Homework 4 Question 1

Preey Shah and Sahil Shah

November 2017

The code that was used to estimate the values of pi is in p1.m. After we came to a correct value for m, we used the code in forSpecifiedM.m to verify our value

Both X_1, X_2 have uniform distributions, with $f(X1) = 1/2,\ x_1 \in [-1, 1]$, 0 otherwise. Since they are independent,$f(X_1, X_2) = 1/4$. Therefore probability that lies in the circle is $\int\int_{x_1^2+x_2^2<1} \frac{1}{4} dxdy = \frac{Area\ of\ circle}{4} = \frac{\pi}{4}$. This can also be viewed as a ratio of areas. X can be taken uniformly from a sqaure of side length=2, and we have to find that the probability of the point lying in a circle of radius 1 contained in that square. This ratio is given by $\frac{area\ of\ circle}{area\ of\ square} = \frac{\pi}{4}$ Now, to estimate π, we should generate a large number of points in a square of area 4 around the origin and calculate the number of points lying in the circle. This way, equating this to $\frac{\pi}{4}$, π can be estimated.

The observed values of π were observed as

Size of sample data	Estimate of pi
10	2.0000
100	3.0400
1000	3.2360
10000	3.1420
100000	3.1397
1000000	3.1418
10000000	3.1417
100000000	3.1418

In the case when $N = 10^9$, it takes around 1 minute to calculate the value. However, had we created an array and then typecast it into single, it would run out of memory(as it went out of memory initially). However, to circumvent this, instead of making an array of random numbers in the beginning, we generated random numbers inside the loop, which was taking a lot of time. So, we finally implemented the given version of code.

Part (d)

Consider a Bernoulli random variable that has a value 1 if it lies inside the circle of radius 1 and 0 if it lies outside it. The value of the parameter p for this Bernoulli random variable is $\pi/4$. The variance of each of this variable is $p(1-p) = \frac{\pi}{4}(1-\frac{\pi}{4})$, and the expected value is $E(X_i) = \frac{\pi}{4}$ We denote these variables by X_i

We will now apply the central limit theorem (as we expect the number of trials required to be fairly large) and hence approximate the distribution of the sum of many such variables to be a Gaussian with mean equal to the expected value. According to the Central Limit Theorem,

$$\sqrt{n}(\frac{\sum_{i=0}^{n} X_i - n\frac{\pi}{4}}{n}) \sim Gaussian(0, \frac{\pi}{4}(1-\frac{\pi}{4}))$$

$$(\frac{\sum_{i=0}^{n} X_i - n\frac{\pi}{4}}{\sqrt{n\frac{\pi}{4}(1-\frac{\pi}{4})}}) \sim Gaussian(0,1)$$

We now need that $\frac{\sum_{i=1}^{n} X_i}{n}$ should lie between $\frac{\pi-0.01}{4}-$ and $\frac{\pi+0.01}{4}-$ or

$$n(\frac{\pi}{4} - 0.0025) \leq \sum_{i=0}^{n} X_i \leq n(\frac{\pi}{4} + 0.0025)$$

Hence, we need:

$$\sqrt{n}\frac{-0.0025}{\sqrt{n\frac{\pi}{4}(1-\frac{\pi}{4})}} \leq (\frac{\sum_{i=0}^{n} X_i - n\frac{\pi}{4}}{\sqrt{\frac{\pi}{4}(1-\frac{\pi}{4})}}) \leq \sqrt{n}\frac{0.0025}{\sqrt{\frac{\pi}{4}(1-\frac{\pi}{4})}}$$

$$\sqrt{n}\frac{-0.0025}{\sqrt{\frac{\pi}{4}(1-\frac{\pi}{4})}} \leq Gaussian(0,1) \leq \sqrt{n}\frac{0.0025}{\sqrt{\frac{\pi}{4}(1-\frac{\pi}{4})}}$$

We now need α, such that $-\alpha \leq Gaussian(0,1) \leq \alpha$ with probability 0.95. The needed value is 1.96 (as found on wikipedia). Therefore we have:

$$\sqrt{n}\frac{0.0025}{\sqrt{\frac{\pi}{4}(1-\frac{\pi}{4})}} = 1.96$$

$$n = \frac{\pi}{4}(1-\frac{\pi}{4})(\frac{1.96}{0.0025})^2$$

Therefore, we get

$$m = 103599$$

We ran the code with the computed value of m and found that the estimate of π was 3.1401, which is within 0.01 of the actual value.