STA 3431 (Monte Carlo Methods), Fall 2019

Homework #1 Assignment: worth 20% of final course grade.

Due: In class at 10:10 a.m. **sharp** on Monday October 21.

Or: By e-mail to j.rosenthal@math.toronto.edu as a single pdf file (of reasonable size) by 9:30 a.m. on Monday October 21.

GENERAL NOTES:

- **Late homeworks, even by one minute, will be penalised!**
- 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, justifying 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. For this first assignment, you must use R for Questions 2 and 3 (though if you wish then you can also solve them using a different language and then compare which is better). For the other questions, you may use another standard computer language like C or C++ or Java or Python if you explain that.
  - 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 the accuracy and consistency of the answers you obtain.

THE ACTUAL ASSIGNMENT:

1. Tell me your name and student number and department and program and year, and also your e-mail address.

2. [6] **(a)** Write a computer program in R to generate pseudorandom Uniform[0,1] numbers, using a method of your choice. Your program should just use simple arithmetic, and should not use any built-in randomness functions. It should not be identical to the version presented in class. Explain your reasons for your choice of method.

(b) Use your program to generate and plot 500 independent pseudorandom Uniform[0,1] numbers. Discuss the extent to which your plot appears to show true Uniform[0,1] numbers.

(c) Perform (with explanation) a few statistical tests of your own choosing, to see how random/uniform/independent your generator "seems" to be.

3. [6] **(a)** Write a computer program in R to compute a good "classical" (i.i.d.) Monte Carlo estimate (including standard error and 95% confidence interval) of $\mathbf{E}[YZ^6 \sin(YZ^2)]$, where $Y \sim \text{Exponential}(3)$ and $Z \sim \text{Normal}(0,1)$ are independent. Your program should use your own pseudorandom function from the previous question, and should not use any built-in randomness functions (for any distributions).

(b) Run your program several times, and produce a final estimate.

(c) Using only Monte Carlo (not numerical integration), discuss how accurate you think your estimate is.

4. [6] Re-write the integral

$$\int_1^\infty \left(\int_{-\infty}^\infty \left[1 + x^4 + (y+5)^2 + \sin(xy^2)\right]^{-x^2-y^4+2} dy \right) dx$$

as some expected value, and then write a computer program (with explanation) to estimate its value using a Monte Carlo algorithm of your choice. (You may use the computer's built-in pseudorandom functions if you wish.) Then, produce a final estimate, and discuss (using only Monte Carlo, not numerical integration) the extent to which your algorithm does or does not work well, and how accurate you think your estimate is.

5. For this question, let A, B, C, and D be the last four digits of your student number, in order. (So, for example, if your student number were 840245070*, then $A = 5$, $B = 0$, $C = 7$, and $D = 0$.) And, let $g : \mathbf{R}^5 \to [0, \infty)$ be the function defined by:

$$g(x_1, x_2, x_3, x_4, x_5) =$$

$$(x_1 + A + 2)^{x_2+3}\Big(1 + \cos\big[2x_2 + 3x_3 + 4x_4 + (B+3)x_5\big]\Big)\, e^{(12-C)x_4}\, e^{-(D+2)(x_4-3x_5)^2} \prod_{i=1}^{5} \mathbf{1}_{0<x_i<1}\,.$$

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.

(a) Identify the values of A, B, C, and D. (This should be easy!)

(b) [6] Write a computer program to get a good estimate of $\mathbf{E}_\pi[(X_1 - X_2)/(3 + X_3X_4 + X_5)]$ using an importance sampler with your choice of function "f". (You may use the computer's built-in pseudorandom functions if you wish.) Discuss the reasons for your choice of f, and the extent to which your algorithm does or does not work well. Then, produce a final estimate, and discuss how accurate you think your estimate is.

(c) [6] Write a computer program to get a good estimate of $\mathbf{E}_\pi[(X_1 - X_2)/(3 + X_3X_4 + X_5)]$ using a rejection sampler with your choice of function "f". (You may again use the computer's built-in pseudorandom functions if you wish.) Discuss the reasons for your choice of f, and the extent to which your algorithm does or does not work well. Then, produce a final estimate, and discuss how accurate you think your estimate is.

[END; total points = 30]

---

*(Historical note: this was the instructor's actual student number when he was a UofT undergraduate student during 1984–88.)