Assignment 7 DS303 on Simulation

Tarun Singla B19198

Question1

Write a python program to estimate the value of π using Monte Carlo simulation (raindrop experiment). Generate animation of the simulation where estimates of π converges with an increase in the number of drops/samples. Plot the Monte Carlo estimate of π with 90% confidence interval where x-axis represent number of sample (1 to 3000) and y-axis represent estimate of π.

STEPS:

1. Importing the Libraries matplotlib (plotting), animation(for animation) and numpy (for maths)
2. I am calling AnimationFunc and calling different functions inside it.
3. We have gen_data() function it will basically generate values x and d y and for every value of x and y we will check whether this lies inside the circle we will update the count and the we will calculate estimated value of pi using monto carlo method
   Let number of points inside the function be x
   Let total number of points generated be y
   Area of circle be a1 = pi
   Area of square be a2 which is equal to 4 (side take is 2)
   x/y = a1/4
   a1 = 4*(x/y) = estimated value of pi.
4. Last part is to calculate the confidence interval whose formula is p^ +- z*sigma/root(n) where sigma is root(p^(1-p^)) p^ = number of points inside circle
5. After that plotted the graph of estimated of pi along with the confidence interval.

Question2

$$\int_0^1 f(x)dx \quad \text{with} \quad f(x) = \frac{1}{27}\left(-65536x^8 + 262144x^7 - 409600x^6 + 311296x^5 - 114688x^4 + 16384x^3\right)$$

using a Monte Carlo approach. Graphically show that the graph of f(x) is fully contained in unit square [0, 1]2 . Generate animation of the simulation where estimates of the integral converges with an increase in the number of drops/samples. Plot the Monte Carlo estimate of integral with 90% confidence interval where x-axis represent number of sample (1 to 2000) and y-axis represent estimate of R integral of f(x)

STEPS:

1. Importing the Libraries matplotlib (plotting), animation(for animation) and numpy (for maths)
2. I am calling AnimationFunc and calling different functions inside it.
3. We have gen_data() function it will basically generate values x and d y and for every value of x and y we will check whether this lies inside the function we will update the count and the we will calculate estimated value of integral using Monto Carlo method
   Let number of points inside the function be x
   Let total number of points generated be y
   Area of f(x) = a1
   Area of square be a2 which is equal to 1 (rectangle of x from 0 to 1 and y from 0 to 1)

x/y = a1
a1 = 1*(x/y) = estimated value of a1 (integral).

4. Last part is to calculate the confidence interval whose formula is p^ +- z*sigma/root(n) where sigma is root(p^(1-p^)) p^ = number of points inside circle

After that plotted the graph of estimated of pi along with the confidence interval

Question3

Sample points from N(0, 1) distribution (target distribution) with destiny using rejection sampling method by taking Cauchy distribution with density as a proposal distribution. Find the smallest value of M such that f(x) ≤ Mg(x). Generate animation of rejection sampling containing graphs of target distribution, proposal distribution, scale proposal distribution and histogram of selected random samples. The histogram of the selected random samples should keep on updating with number of iterations/ samples

F(X) = standard normal distribution **G(X) = Cauchy distribution**

STEPS

1. First Step is to find the value of M which will be sqrt(2*pi/e) on solving it will become 1.5 approx.
2. Plotted graph of f(x) and m*g(x)
3. We will be using definition of rejection sampling
4. Now we apply functions in AnimationFunc which will generate the samples of cauchy distribution and standard normal distribution after that we are using check the condition
   Let samples generated from standard normal distribution be x
   Let samples generated from Cauchy distribution be y
   We will check x < (f(y)/M*g(y) we will accept the samples otherwise we will reject the samples
5. After that we are plotting the histogram of it.

Question4

P(x) is a density function of target distribution and N(0, 4) is density function of proposal distribution. Write a python program to estimate E[X], E[X2] and E[X5] where X ～p using importance sampling technique. Report the values of estimates for N = 100, 500, and1000, also report the error in each estimate

P(X) = 1/2*e-|x|

STEPS

1. First defined function f(x) is the proposal distribution which is equal to p(x), g(x) which is normal distribution and h(x) which will be calculating exponent of X to find the value of E[X^p] for different values of p.
2. For different values of N we will be calculate using the formula of of importance sampling by creating the random samples from normal distribution.
3. For calculating the actual values we know when E[X^j] when j is odd is an odd function so its integration is will be zero now for calculating actual value when j = 2 I calculated using pen and paper which came out to be 2.
4. After that I have calculated Error which abs(actual – estimated)

Question5

Let f(X) = 1 40 × (2x + 3), 0 < x < 5 be a density function. Sample 1000 random draws from this density using the inverse transform sampling method. Plot the graph of given density function and histogram plot of the drawn samples in a single figure

STEPS

1. Defined the function f(x) and inverse CDF f(x) for inverse transform sampling
2. Generated the random uniform samples and calculated the values of inverse function using this which gave random samples with distribution f(x)
3. Plotted the histogram of it gave use the distribution

Question6

Simulate a 2D random walk inside given upper and lower bound of x and y-axis for given number of steps. Write a python code to generate animation of the simulated 2D random walk

STEPS

1. Defined the random walk function to generate 2d points of randomly with increments and decrements of either +1 or -1 after setting the minimum value and maximum value of the coordinates. For better result Normalized the random walk
2. After that Plotted and animated generated the random walk

Question7

Generate 100000 random samples from bivariate normal distribution and draw scatter plot using following covariance matrices

(a) cov = c ∗ I, where c ∈ R + and I is 2 × 2 identity matrix.

(b) cov = diag(a1, a2) where a1 and a2 ∈ R +.

(c) cov = A2×2 where A is 2 × 2 symmetric matrix with positive diagonal entries. Repeat the experiment for fixed value of diagonal entries and by taking different value for off diagonal entries like ±0.2, ±0.5, ±0.8, ±1, ±2, ±3, ±5, ±10 etc.

Write the insights you get from all these scatter plot experiments. Is this insight is also applicable to other distributions

HINTS

1. Made function to check whether matrix is positive definite.
2. For part(a) generating random matrix c*I where I is identity matrix. Update the value till matrix is not positive definite.
3. Plotted bivariate normal distribution.
4. For part(b) generating diagonal matrix and till matrix is not positive definite and after that plotted the scatter plot
5. For part(c) generating symmetric matrix and after that we are generating the multivariate normal distribution and then plotting the scatter plot.

References

1. https://isquared.digital/blog/2020-04-12-random-walk/

2. http://www.acme.byu.edu/wp-content/uploads/2016/12/Vol1B-MonteCarlo2-2017.pdf
3. https://matplotlib.org/stable/api/animation_api.html
4. http://www.ime.unicamp.br/~dias/chap4.pdf