STAT 243 HW#4

1. We use the Monte Carlo method to evaluate each of the following integrals:

a.

$$\int_0^1 x^2 dx$$

Here, we let $f(x) = \frac{1}{1-0}$, a Uniform(0,1) distribution, and we let $h(x) = x^2$. We take 1,000,000 samples from a Uniform(0,1), and using the Monte Carlo method we obtain an estimate of the evaluated integral as .333328, or equivalenty $\frac{1}{3}$.

b.

$$\int_0^1 \int_{-2}^2 x^2 \cos(xy) \ dxdy$$

Here, since we have two variables, instead of having $f(x)$ and $h(x)$, we will instead have $f(x,y)$ and $h(x,y)$. Here, we let $f(x,y) = \frac{1}{4}$ and we let $h(x,y) = 4x^2\cos(xy)$. Thus, we see that $\int_0^1 \int_{-1}^2 \frac{1}{4} = 1$, which satisfies our condition for $f(x,y)$. We sample from a Uniform(-2,2) for x and from a Uniform(0,1) for y, each 1,000,000 times, and we plug these values into our function $h(x,y)$. Thus, taking 1,000,000 samples from the two Uniform distribution and using the Monte Carlo method, we obtain an estimate of the evaluated integral of 3.480481.

c.

$$\int_0^\infty \frac{3}{4} e^{-\frac{x^3}{4}} dx$$

Here, we see that the integral is over a very large range (0 to ∞), compared to the other integrals that we have evaluated. As a result, we do not expect to obtain as accurate results for evaluating the integral as we did in parts a and b. Thus, here we estimate the integral 100 times using 1,000,000 samples and the Monte Carlo method, and plot the density to see the value that the estimates center themselves around. For this integral, we let $f(x)$ and $h(x)$ be defined as:

$$f(x) = \frac{\frac{1}{4}^5}{\Gamma(5)} \cdot x^4 \exp\left(-\frac{x}{4}\right) \sim \text{Gamma}(5, 0.25)$$

$$h(x) = \frac{3}{4} \cdot \frac{\Gamma(5)}{\frac{1}{4}^5} \cdot \exp\left(\frac{x}{4} - \frac{x^3}{4}\right)$$

Thus, we take 1,000,000 samples of x from $x \sim \Gamma(5, 0.25)$ and plug these into $h(x)$ to be used in the formula for estimating the integral. Repeating this process 1000 times, we obtain the following density plot of evaluated integral estimates. The vertical red line represents the maximum of the density, which occurs at the value 2.25775.

Monte Carlo Method Integral Estimates

N = 1000   Bandwidth = 0.04357
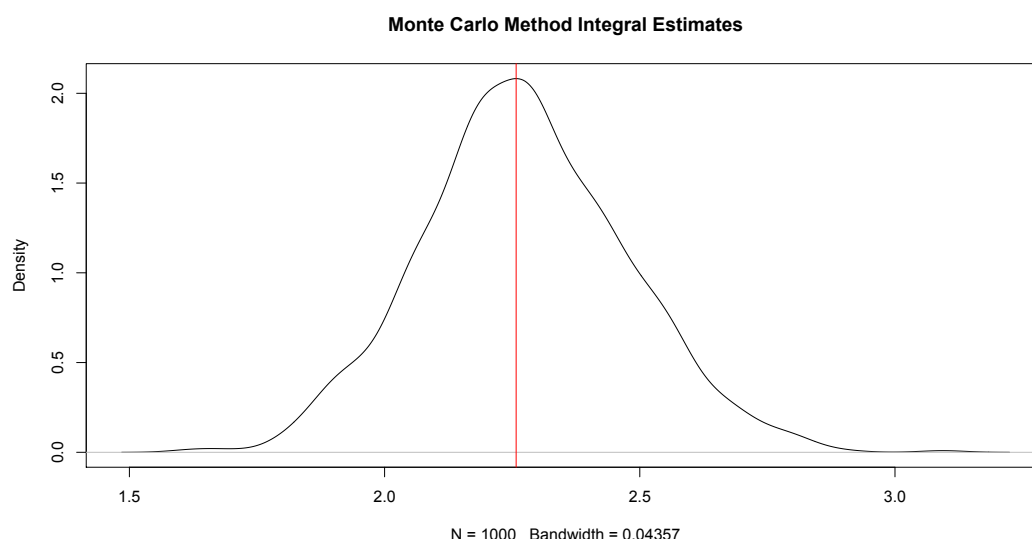
2. We let

$$I = \frac{1}{\sqrt{2\pi}} \int_1^2 e^{-\frac{x^3}{2}} dx$$

and estimate I using importance sampling. We take g to be $N(1.5, \nu^2)$ with $\nu = 0.1$ and 10.

3. We will approximate the following integral using both the simple Monte Carlo integration and the control variate method:

$$I = \int_0^1 \frac{1}{1+x} dx$$

a. We let $h(x) = \frac{1}{1+x}$ and $U_1, \dots U_n$ be iid Unif[0,1].

    b.

    c.

    d.

4. We consider a common application in statistics: three different treatments are to be compared by applying them to a randomly selected experimental units. However; instead of the usualy assumption associated with e_{ij} in the typical analysis of variance methods of comparison, we assume that the e_{ij} have independent and identical double exponential distributions centered on zero.

    a.

    b.

5. We consider the zero-inflated Poisson(ZIP) model, in which random data $X_1, \ldots, X_n$ are assumed to be of the form $X_i = R_i Y_i$, where Y_i's have a Poisson(λ) distribution and the R_i's have a Bernoulli(p) distribution, all independent of each other. Given an outcome $x = (x_1, \ldots, x_n)$, the objective is to estimate both λ and p. We consider the following hierarchical Bayes model:

- $p \sim \text{Uniform}(0,1)$ (prior for p)

- $(\lambda|p) \sim \text{Gamma}(a,b)$ (prior for λ)

- $(r_i|p,\lambda) \sim \text{Bernoulli}(p)$ independently (from the model above)

- $(x_i|r,\lambda,p) \sim \text{Poisson}(\lambda r_i)$ independently (from the model above)

where a and b are known parameters, and $r = (r_1, \ldots, r_n)$. It follows that:

$$f(x,r,\lambda,p) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \prod_{i=1}^{n} \frac{e^{-\lambda r_i}(\lambda r_i)^{x_i}}{x_i!} p^{r_i}(1-p)^{1-r_i}$$

We wish to sample from the posterior pdf $f(\lambda, p, r|x)$ using the Gibbs sampler.

    a.

    b.

    c.

6. The Independence - Metropolis - Hastings Algorithm is an importance-sampling version of MCMC. We draw the proposal from a fixed distribution g. Generally, g is chosen to be an approximation to f. The acceptance probability becomes:

$$r(x, y) = min \left\{ \frac{f(y)}{f(x)} \cdot \frac{g(x)}{g(y)}, 1 \right\}$$

A random variable Z has an inverge Gaussian distribution it it has density

$$f(z) \propto z^{-3/2} \exp \left\{ -\theta_1 z - \frac{\theta_2}{z} + 2\sqrt{\theta_1 \theta_2} + \log\sqrt{2\theta_2} \right\}, \quad z > 0$$

where $\theta_1 0$ and $\theta_2 > 0$ are parameters. It can be shown that

$$E(Z) = \sqrt{\frac{\theta_2}{\theta_1}} \quad \text{and} \quad E\left(\frac{1}{Z}\right) = \sqrt{\frac{\theta_1}{\theta_2}} + \frac{1}{2\theta_2}$$

We let $\theta_1 = 1.5$ and $\theta_2 = 2$, and draw a sample of size 1,000 using the independence-Metropolis-Hastings algroithm. We use a Gamma distribution as the proposal density.