Monte Carlo, Fall, 2005
Homework 2
December 22

This is a different approach to some of the problems of the first assignment. Suppose Y is a random variable with $|Y| \leq 1$ and probability density $h(y) = \frac{3}{4}(1-x^2)$ for $-1 \leq y \leq 1$. Let $X = (Y_1, \cdots, Y_n)$ with the $Y_k \sim h$ independent. You wrote a sampler for Y in the previous homework. The density for X is $f(x) = \prod_k h(y_k)$. We want to sample the probability density of X given that $S = Y_1 + \cdots + Y_n > nb$. Let $\chi_b(X) = 1$ if $S \geq nb$ and $\chi_b(X) = 0$ if $S < nb$. Let $Z(b) = P(S > nb)$. The conditional probability density for X is

$$f_b(X) = \frac{1}{Z(b)} \chi_b(X) f(x) \ . \tag{1}$$

Note that none of the calculations require knowing Z.

1. Write a program to sample $f_b(x)$ using the following Metropolis strategy. Cycle through the components, X_k. For each k let $\widetilde{X}_k$ be an independent sample from the density h. If $\widetilde{S} \geq nb$, accept the new component. Otherwise reject it and keep Y_k. You will need to chose an initial sample that satisfies $S \geq nb$.

2. Explain why this strategy preserves the density (1).

3. Test the correctness of your sampler by checking that it produces the correct conditional densities $h(y \mid S \geq nb)$ that are predicted by Cramer's theorem: $h(y \mid S \geq nb) \approx h_\lambda(y)$ for an appropriate λ. Use a histogram. Try a few values of b and n. For small n, the approximation $h(y \mid S \geq nb) \approx h_\lambda(y)$ is not accurate.

4. We want to estimate

$$A = E\left[S \mid S \geq nb\right] \ .$$

Do a length L run and let $S(t)$ be the value of S after t Metropolis sweeps through the lattice. That is, measure S after each Y_k has been resampled once, regardless of whether the change in Y_k was accepted or rejected. Estimate the autocorrelation function of $C(s) = \text{cov}(S(t), S(t+s))$ for large t as explained in the notes. Estimate the autocorrelation time using $\widehat{C}(s)$ for $s \leq 10$. A run of length $L = 10^6$ should be plenty.

5. Estimate the quantity $\text{var}(\widehat{A_L})$ using $M = 10^3$ independent MCMC runs. How close is this direct estimate to the one from part 4? Use $b = 3$ and $n = 50$.