ,$$

for all x. Suppose we have an h sampler and want an f sampler based on rejection from h. Consider the following two methods:

Direct method: Generate h samples and accept with probability proportional to $f(x)/h(x)$.

Indirect method: Generate g samples by rejection from h, then reject from these g samples to create f samples.

Measure the efficiency of either algorithm by the expected number of h samples needed to produce an f sample. Is it possible that an indirect method is more efficient than the direct method? Either give an example or show that it is not possible.

2. Program the Box-Muller algorithm for sampling from a standard normal density. Verify that your samples have the correct distribution using either the histogram method or the kernel density estimation method.

3. Use the two dimensional histogram or kernel density estimation algorithms to verify that the pair (X, Y) produced by Box-Muller is has the two dimensional density $\frac{1}{2\pi}e^{-(x^2+y^2)/2}$.

4. One feature of the rejection algorithm is that it can work for a probability density of the form $f(x) = \frac{1}{Z}e^{-\phi(x)}$ (or related forms) even when the normalization constant Z is not known.

   (a) We want random variables $X = (X_1, X_2, X_3)$ uniformly distributed in the unit ball $X_1^2+X_2^2+X_3^2 \leq 1$. Write a sampler that does this by rejection from uniform density inside the cube of side 2 that contains the unit ball. Use the histogram method to check that $L = X_1+2X_2+3X_3$ and $R = (X_1^2+X_2^2+X_3^2)^{1/2}$ have the correct one dimensional distributions. This sometimes is called checking moments. We do it because it is hard to do density estimation directly on a three dimensional distribution. Note that L is a linear functional of X and all linear functionals have the same distribution, modulo scaling. Why do we test a complicated functional rather than, say, X_1?

   (b) Let $g(x)$ be the probability density sampled in part (a). Now we want to sample $f(x) = \frac{1}{Z(\lambda)}e^{\lambda x_1}g(x)$ without calculating Z. Suggest and program a way to do this by rejection from g samples and verify that the results are correct for moderate values of λ using a histogram or kernel method. Show computationally that this method is inefficient (many g samples per f sample) if λ is large. Why is this true?