STA410/2102 (Statistical Computation), Fall 2007

Homework #4

Due: In class by 6:10 p.m. **sharp** on Tuesday December 4. (If you prefer, you may bring your assignment to the instructor's office, Sidney Smith Hall room 6024, any time before it is due; slide it under the door if he is not in.)

Warning: Late homeworks, even by one minute, will be penalised!

Reminder: The final exam will be 7–10 p.m. on Wed Dec 12, in NR25.

Note: 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.

Include at the TOP of the first page: Your name and student number.

Reminder: When writing programs in R for homework, you should include some comment lines to explain what you are doing. Also, you should hand in both the program itself, and the program's output.

The assignment:

1. [10 points] Run a linear congruential pseudorandom number generator in R to obtain Monte Carlo estimates of each of $\mathbf{E}(U)$, $\mathbf{E}(U^2)$, and $\mathbf{E}(\cos(U))$, where $U \sim$ Uniform$[0,1]$, and also of $\mathbf{E}(U_{n-1}U_n)$ where $\{U_n\}$ are the outputs from the generator. Run this for several different choices of seed values, and several different choices of the parameters m, a, and b (including the choice $m = 2^{32}$, $a = 69069$, $b = 23606797$). For each choice, discuss how accurately the results approximate those of a true i.i.d. Uniform$[0,1]$ sequence. Try to find one choice for which the results are very good, and another choice for which the results are poor. [Hint: you may use the file "Rrng" if you wish.]

2. [5 points] Use your linear congruential pseudorandom number generator from the previous question to obtain a Monte Carlo estimate of $\mathbf{E}(Z^4)$, where $Z \sim N(0,1)$. [Note: do not use R's built-in functions "dnorm", "dunif", etc.]

3. Consider the Variance Components Model, with prior parameter values $a_1 = b_1 = a_3 = 0, a_2 = 0.5, b_2 = 1$, and $b_3 = 10^{12}$. Suppose $K = 2$ and $J = 3$, with $Y_{ij} = i + j^2$. Compute (with full explanation of your methods used) the posterior mean of μ, in each of four different ways, as follows (over). [You may use the file "Rvarcomp" if you wish.]

(a) [20 points] By numerical integration. [Hint: you will have to "cut off" the infinite integrals; be sure to discuss how you approached that. Also, it may be difficult to get a very accurate estimate or use a very large value of "M"; just do your best, and explain the difficulties that arise.]

(b) [10 points] By Monte Carlo integration. [Hint: first re-write the corresponding integrals as expected values.]

(c) [10 points] By a Metropolis algorithm, with full proposal increment distribution given by a standard multi-variate normal.

(d) [10 points] By a Metropolis-within-Gibbs algorithm, with proposal increment distributions of your choice.

4. [15 points] Consider again the Variance Components Model, again with prior parameter values $a_1 = b_1 = a_3 = 0, a_2 = 0.5, b_2 = 1$, and $b_3 = 10^{12}$. Suppose $K = 6$ and $J = 5$, with $\{Y_{ij}\}$ given by the famous "dyestuff" data set (defined as "Ydye" in the file "Rvarcomp"). Estimate the posterior means of V, W, and μ. [Hint: this question is somewhat open-ended; you may use whatever method(s) you wish to try to get an accurate estimate. Explain all your steps. If you use a Metropolis or Metropolis-within-Gibbs algorithm, then ideally you should use a convergence diagnostic too.]