Monte Carlo Simulation (MA323)
A Model Report
End-semester Examination (Lab)

1. To generate a sample path of a standard Brownian motion in the time interval [0, 1] with 5000 generated values, we can use the following algorithm.

   1: Generate $Z_1, Z_2, \ldots, Z_{5000}$ from $N(0, 1)$ distribution using Box-Muller method.
   2: Set $W_0 = 0$.
   3: **for** $i = 1, 2, \ldots, 5000$ **do**
   4: Set $W_i = W_{i-1} + \frac{Z_i}{\sqrt{5000}}$.
   5: **end for**
   6: Return W.

   The above algorithm is implemented in R software to generate one sample path. Therefore, the above algorithm is ran for 10 times to generate 10 sample paths. The seed is taken to be 123. The plots of 10 sample paths are given in Figure 1. The function `rnorm` and packages `tidyr` and `ggplot2` are used.

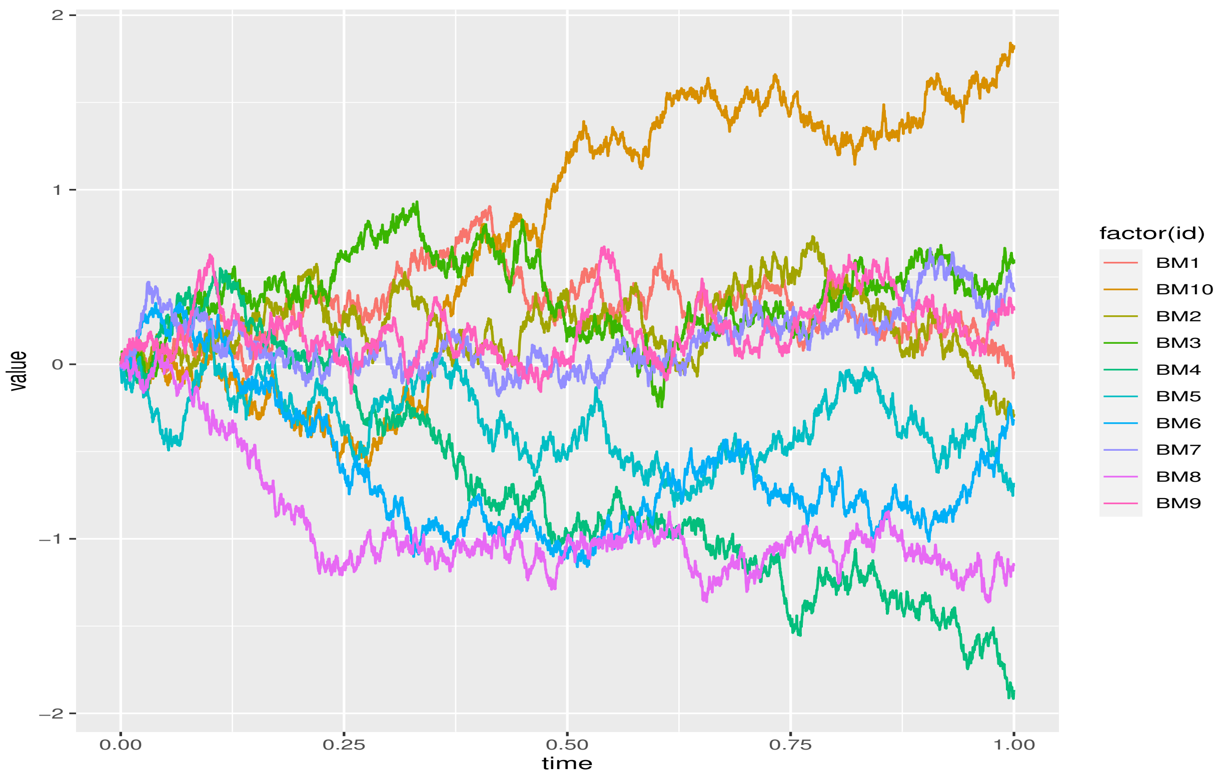

Figure 1: Plot of Sample Paths of Standard Brownian Motion

2. (a) To find Monte Carlo estimate of $P(X > 4)$, $N = 10^4$ random numbers are generated from $N(0,\, 1)$ distribution. Then, $P(X > 4)$ is approximated by

$$P(X > 4) \approx \frac{1}{10^4} \sum_{i=1}^{10^4} 1_{\{X_i > 4\}} = \widehat{\mu}, \text{ say,}$$

where X_i are the generated random numbers.

To compute the estimated standard error, we form an array of the values in the above sum, and calculate the standard deviation of the array. We finally divide that by $\sqrt{N}$. A 99% confidence interval for $P(X > 4)$ is given by as $[\widehat{\mu} - 2.58s,\, \widehat{\mu} + 2.58s]$, where s is estimated standard error.

This is implemented in R software with the help of `rnorm, mean`, and `sd` functions. The seed value is taken to be 111. The outputs are reported in Table 1. Note that probability is always a non-negative quantity. However, the lower bound of confidence interval turns out to be negative. Therefore, it is replaced by 0.

(b) To implement the importance sampling technique, the importance density is taken to be

$$q(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-5}{3}\right)^2} \quad \text{for } x \in \mathbb{R}.$$

The choice of this density was motivated by the fact that our region of interest is $[4,\, \infty)$. The variance was chosen as 9 because we also want some variability of values in the domain.

To do this, N random numbers are generated from $N(5,\, 9)$ distribution, and then, $P(X > 4)$ is approximated by

$$\widehat{\mu}_{\text{imp}} = \frac{1}{N} \sum_{i=1}^{N} 1_{\{X > 4\}} \frac{p(X_i)}{q(X_i)} = \frac{3}{N} \sum_{i=1}^{N} 1_{\{X_i > 4\}} \exp\left[-\frac{1}{18}(8X_i^2 + 10X_i - 25)\right],$$

where $p(\cdot)$ and $q(\cdot)$ are, respectively, the densities of $N(0,\, 1)$ and $N(5,\, 9)$ and X_i's are the generated random numbers.

To compute the estimated standard error and confidence interval, we follow the same steps as Part (a). The results are reported in the Table 1.

Table 1: Results of Question 2

	Estimate	Standard Error	Confidence Interval
Simple MC	10^{-4}	10^{-4}	$(0.00,\, 3.58 \times 10^{-4})$
Importance Sampling	3.36×10^{-5}	1.29×10^{-6}	$(3.03 \times 10^{-5},\, 3.69 \times 10^{-5})$

(c) As is easily observed, the ratio of standard error of simple Monte Carlo to that of importance sampling is 7.75 (approx.). This is an improvement. The confidence intervals are accordingly much smaller in the importance sampling case.

3. **Correction:** The probability density function given in the question is not correct. The correct probability density function is

$$f(x,\, y) = \begin{cases} \frac{9}{4} y^2 e^{-\left(\frac{3}{2}+x\right)y} & \text{if } x > 0,\, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The following solution is written based on the above probability density function.

(a) The conditional probability density function of X given $Y = y > 0$ is

$$f_{X|Y}(x|y) \propto e^{-yx} \quad \text{for } x > 0.$$

Therefore, $X|Y = y \sim Exp(y)$ for $y > 0$. Now,

$$E\left[h(X,\, Y)|Y\right] = \frac{\ln(1+Y)}{Y}.$$

(b) Note that

$$E\left[h(X,\, Y)\right] = EE\left[h(X,\, Y)|Y\right] = E\left[\frac{\ln(1+Y)}{Y}\right].$$

The marginal probability density function of Y is

$$f_Y(y) = \begin{cases} \frac{9}{4} y e^{-\frac{3}{2}y} & \text{if } y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $Y \sim Gamma(2,\, \frac{3}{2})$. Now, the following algorithm can be used to generate a random number from $Gamma(2,\, \frac{3}{2})$ distribution.

1: Generate U_1 and U_2 from $U(0,\, 1)$ distribution.
2: Set $Y = -\frac{2}{3} \ln\left(U_1 U_2\right)$.
3: Return Y.

Therefore, using conditioning technique, $\mu = E\left(h(X,\, Y)\right)$ can be approximated by

$$\widehat{\mu} = \frac{1}{N} \sum_{i=1}^{N} \frac{\ln(1+Y_i)}{Y_i},$$

where Y_i's are random number generated from $Gamma(2,\, \frac{3}{2})$ distribution using above algorithm.

The above Monte Carlo method is implemented in R with $N = 10^4$ and seed 123. The estimate of expectation of $h(X,\, Y)$ is 0.67 and the corresponding estimate of variance is 1.8×10^{-6}.