CS302: Modeling And Simulation
Lab 5 Report

Rajdeep Pinge 201401103, Aditya Joglekar 201401086

March 2, 2017

1 Q1: Implementing Monte carlo method

In the given problems in part 1, we need to use the Monte Carlo techique to estimate the quantities asked. Note that these are toy examples which we can easily solve by hand. The goal however, is to understand the basics of the method and quickly verify if the estimation converges to the actual answer or not.

The Monte Carlo method roughly includes the following steps:

- For geometrical applications, we start by identifying the random variables and their bounds.
- The required quantity is expressible as the ratio of the count of possible scenarios involving the random variables.
- We make use of the Law of Large Numbers and the ability to generate uniformly distributed random numbers.
- Note that in practice, we really have access to pseudo-random numbers. But, using standard implementations, these are good enough approximations.

1.1 Area between the curve for $f(x) = x^2$ and the x-axis from x = 0 to x = 2.

Here, we need to calculate the area under $f(x) = x^2$, from x=0 to x=2.

To put the problem in the general MC(Monte Carlo) framework, we need to define geometric bounds. Here, that is easy enough. We know that that x will go from x=0 to x=2. Thus, y must lie between y=0 and y=4. Thus, we have a "box like frame" around our curve of interest.

We will now think of the above as a dart board and imagine that we are throwing darts at it. The dart will hit a point somewhere on the board. We keep track of where these darts land.

Essentially, we will pick the x and y coordinates from two uniform distributions, one ranging from [0,2] and the other [0,4]. The point P(x,y) will be a dart mark on our imaginary board.

Since, the numbers are uniformly distributed, the dart is equally likely to fall anywhere on the board. We will divide the board into two regions. One under the curve and one above it.

Because of uniformity we can show that the ratio of the points in the two regions is approximately equal to the ratio of the areas provided we take a large number of points.

Thus, the area under the curve can be estimated by

$$\frac{Area\ of\ the\ rectangle}{Area\ under\ the\ curve} = \frac{8}{Area\ under\ the\ curve} = \frac{n_t}{n_i} \tag{1}$$

$$Area\ under\ the\ curve = \frac{8n_i}{n_t} \tag{2}$$

where n_i and n_t are the points lying under the curve and the total number of points considered respectively.

There are two approaches to use the MC approximation.

1.1.1 Long term behaviour of system for a large number of iterations

The first one is the one we have been discussing. Here, an important consideration is the number of points taken. By the law of large numbers, the larger the number, the better is our approximation. In the limiting case of infinite points the answer would converge to the actual analytical answer. We thus, want to examine how does the estimate approach the theoretical answer.

To do this we will run the experiment with different number of points taken. We then plot the answer obtained.

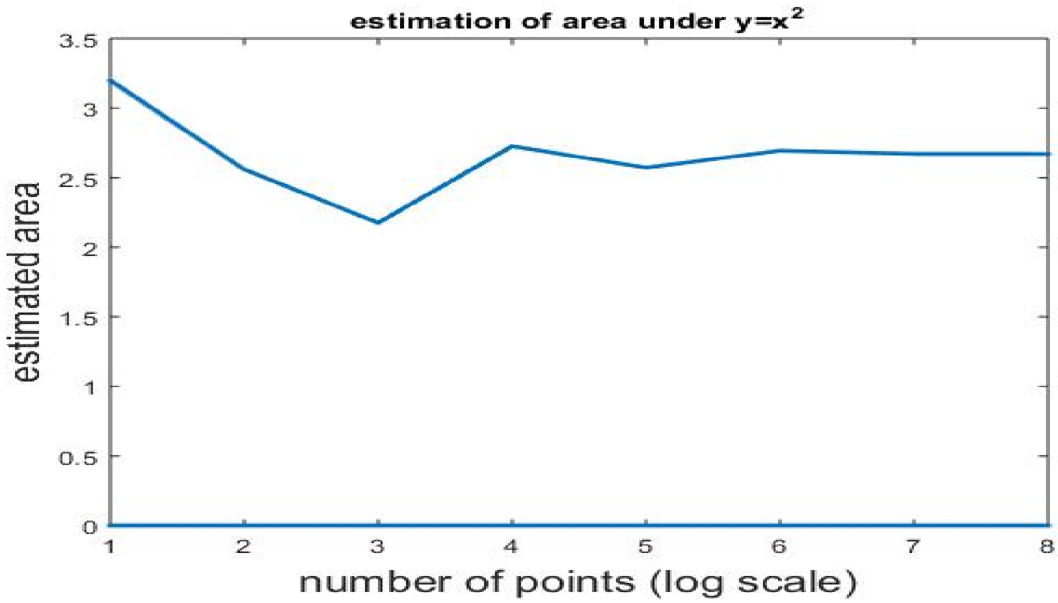

Figure 1: Convergence of the estimate to the analytical answer

- We know that the analytical answer is 8/3 or around 2.667.
- We notice that for lesser number of points considered, the estimate is way off the mark.
- But as we increase the number of points taken, we notice that the fluctuations go down and that it converges to the theoretical answer.
- This is what we expect from law of large numbers. When, we take a large number of samples, the ratio of the number of points in a region to the total number of points starts to approach the probability of the point lying in that region. Moreover, probability of the point lying in a region is proportional to its area due to the uniform assumption. Hence, we get better and better results for larger number of observation s.
- Intuitively, if we are dealing with things whose outcomes are known but are random in nature, if we observe them a large number of times, our estimate about their behaviour improves, we can predict how they will behave on "an average". The large number of observations cover all possibilities giving us a more complete picture.

1.1.2 Short term behaviour of a system for a large number of experiments: Ensemble Average

There is another approach to this problem called "ensemble averaging". Here, instead of taking a large number of points to estimate, we take a small number of points but repeat the experiment a large number of times. The average of the estimates at all the iterations is expected to converge to the analytical answer.

The intuition behind the idea is that each time we conduct the experiment, we get an estimate with the points chosen. But, thse points are from the sample space which the 2-D region being considered. Thus, if we do the experiment a large number of times, we expect all possible variations will be captured and thus the average of these cases will be close to the theoretical solution.

It is similar to the situation in which seeing bits and pieces of a dark room with a matchstick a large number of times gives us a general idea of our surroundings.

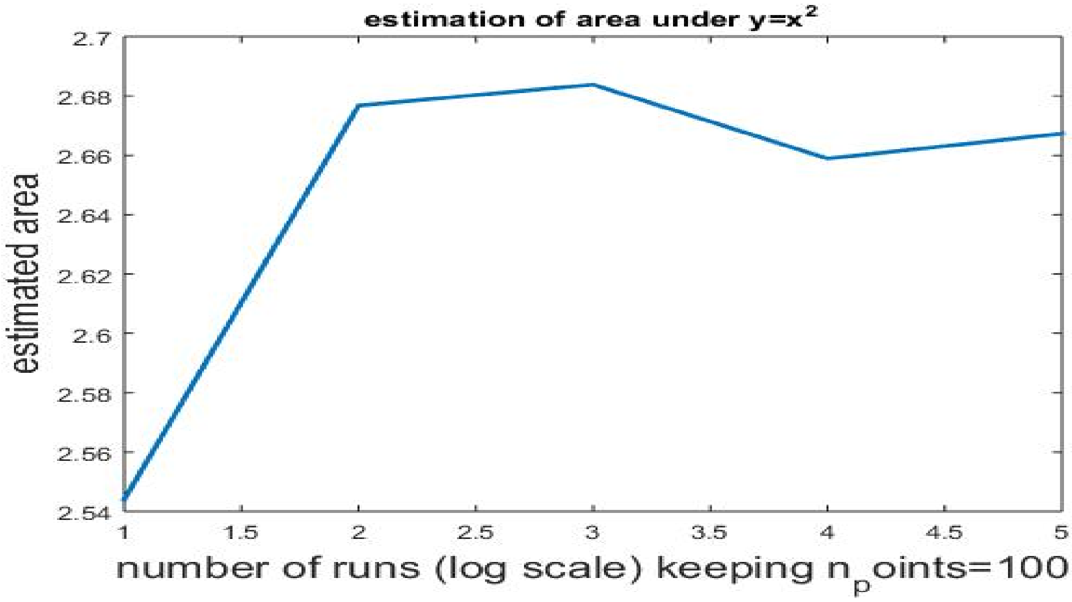

Figure 2: Convergence of the ensemble estimate to the analytical answer

Clearly, for large number of runs, we get a good estimate of the actual answer.

1.2 An estimate of π.

To estimate the area of π, we make use of the standard unit circle (radius = 1 unit) and a square circumscribing it (length of side = 2 units).

Once again, we use similar arguments as in the previous section and connect the areas to the probabilities of the points lying in specific regions. Further, we use Law of Large Numbers to say that the ratio of the probabilities will be approximated by the ratio of the number of points drawn from a uniform distribution. Basically, similar arguments as above.

Thus, the value of π can be estimated by

$$\frac{Area\ of\ the\ circle}{Area\ of\ the\ square} = \frac{\pi * r^2}{2r * 2r} = \frac{\pi * 1^2}{2 * 2} = \frac{n_i}{n_t} \tag{3}$$

$$\pi = \frac{4n_i}{n_t} \tag{4}$$

where n_i and n_t are the points lying under the curve and the total number of points considered respectively.

Again, we get two plots to check the converge of the two approaches.

1.2.1 Long term behaviour of system for a large number of iterations

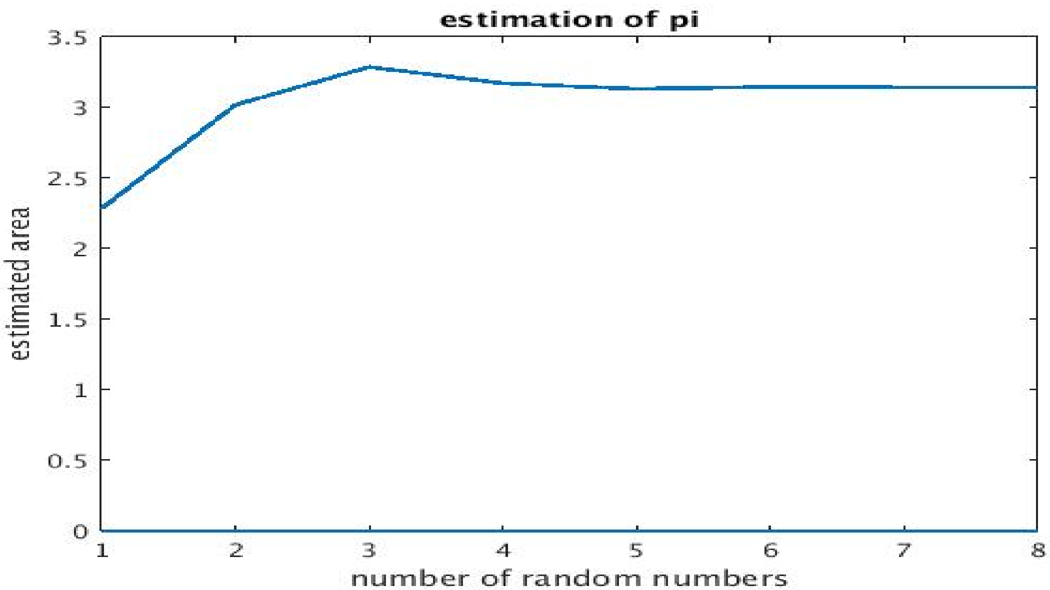

Figure 3: Convergence of the estimate to the analytical answer

1.2.2 Short term behaviour of a system for a large number of experiments: Ensemble Average

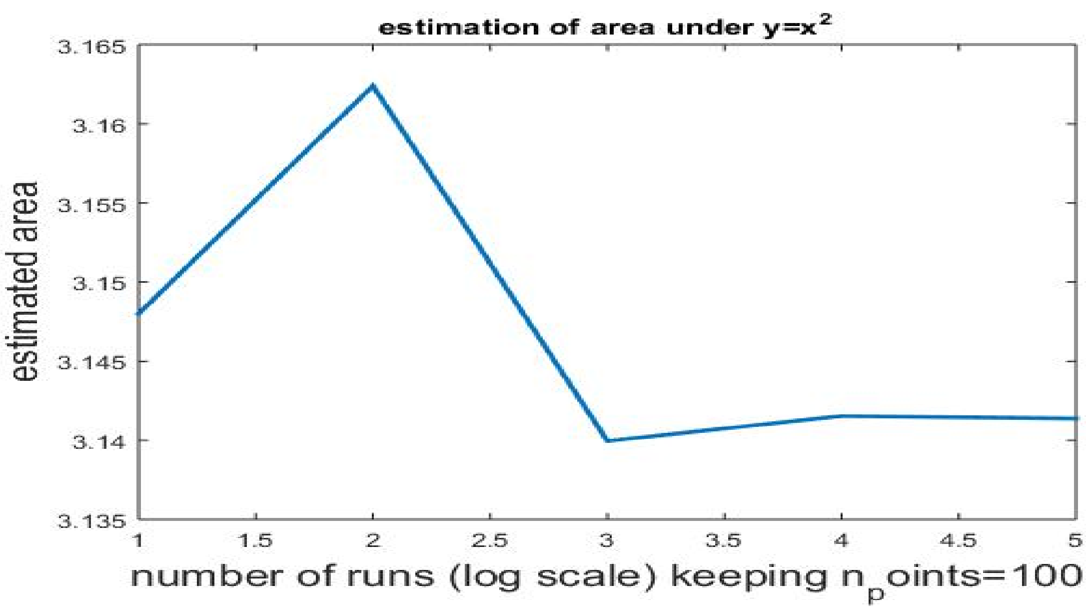

Figure 4: Convergence of the ensemble estimate to the analytical answer

1.3 Volume of a Sphere.

To estimate the volume, we make use of the standard unit sphere (radius = 1 unit) and a cube encompassing it (length of side = 2 units).

Once again, we use similar arguments as in the previous section and connect the Volumes to the probabilities of the points lying in specific regions. Further, we use Law of Large Numbers to say that the ratio of the probabilities will be approximated by the ratio of the number of points drawn from a uniform distribution. Basically, similar arguments as above with the only difference being that we have a third dimension and thus 1 more random variable.

$$\frac{Volume\ of\ the\ sphere}{Volume\ of\ the\ cube} = \frac{Volume\ of\ the\ sphere}{2*2*2} = \frac{n_i}{n_t} \tag{5}$$

$$Volume\ of\ the\ sphere = \frac{8n_i}{n_t} \tag{6}$$

Again, we get two plots to check the converge of the two approaches.

1.3.1 Long term behaviour of system for a large number of iterations

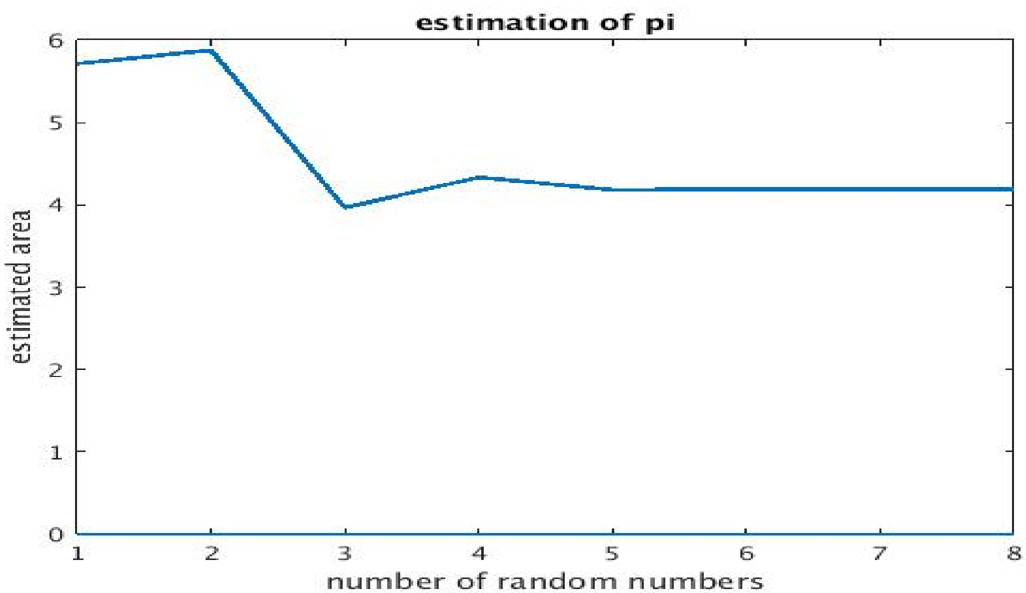

Figure 5: Convergence of the estimate to the analytical answer

1.3.2 Short term behaviour of a system for a large number of experiments: Ensemble Average

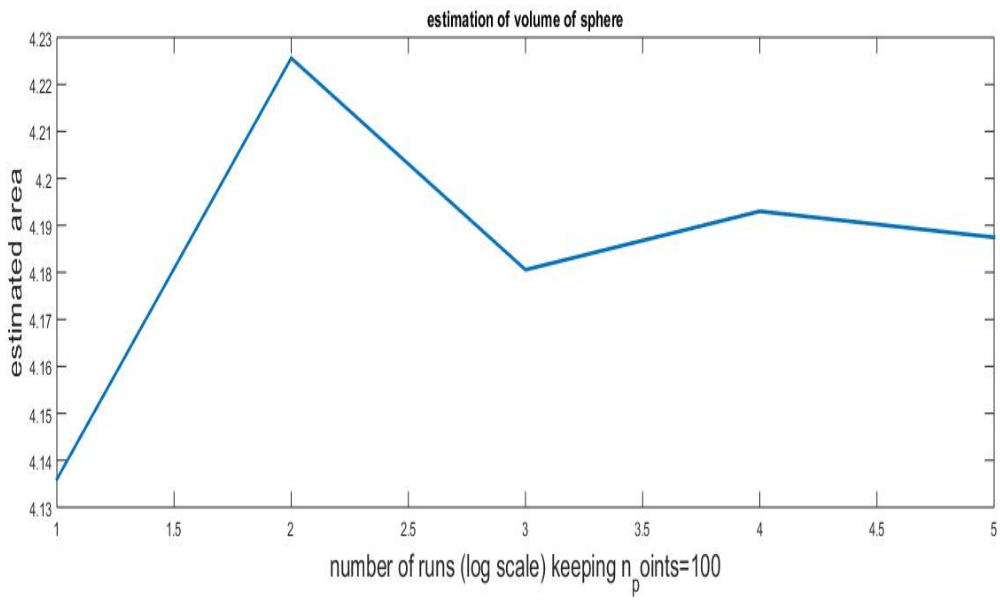

Figure 6: Convergence of the ensemble estimate to the analytical answer

2 Q2: Starting from uniformly distributed random numbers between 0 and 1, generate random numbers and plot a histogram for PDF and CDF for the following distributions

2.1 Normal (N(μ,σ^2)) (Use Box-Muller Algorithm)

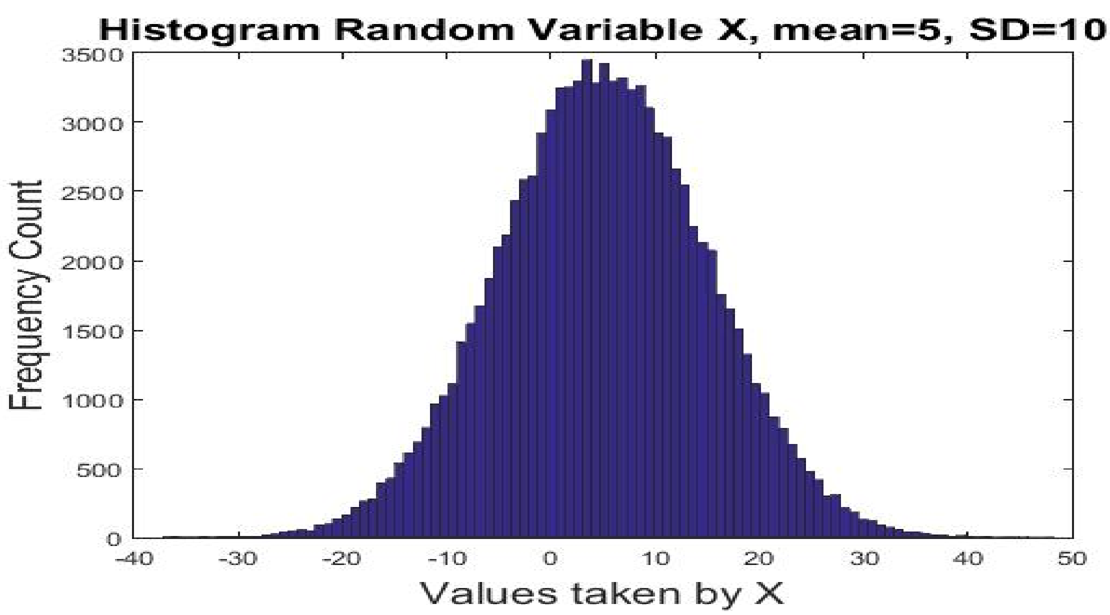

Figure 7:

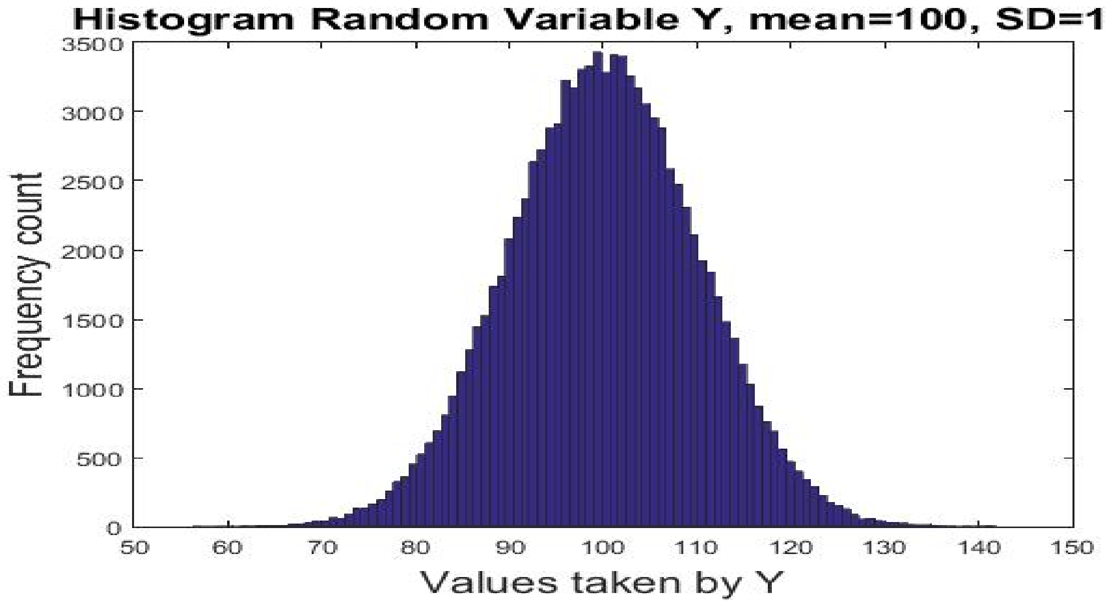

Figure 8:

2.2 Exponential $\mathbf{F}_X(x) = \mathbf{1}\ \mathrm{e}^{-(x*\lambda)}$ (Use Inverse Transform Method)

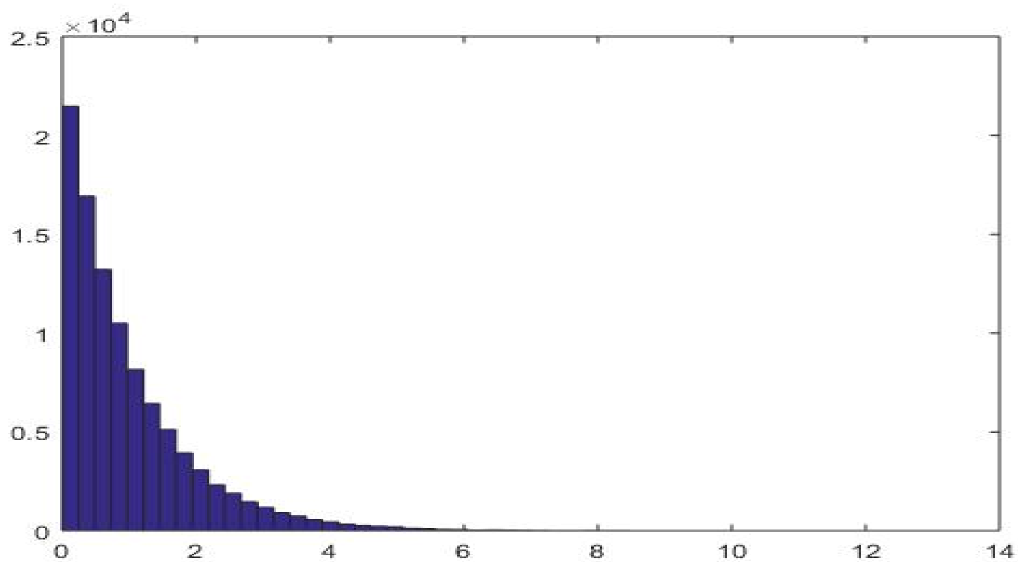

Figure 9: PDF for given distribution

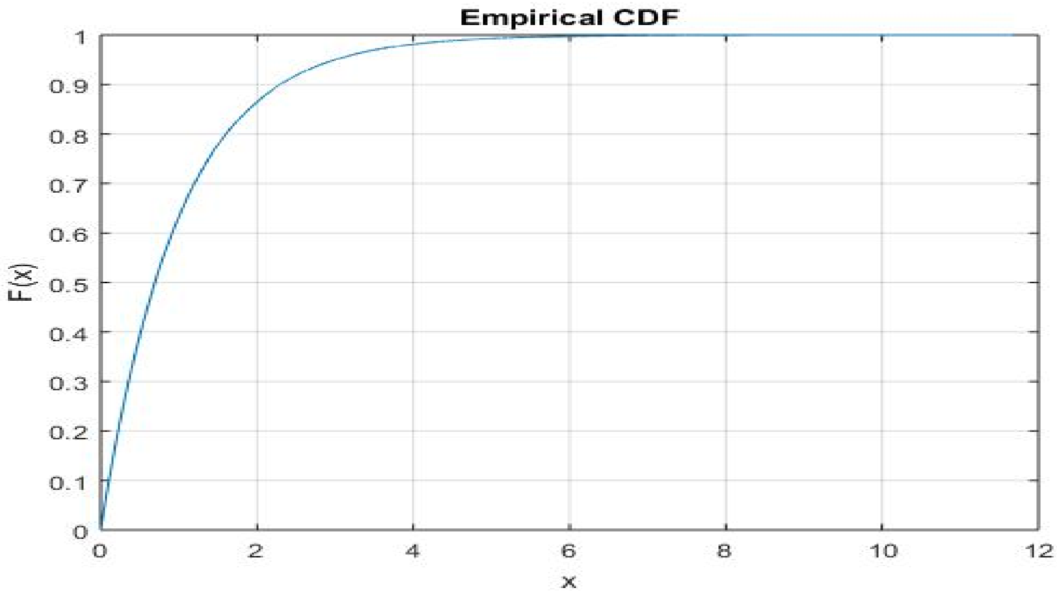

Figure 10: CDF for given distribution

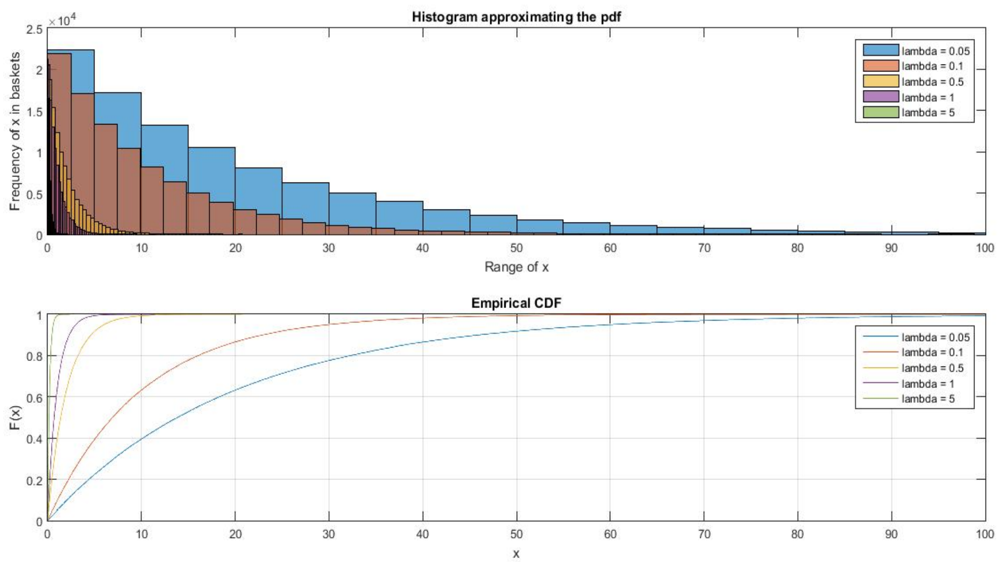

Figure 11: graphs for varying lambda

2.3 Weibull $F_X(x) = 1\ e^{-(x/\lambda)k}$, for x > 0. (Use Inverse Transform Method)

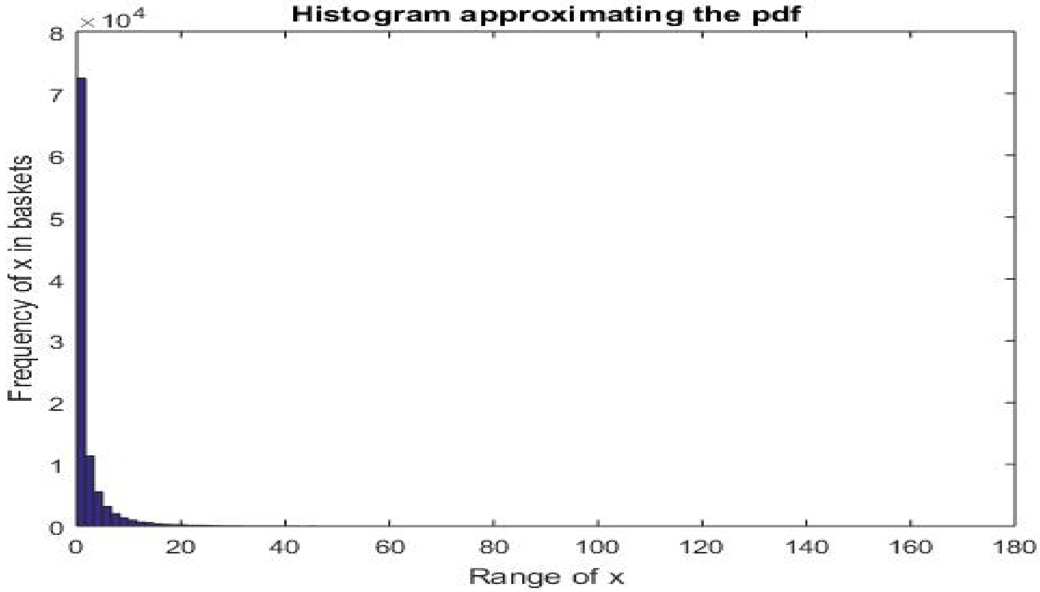

Figure 12:

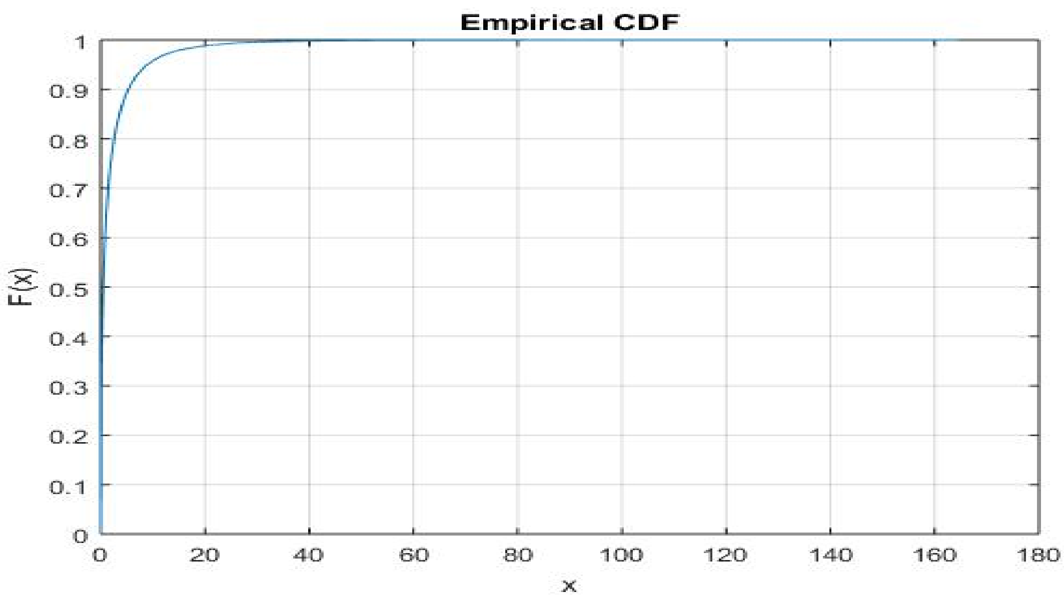

Figure 13:

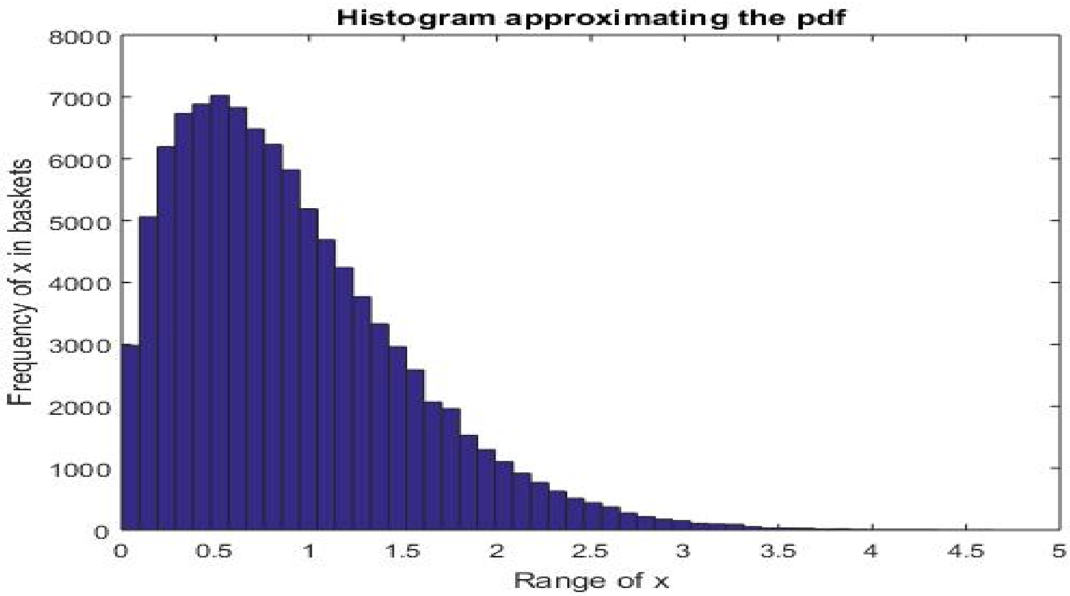

Figure 14:

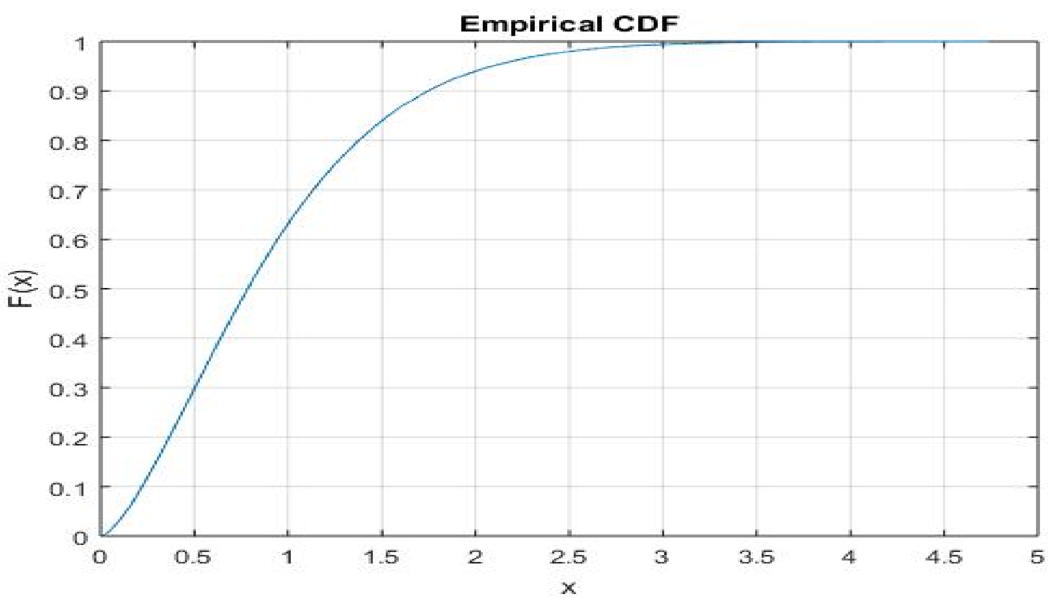

Figure 15:

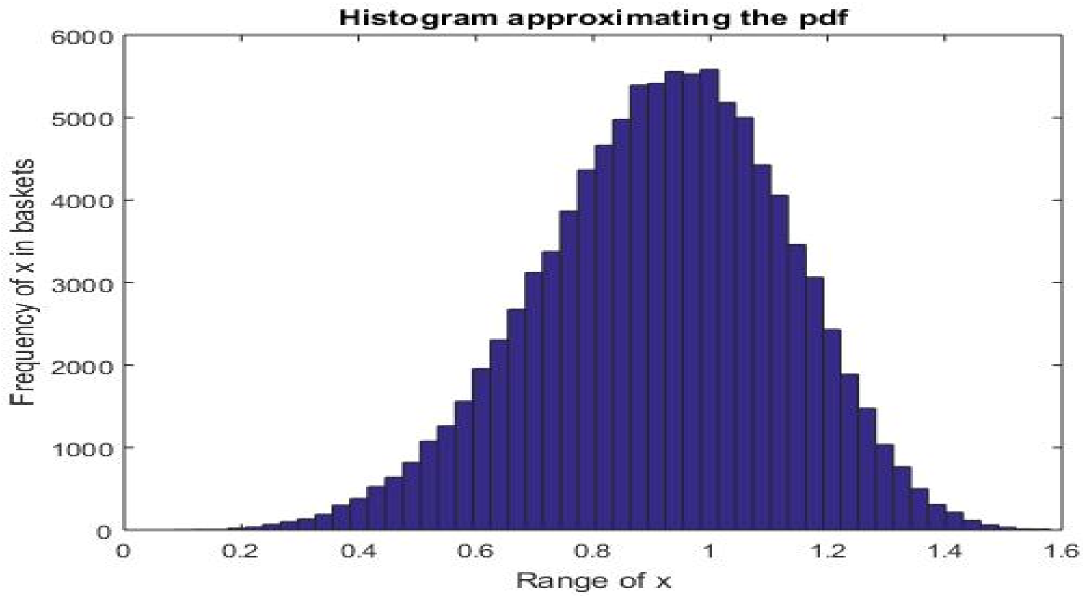

Figure 16:

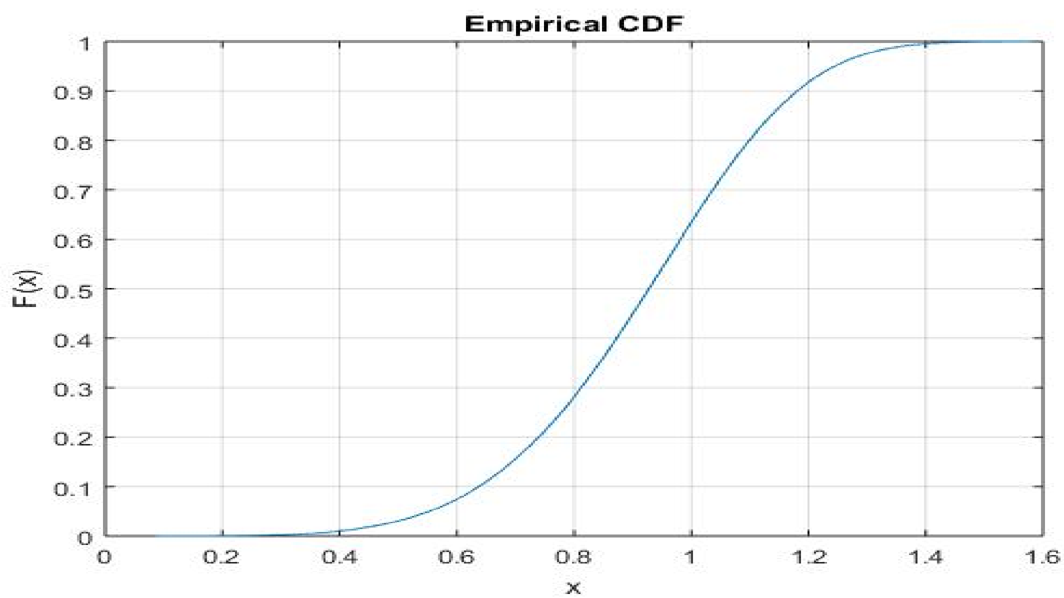

Figure 17:

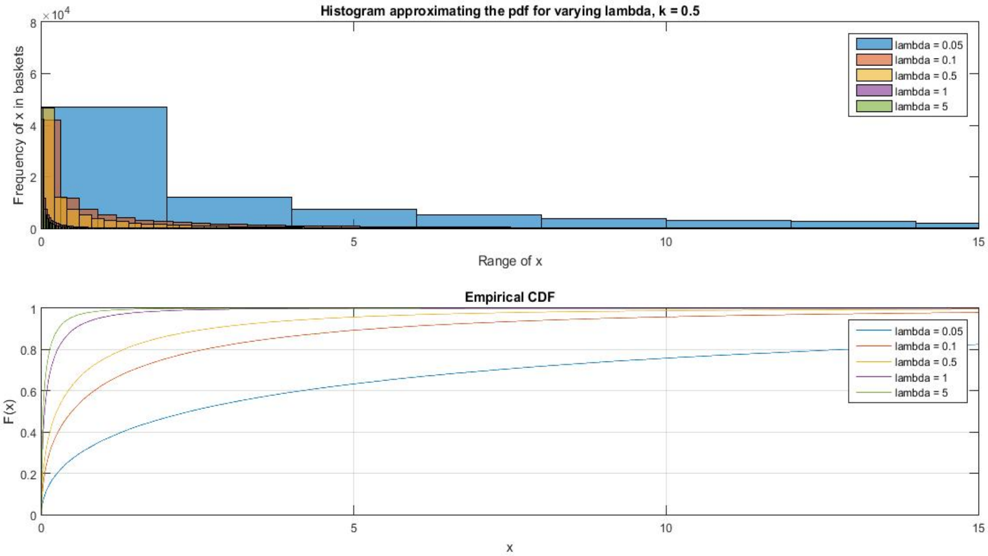

Figure 18: graphs for varying lambda

The pdfs have been generated using Inverse Transform method. We use a pdf known to generate the new pdf by using the inverese relation between them. These pdfs are parameter dependent, hence it is clearly seen that with different parameters, the pdfs change according to what we expect mathematically.

3 Q3: Using Acceptance-Rejection Method obtain random numbers whose probability density function is given by f(x) = 2π $\sin(4\pi x)$ in the range 0 to 0.25. Generating sufficient number of values plot the histogram.

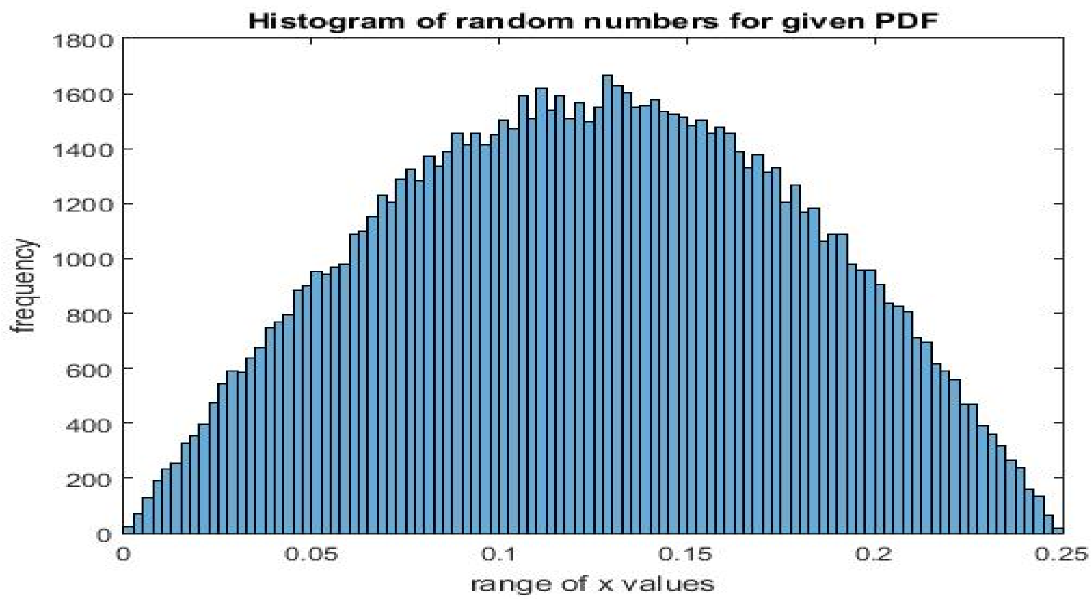

Figure 19: Histogram of random numbers for given PDF

For large enough observations, the above technique gives us a pretty good approximation of the given pdf. The acceptance rejection basically generates the required pdf by generating a known pdf function and based on the relative ratio between the pdfs, either accepting or rejecting the generated number.