## STA3431H (Monte Carlo Methods), Winter 2009

Homework #3

**Due:** In class by 2:10 p.m. **sharp** on Monday March 23.

**NOTE:** All the same "Notes" from HW#1 still apply. In particular, **homeworks which are late, even by one minute, will be penalised.**

**The assignment:**

**1.** Let $\mathcal{X} = \{1, 2, 3\}$, and $\pi(1) = 1/6$, $\pi(2) = 1/3$, $\pi(3) = 1/2$, with $P(i, \{j\})$ given by the matrix:

$$P = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

(so e.g. $P(2, \{1\}) = 1/4$ and $P(1, \{2\}) = 0$). Determine (with explanation) the following.

**(a)** Is this Markov chain reversible with respect to $\pi$?

**(b)** Is $\pi$ a stationary distribution for this Markov chain?

**(c)** Is this Markov chain irreducible?

**(d)** Is this Markov chain aperiodic?

**(e)** Does $\lim_{n\to\infty} P^n(i, \{j\}) = \pi(j)$ for all $i, j \in \mathcal{X}$?

**(f)** Compute $P(3, \{1\})$ and $P^2(3, \{1\})$ and $P^3(3, \{1\})$ analytically (i.e., precisely).

**(g)** Compute $P^8(3, \{1\})$ numerically, using a computer program. (Do not use any pre-programmed matrix function packages; instead write the program yourself, from scratch, perhaps using mathematical induction.)

**(h)** Compute (numerically) the value of $|P^8(3, \{1\}) - \pi(1)|$.

**2.** Consider an independence sampler algorithm on $\mathcal{X} = (1, \infty)$, where $\pi(x) = 3\,x^{-4}$ and $q(x) = r\,x^{-r-1}$ for some choice of $r > 0$, with identity functional $h(x) = x$.

**(a)** For what value of $r$ will the algorithm provide i.i.d. samples?

**(b)** For what values of $r$ will the sampler be geometrically ergodic?

**(c)** For $r = 1/20$, find a number $n$ such that $D(x, n) < 0.01$ for all $x \in \mathcal{X}$.

**(d)** Run the algorithm (on a computer) for the cases $r = 1/20$ and $r = 12$, each with $M = 10^5$ and $B = 10^4$, to estimate $\mathbf{E}_\pi(h)$. [Hint: if $U \sim \text{Uniform}[0, 1]$, then what is the density of $U^{-1/r}$?]

**(e)** For both cases, use repeated runs (from an overdispersed starting distribution) to approximate the standard error of the estimate.

**(f)** For both cases, use autocorrelation estimates to compute (from scratch, i.e. without using any pre-programmed correlation estimator) the standard error factor "varfact", and hence obtain different approximations of the standard error of the estimate.

**(g)** Compare the results of parts (e) and (f).

**(h)** Use the results of parts (e) and (f) to compare the two cases, to determine (with explanation) which one is a "better" sampling algorithm.

**3.** Repeat question 3 from HW#2 (sampling from the variance components model for the dyestuff data, using a usual Metropolis algorithm), except this time using (a) a systematic-scan Gibbs sampler, and (b) a Metropolis-within-Gibbs algorithm (with symmetric normal proposal increments having fixed appropriately-chosen variance), instead of the previous usual Metropolis algorithm. Then, for all three algorithms (Metropolis, and Gibbs, and Metropolis-within-Gibbs), compute the "varfact" factor, and discuss the relative merits of the three algorithms for this example.

**4.** Let $\mathcal{X} = \mathbf{R}$, and let $\pi(x) = c\, g(x)$, where $g(x) = e^{-|x|/10}(1 + \cos(x)\sin(x^3))$, and let $h(x) = x + x^2$. With fixed appropriate choice of $M$ and $B$ and the starting distribution $\mathcal{L}(X_0)$, estimate $\mathbf{E}_\pi(h)$ in each of two different ways:

**(a)** With a usual Metropolis algorithm for $\pi$, with usual proposals $Y_n \sim N(X_{n-1}, 1)$.

**(b)** With a Metropolis-Hastings algorithm with proposals $Y_n \sim N(X_{n-1}+\frac{1}{2}g'(X_{n-1}),\, 1)$. [Notes: Here $g'(x)$ is the usual derivative of $g$. The resulting algorithm is called a *Langevin algorithm*, since it is motivated by a Langevin diffusion.]

**(c)** For both algorithms, compute the standard error factor "varfact", and use this to compare the two algorithms to determine which one is better.