STA 3431 (Monte Carlo Methods), Fall 2017

Homework #2 Assignment: worth 20% of final course grade.

Due: In class at 10:10 a.m. sharp on Monday November 13.

GENERAL NOTES:

- **Late homeworks, even by one minute, will be penalised!**
- Include at the top of the first page: Your name and student number and department and program and year and e-mail address.
- Homework assignments are to be solved by each student individually. You may discuss assignments in general terms with other students, but you must solve it on your own, including doing all of your own computing and writing.
- For full points, you should provide very complete solutions, including explaining all of your reasoning clearly and neatly, performing detailed Monte Carlo investigations including multiple runs and error estimates as appropriate, justifying all of the choices you make, etc.
- You may use results from lecture, but clearly indicate when you do so.
- When writing computer programs for homework assignments:
  - R is the "default" computer programming language and should normally be used for homework (and tests). You may perhaps use other standard computer languages like C and C++ and Java and Python with prior permission from the instructor.
  - You should include your complete source code and your program output.
  - Programs should be clearly explained, with comments, so they are easy to follow.
  - You should always consider such issues as the accuracy and consistency of the answers you obtain.

THE ACTUAL ASSIGNMENT:

1. [6] For this question, again let A, B, C, and D be the last four digits of your student number, and again let $g : \mathbf{R}^5 \to [0, \infty)$ be the function defined by:

$$g(x_1, x_2, x_3, x_4, x_5) \;=\; (x_1 + A + 2)^{x_2+3}\Big(1 + \cos[(B+3)x_3]\Big)(e^{(12-C)x_4})|x_4 - 3x_5|^{D+2} \prod_{i=1}^{5} \mathbf{1}_{0<x_i<1}\,.$$

Let $\pi(x_1, x_2, x_3, x_4, x_5) = c\, g(x_1, x_2, x_3, x_4, x_5)$ be the corresponding five-dimensional probability density function, with unknown normalising constant c. Write a program to estimate $\mathbf{E}_\pi[(X_1 - X_2)/(1 + X_3 + X_4X_5)]$ using an MCMC algorithm of your choice, and obtain the best estimate you can. Include some discussion of the reasons for your choice, accuracy, uncertainty, standard errors, etc. Also, discuss the advantages and disadvantages of your approach compared to the methods used for this problem on Homework #1.

2. Consider an independence sampler algorithm on $\mathcal{X} = (1, \infty)$, where $\pi(x) = 5\, x^{-6}$ and $q(x) = r\, x^{-r-1}$ for $x > 1$ for some choice of $r > 0$, with identity functional $h(x) = x$.

(a) [2] Determine mathematically what value of r will provide i.i.d. samples.

(b) [3] Determine mathematically what values of r will make the sampler be geometrically ergodic.

(c) [3] For $r = 1/20$, find (with explanation) a number n such that $D(x, n) < 0.01$ for all $x \in \mathcal{X}$.

(d) [6] Write and run a computer program to estimate $\mathbf{E}_\pi(h)$ with this algorithm in the two cases $r = 1/20$ and $r = 10$, each with $M = 10^5$ and $B = 10^4$. Estimate the corresponding standard errors by two different methods: (i) using "varfact", and (ii) from repeated independent runs.

(e) [3] Discuss and compare the standard errors estimated by each of the two methods in each of the two cases, including discussion of which method is "better" for assessing uncertainty, and which case is a "better" sampling algorithm.

3. Consider the standard variance components model described in lecture, with $K = 6$ and $J_i \equiv 5$, and $\{Y_{ij}\}$ the famous "dyestuff" data (from the file "Rdye"). Consider two sets of prior values: (i) the "reasonable" values $a_1 = a_2 = a_3 = 1500$, $b_1 = b_2 = b_3 = 1500^2$; and (ii) the "unreasonable" values $a_1 = a_2 = a_3 = b_1 = b_2 = b_3 = 100$. For each set of prior values, estimate (as best as you can, together with a discussion of accuracy etc.) the posterior mean of W/V, in each of three ways:

(a) [8] With a random-walk Metropolis algorithm.

(b) [8] With a Metropolis-within-Gibbs algorithm.

(c) [8] With a Gibbs sampler. [Note: first derive from scratch all of the conditional distributions, whether or not they were already described in lecture.]

(d) [3] Finally, discuss the relative merits of all three algorithms for this example, for each of the two sets of prior values.

[END; total points = 50]

Reminder: There will be a sit-down test on Monday Nov 20 in class, worth 30% of your final grade. No aids allowed.