Lab Report

Lab 1

1. Sourav Ganguly, 2. Gudepu Venkatehwarlu, 3. Jyoti

19/01/2022

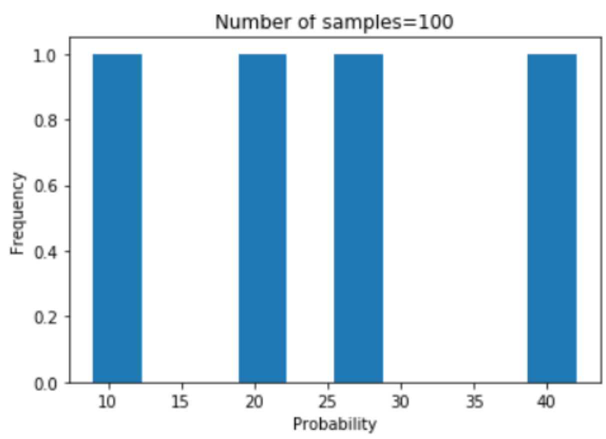
Reinforcement Learning(RL)

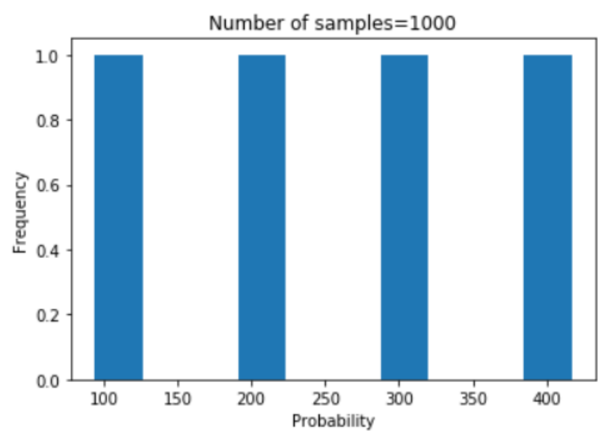
Solutions of the given problems

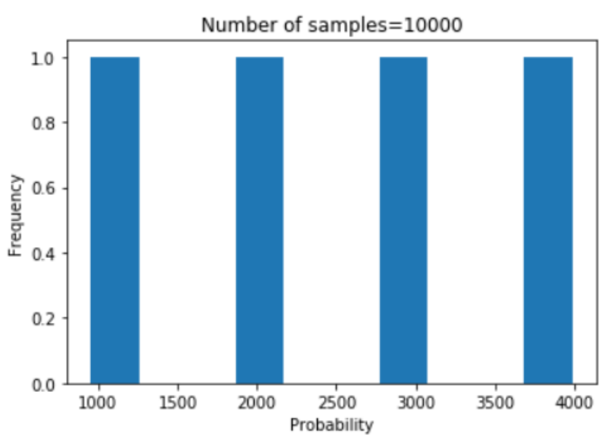
1. Sample n=100, 1000, 10000 points from (a) Multinomial distribution with four outcomes say 1, 2, 3, 4 with corresponding probabilities [0.2 0.4 0.3 0.1] (b) Uniform Distribution in 0 to 1. (c) Gaussian Distribution with mean 0 and variance 1. (d) Exponential Distribution with rate parameter = 0.5

   Solution: We create a list containing all the three values n=[100,1000,10000]. Post that, we use the inbuilt function in python to generate the distributions by sampling from the appropriate functions.

   (a) Since we are given the probabilities of each of the 4 points as [0.2,0.4,0.3,0.1], we pass it in the inbuilt function to get the below curves. As shown here, we have



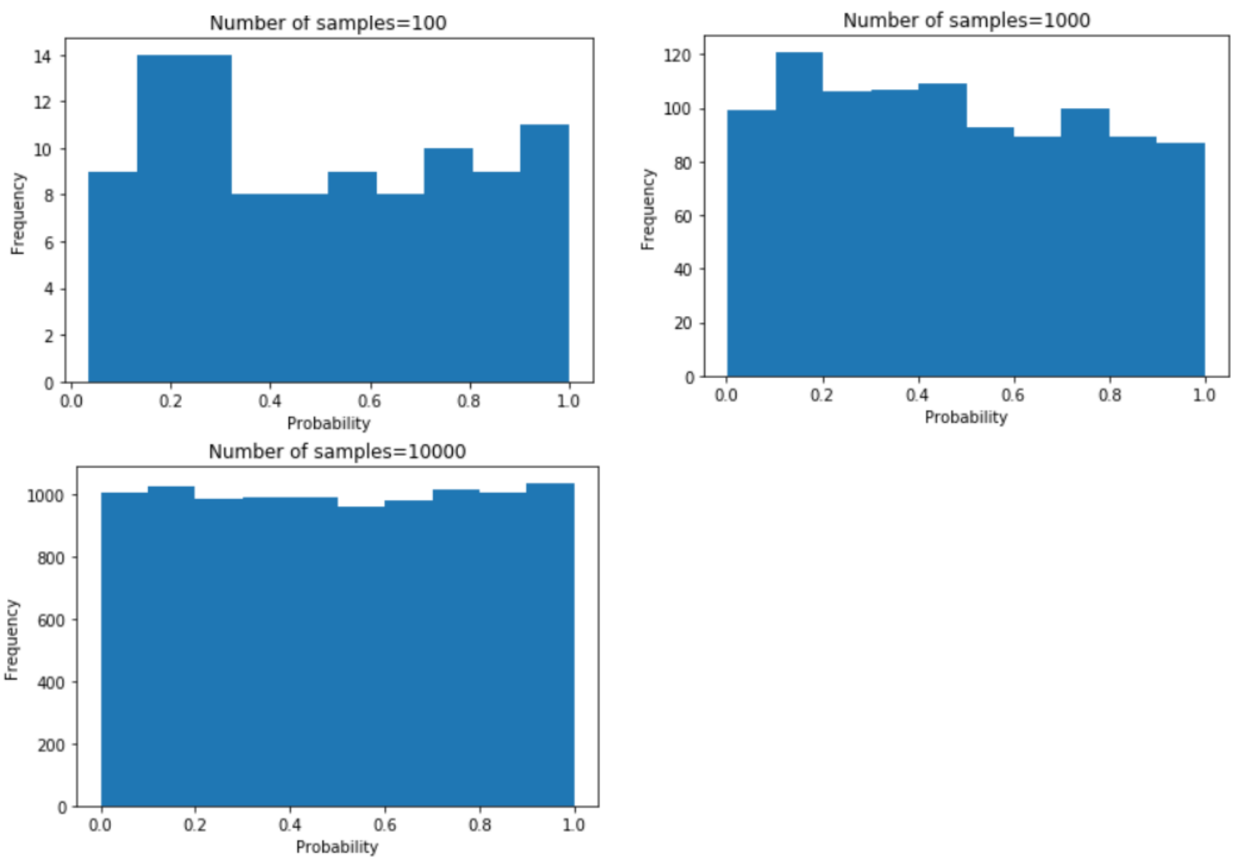




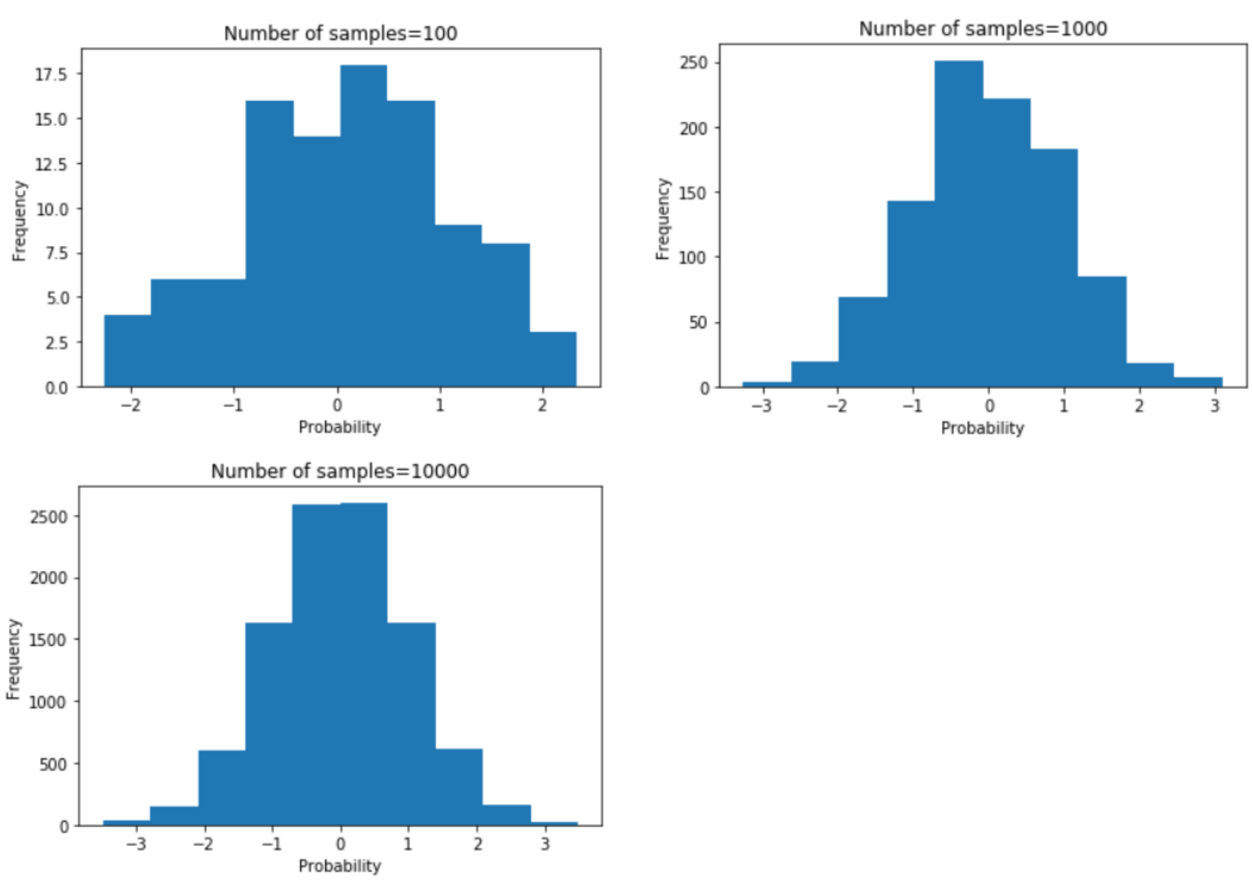
   4 objects were given so, we generate the multinomial distribution changing the number of samples only and then plot a histogram with default bins as shown in the figures.

   (b) Next we need to repeat the same experiment but this time the distribution is changed. Now we have to do for the distribution named as Uniform distribution. To define the uniform distribution, the requiremnt of the two extremes is of vital importance. As given in the question, we define our two extremes as 0 and 1. The area of the distribution is 1. So the points must lie in between 0 and 1 points on the x-axis and not go beyond it. The distribution is repeated with number of samples as 100,1000 and 10000 respectively. The plots are displayed below.

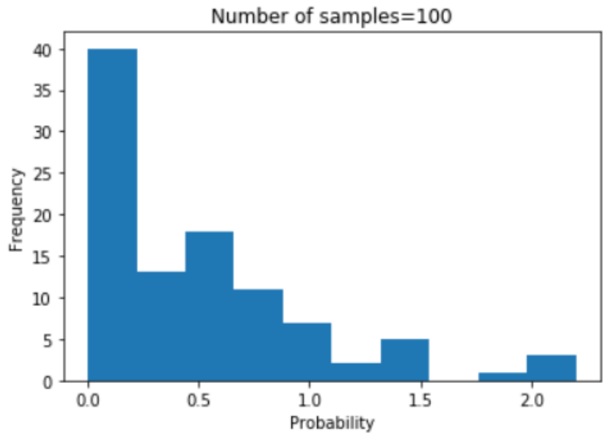
   (c) Next our question asks us to generate a Normal or Gaussian distribution with number of samples as 100, 1000 and 10000. Again we use the inbuilt numpy.random.normal() to generate the required plots. Normal distribution is defined by the mean and the variance of the distribution. As provided in te question, we give to the function the mean and standard deviation which is essentially the square root of variance. The alias used in the inbuilt function for mean and standard deviation is location

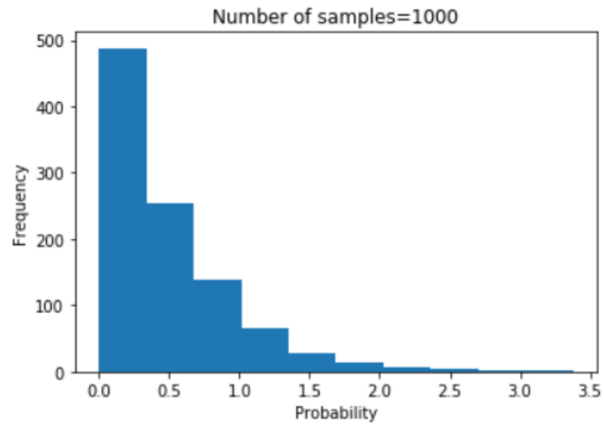


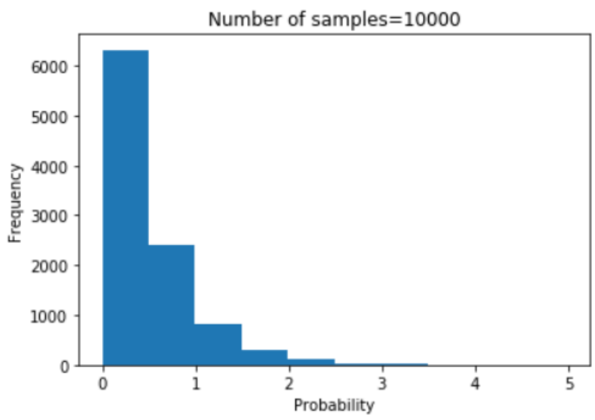
(loc) and scale. We the pass number of samples to be taken and then we get our sampled points which on plotting with the help of histograms give the following structure.



(d) The last subpart dictates us to create an exponential distribution. Now for exponential distribution the most important factor is the scaling factor(rate parameter) which is essentially the inverse of mean of the distribution. Now as the rate parameter was already given, we use it to generate the respective plots as shown below.

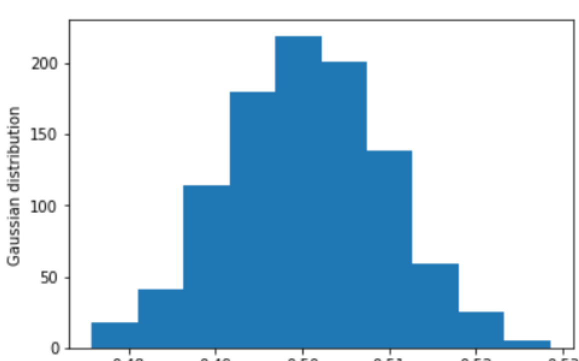






2. Generate normal random variable samples with mean µ and variance σ from uniform random variable. Here, assume you get uniform samples from the interval $(0, 1)$ from inbuilt library functions. Verify if your method indeed generates normal random samples. Justify your procedure clearly.
   Solution: We solve this question in two different ways.

   (a) For this question we use the **Central Limit Theorem** to generate the given Normal distribution from uniform distribution. Central Limit Theorem is a very powerful theorem. Using this theorem we can actually convert any particular distribution into Normal distribution. Now based on this theorem, it states that if we take large number of samples from any distribution and take an average on those samples, it will finally converge to give the Normal distribution. Hence, following , we generated 1000 samples of 1000 samples. The statement might sound confusing. But what we did is that we took 1000 samples between 0 and 1 uniformly and took the mean of the samples. Let this value be μ_1. Now we repeat the same step and found mean from 1000 uniformly generated samples for 999 other occurrences and name them from $\mu_2 ... \mu_{1000}$.
   Finally used these means namely from $\mu_1 ... \mu_{1000}$ to generate our required histogram as shown below.



(b) Again we solve the same question using another technique called the **Box-Muller** technique. In this technique we consider a point in the space and then find its Angular coordinates. By angular coordinate I mean we transform from $(x, y) \implies (s, \theta)$ system. So, from this we know $x = s\cos(\theta)$ *and* $y = s.\sin(\theta)$. Let us suppose (x,y) is a point on the normal distribution with mean 0 and standard deviation 1. $f_r(R) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{r^2}{2})}$. Ok now $r^2 = x^2 + y^2$ by Pythagoras theorem. So, we use it to get the following $f_{x,y}(X,Y) = f_x(X).f_y(Y)$. Since, x and y are independent Finally we, get $f_{x,y}(X,Y) = f_x(X).f_y(Y) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{x^2}{2})}.\frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{y^2}{2})}$ Combining both, we get, $f_{x,y}(X,Y) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{x^2+y^2}{2})}$ Finally we write it in form of (s, θ)

$$f_{x,y}(X,Y) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{x^2+y^2}{2})}$$
$$f_{x,y}(X,Y) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{s^2}{2})}$$

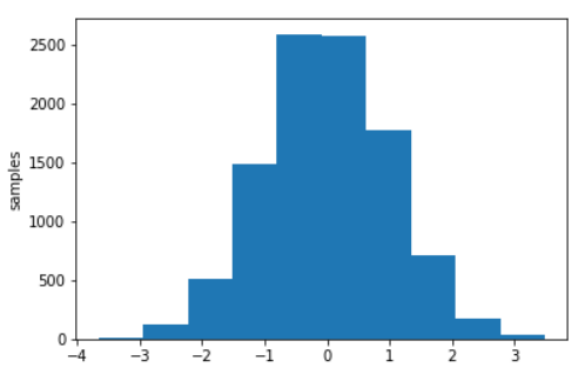
Let $h = \frac{s^2}{2}$ or we can say $s = \sqrt{2h}$ So, we get the following,

$$f_{s,\theta}(S,\Theta) = \frac{1}{\sqrt{2.\pi}}.\exp{(-\frac{s^2}{2})}$$
$$f(h,\theta) = \frac{1}{\sqrt{2.\pi}}.\exp{(-h)}$$
$$or \quad f(h,\theta) = U \sim (0, \frac{1}{2\pi}).Exp \sim (h)$$

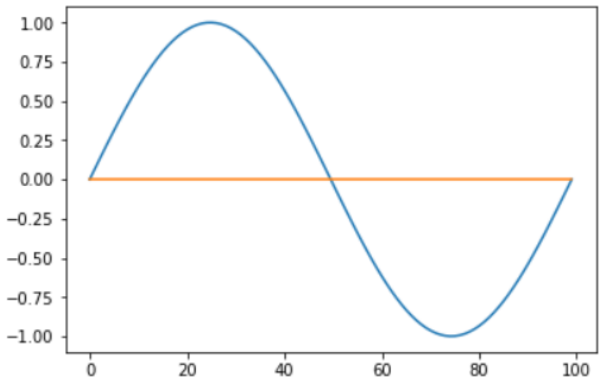
Finally we sample 2 N different Uniform distributions, or as we did here sample 2N samples from uniform distribution and divide it into two sets U1 and U2 of N samples each. So, finally we get two distributions. Then after applying inverse distribution transform to the above equations, we get

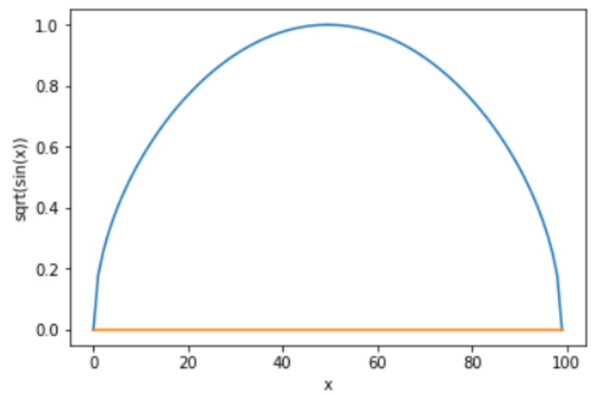
$$x = \sqrt{(-\ln(U_1)}\cos{(2.\pi.U_2)}$$
$$y = \sqrt{(-\ln(U_1)}\sin{(2.\pi.U_2)}$$

Finally we concatenate it into N (x,y) pairs and finally plot it as a histogram.



3. (a) Plot the $\sqrt(\sin(x))$ function. Find the area under the curve of $\sqrt(\sin(x))$ in the interval $(0, \pi)$ using concepts taught in class not using numerical techniques.
Solution: Here in this question, we use the Monte Carlo method of integration. So we capture the fact that in the given question, we have to find the area of the curve $\sqrt(\sin(x))$. Before moving to the technique of calculating the value, let us first take a look at the curve. Below is drawn two distinct curves one of $\sin(x)$





and the second is of $\sqrt(\sin(x))$. Now as we can see that the given curve extends between 0 to 1 in the y-axis and 0 to π in the x-axis. Now as we are aware of that

$$E(g(x)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$
$$E(\pi.\sqrt(\sin(x))) = \int_{0}^{\pi} \pi.\sqrt(\sin(x)).\frac{1}{\pi}dx$$

Now in the above equation, we can see that f(x) is our distribution so we can take it as uniform random values between 0 and π. Now the distribution becomes, $f(x) = \frac{1}{\pi}$. Hence, we generate uniform distribution values between 0 and π for 10000 times. Now we multiply the same by π since our new function is $g^{'}(x) = \pi.\sqrt(\sin(x))$. After that we take the mean of the samples to get our required area. The value of the Area comes out to be 2.398255227208059. Now to cross verify, we simulate the area by taking small boxes from points(equidistant points) in between 0 and π for 10000 instances and find the area of the rectangles formed by taking the product of the minimum length of the two adjacent heights and base($\frac{\pi-0}{10000}$). Thus, adding all the rectangle's area we get the result as 2.396248980018034 which is our actual method to integrate. Thus, we can see that the values are quite close. Hence, verified.

(b) Here, we approach the solution in two ways. The first is we take our solution similarly as we did the above one. So we are already familiar with the equation $E(g(x)) = \int_{-\infty}^{\infty} g(x)f(x)dx$. Now here f(s) is our distribution function. We mention the techniques followed in simplifying the problem and finally solving it.

i. **Technique 1:** Here, we sample our values from **Uniform Distribution**. So we take N=10000 samples and then get the values in $[0, 1]$ range. So then we solve it in the way mentioned below

$$E(g(x)) = \int_{-\infty}^{\infty} g(x).f(x)dx$$

Let us consider $g^{'}(x) = \pi.g(x)$

$$E(g^{'}(x)) = \int_{-\infty}^{\infty} g^{'}(x).f(x)dx$$
$$E(g^{'}(x)) = \int_{0}^{\pi} \pi.(\sqrt{\sin(x)}.\exp(-x^2)).\frac{1}{\pi}dx$$

Hence, required integral is actually formed. So, we now pass our samples through the function $g^{'}(x) = \pi.\sqrt{\sin(x)}.\exp(-x^2)$. Then finally take the average or mean of the samples formed, we shall get the required area.The answer given by this method is 0.5803785639888467.

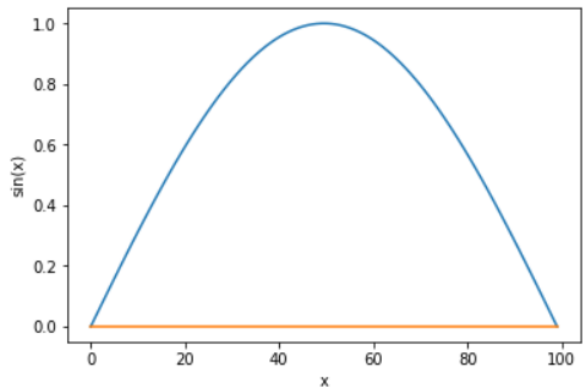
ii. **Technique 2:** In this technique we take our sample not from a uniform distribution but from a **Normal Distribution**. Since Normal distribution has more samples in the vicinity of the mean so the samples generated will be more accurate as compared to the uniform distribution. But in this technique we have to take care of certain parameters. Let us see how to do this. Since here our distribution is Normal distribution, our required distribution function namely f(x) will have the following form $f(x) = \frac{1}{\sqrt{2.\pi.\sigma}}.\exp{(-\frac{(x-\mu)^2}{2\sigma^2})}$. So, here my g(x) will not be the same as before, in this case our g(x) will be as follows $g(x) = \sqrt(\sin(x))$. Now since the domain of the samples in the Normal distributions is from $(-\infty, \infty)$, we have to restrict with our required domain or $[0, 1]$. So to do this, first we have to take the absolute value of the samples so that all the negative values are converted into positive samples. After taking the absolute samples, we restrict our samples between $[0, \infty)$. So our f(x) changes as follows $E(g(x)) = 2.\int_0^\infty g(x)\frac{1}{\sqrt{2.\pi.\sigma}}.\exp{(-\frac{(x-\mu)^2}{2\sigma^2})}$. So, next we define our mean and standard deviation to define the random distribution to sample from. Since, we need $\exp(-x^2)$ term so, we choose $\mu = 0 \ \ and \ \ \sigma = \frac{1}{\sqrt{2}}$. So we got the function reduced to the form. $E(g(x)) = 2.\int_0^\infty g(x)\frac{1}{\pi}.\exp{(-x^2)}$. So, now if we consider $g^{'}(x) = \pi.g(x)$ and then calculate $E(g^{'}(x)) = 2.\pi.\int_0^\infty g(x)\frac{1}{\pi}.\exp{(-x^2)}$. Now let us divide our integration as given below $E(g^{'}(x)) = 2.\pi.\int_0^\pi g(x)\frac{1}{\pi}.\exp{(-x^2)}+$ $2.\pi.\int_\pi^\infty g(x)\frac{1}{\pi}.\exp{(-x^2)}$. Now if suppose we do not sample from the range when $x < 0 \ \ or x > \pi$. Then we get the second integral namely $2.\pi.\int_\pi^\infty g(x)\frac{1}{\pi}.\exp{(-x^2)} = 0$ as g(x) is 0 in this regions. We did not sample from this region so no question of getting any area in this region. Now our redefined formula becomes $E(g^{'}(x)) = 2.\pi.\int_0^\pi g(x)\frac{1}{\pi}.\exp{(-x^2)}$ or $\pi.E(g(x)) * \frac{1}{2} = \pi.\int_0^\pi g(x)\frac{1}{\pi}.\exp{(-x^2)}$. So finally we take the mean of the samples multiply it by π and then divide it with 2 to get our required area.The answer that comes out in this technique is 0.574131338102294. Finally let us conclude with technique 3

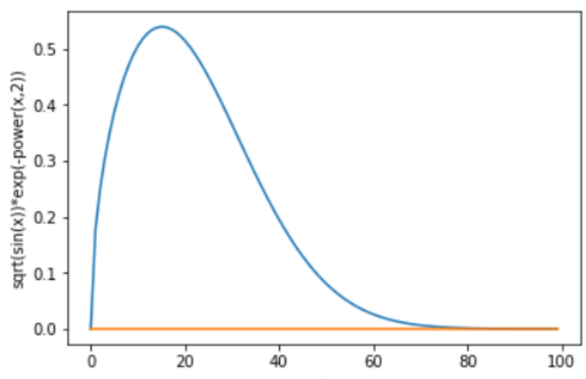
iii. **Technique 3:** Now to cross verify, we simulate the area by taking small boxes from points(equidistant points) in between 0 and π for 10000 instances and find the area of the rectangles formed by taking the product of the minimum length of the two adjacent heights and base($\frac{\pi-0}{10000}$). Thus, adding all the rectangle's area we get the result as 0.5748303993459218 which is our actual method to integrate. Thus, we can see that the values are quite close. Hence, verified.

Finally let us look at the plot of the given function

4. Consider a very peculiar game of Snakes and Ladders shown in Figure 1. The player starts with their pawn in state 0. In each step, a fair six-sided die is tossed. If the die outcome is 1, the player's pawn can move ahead one position, if the outcome is 2, the pawn can move ahead 2 positions, and for any other outcome, the pawn has to stay in the same position. Notice that there is a snake as well as a ladder between the positions 2 and 4. Thus if the pawn visits any of these states, it will get stuck in an infinite loop and the game would never end.

(a) Represent this game as a Markov Chain and draw it clearly showing the states,





transitions and their probabilities. (b) What is the probability that starting from position 0, the player ever reaches the end state? You solve this using simulations. Try to solve it analytically as well and cross check.

Solution:

(a) Here we considered 9 different states running from 0 to 8. Now we are sure to begin from '0' and our target is 8th position. So we follow Markov chain technique(Memorylessness property). Just that if our present state is '2' then it will surely go to state '4' in the next time instant and if the present state is '4' then it will surely move to state '2' in the next time stamp. So in all it is an infinite oscillation between states 2 and 4.

Assumptions:

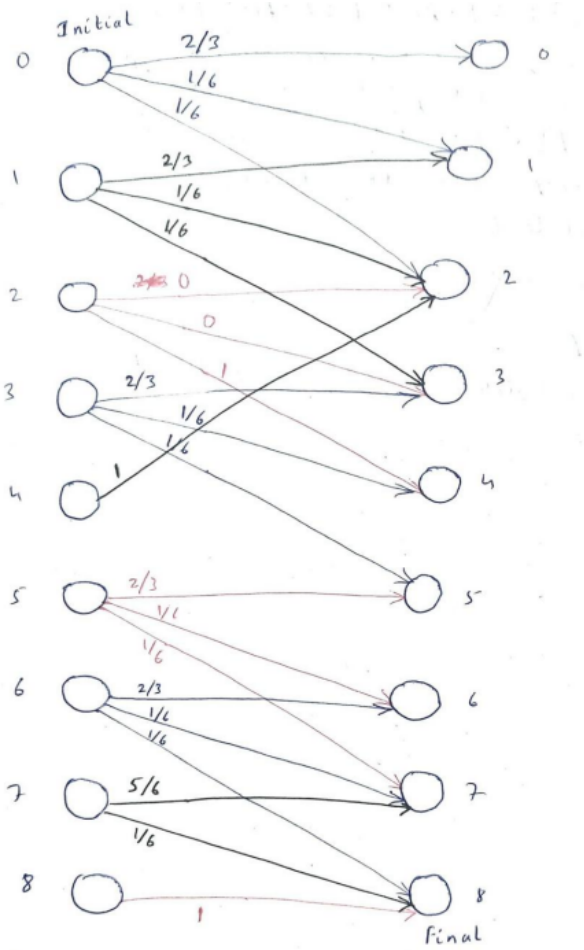
i. Number of trails is **N=10000**

ii. When my pawn is in the 7th step, we assumed that, it needs a dice roll output '1' only to end the game otherwise it will remain in the same state even if'2' comes out. $P(8|7, op = 1) = \frac{1}{6}$
$P(8|7, op = 2, 3, 4, 5, 6) = \frac{5}{6}$
Note: op denotes the outcome of the dice.

iii. When the pawn is in 8th state then it will remain in that state whatever the outcome of the dice is.

$P(1comesout) = \frac{1}{6}$
$P(2comesout) = \frac{1}{6}$
$P(\{3, 4, 5, 6\}comesout) = \frac{4}{6}$
Let us check the transition matrix. Now that we have seen the transition matrix let us find the pictorial representation of the states. Our assumption is that Initial set of states are shown in one column and the next time stamp same states are copied. They are same states. To make the diagram neat, we have created the states for two different instances as shown below. The transition probabilities

$$T = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 2/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2/3 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$



are shown by edges in between the nodes at different levels. We can assume the given graph to be a form of Bipartite graph such that none of the states can communicate with each other in the same time stamp.

(b) For calculating the Probability of winning or that the pawn reaches state 8 we first declare a few points. Here we have 3 final states, namely state 2,4 and 8. Now if the pawn end up at any of these states the episode is said to be complete. When the pawn is in state '8', we consider it to be success for the pawn the game is declared to be a win, so we increase the score of the success variable that is assigned to keep track how many times out pawn is ending at the success state. If the pawn ends up in state 2 or state 4 then it is declared to be failure of the pawn and the game is declared to be a lost. So, we increase a regret variable kept to count the instances how many times our pawn loses. Finally to find the probability of the pawn ending in state 8 or our success probability is found by dividing the value of success variable by the total number of trials (N=10000).

$$\begin{aligned} f_{02} &= TP[02] + \frac{1}{6}.f_{12} + \frac{4}{6}f_{02} \\ Also, \quad f_{12} &= TP[12] + \frac{1}{6}f_{32} + \frac{4}{6}f_{12} \\ or, \quad f_{32} &= \frac{4}{6}f_{32} + TP[34] \end{aligned}$$

Note TP[xy] denotes the transition probability of moving from state 'x' to state 'y'. Thus, finally my equation become

$$\begin{aligned} f_{02} &= \frac{1}{6} + \frac{1}{6}.f_{12} + \frac{4}{6}f_{02} \\ Also, \quad f_{12} &= \frac{1}{6} + \frac{1}{6}f_{32} + \frac{4}{6}f_{12} \\ or, \quad f_{32} &= \frac{4}{6}f_{32} + \frac{1}{6} \end{aligned}$$

On calculating we find that the values of $f_{32} = f_{34} = \frac{1}{2}$
$f_{12} = \frac{3}{4}$
$f_{02} = \frac{7}{8}$ Now since, transition between states 2 and 4 is with probability 1. So reaching any of the states namely 2 or 4, we can declare it will surely traverse the other state. Hence, reaching any one of the failure state is enough to declare that the pawn has failed. P(reaching state 2 or 4 from state 0)=P(failure of the pawn)=$\frac{7}{8}$
P(reaching state 8 from state 0)=P(success of the pawn)=$\frac{1}{8}$
Our simulation replicate the results with the printed answer copied below
Simulation final result:
Probability that our pawn will end in the end state(state 8): 0.12406
Probability that the pawn will end up in state 2 or state 4: 0.87594