Reinforcement Learning **(Date:25-01-2022)**

Lab Assignment 1

Instructor: Dr. Prabhuchandran K J *By:* 211022001, 211022002, 211022005

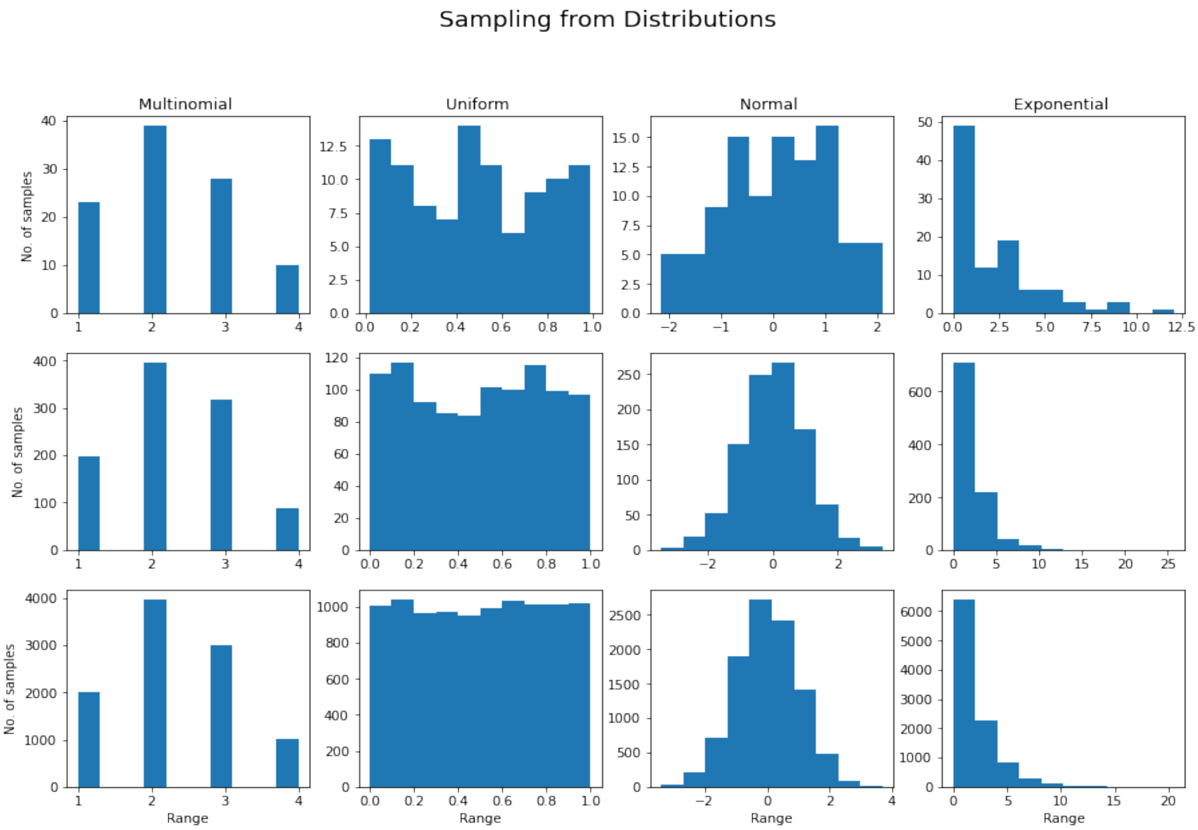
1 Problem 1

1.1 Aim

To plot and observe the generation of samples and convergence of the samples to parent distribution as number of samples increases for the following distribution,

1. Multinomial distribution with four outcomes say $\{1, 2, 3, 4\}$ with corresponding probabilities $[0.2, 0.4, 0.3, 0.1]$.
2. Uniform Distribution in 0 to 1.
3. Gaussian Distribution with mean 0 and variance 1.
4. Exponential Distribution with rate parameter $= 0.5$.

1.2 Results



2 Problem 2

2.1 Aim

To generate samples from normal distribution with mean μ and variance σ^2 using the samples generated from uniform distribution of interval $[0, 1]$.

2.2 Inverse Sampling Method

Inverse Sampling Method uses the inverse of Cumulative Distribution Function (CDF) of the target distribution to generate sample. The samples from uniform distribution is given as input to the inverse CDF function.

Let X be a random variable whose CDF is F_X. The steps for inverse sampling method is as follows :

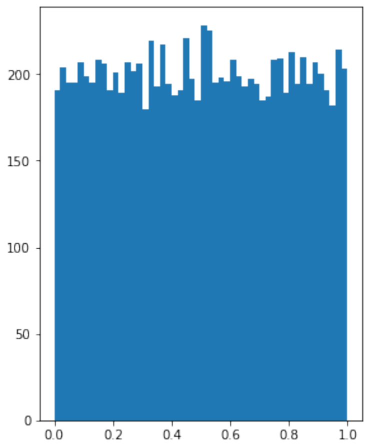
- Generate random samples u from the standard uniform distribution in the interval $[0, 1]$, i.e., from $U \sim \text{Unif}[0, 1]$.
- Find inverse of the CDF i.e., $F_X^{-1}(x)$.
- Compute $X = F_X^{-1}(u)$, the computed X is the samples from the distribution with CDF F_X.

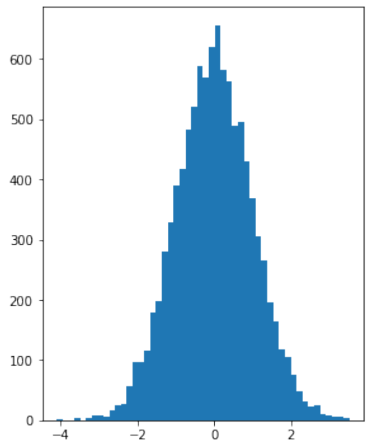
Proof :

$$F_X^{-1}(u) = inf\,\{x|F_X(x) \geq u\} \text{ for } (0 < u < 1)$$
$$P(F_X^{-1}(U) \leq x) = P(U \leq F_X(x)) = F_X(x) \text{ because } P(U \leq y) = y, \text{ when } U \sim \text{Unif}[0, 1].$$

2.3 Result for Inverse sampling Method

For samples n=10000





2.4 Box-Muller Method

The problem with inverse sampling is

- The integral for normal distribution does not have close form solution,
- Computationally complex method.

Box-Muller method is a more simple method to generate samples from normal distribution. In this method we uses the assumption that X and Y are two independent normal distribution whose samples are generated using two uniform distribution which are independent and identically distributed.

If (X, Y) is a pair of independent standard normal distribution, then the joint probability density is a product

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} . \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Since the densities are symmetric, we can go to Polar co-ordinate random variables (Θ, R), where $0 \leq \Theta \leq 2\pi$ and $X = R\cos\Theta$ and $Y = R\sin\Theta$. As Θ is uniformly distributed in $[0, 2\pi]$ it can be sampled using Uniform distribution as

$$\Theta = 2\pi U_1$$

. For distribution of R we can write as

$$G(R) = P(R \leq r) = \int_{r=0}^{r} \int_{\Theta=0}^{2\pi} \frac{1}{2\pi} e^{-r^2/2} r dr d\theta = \int_{r=0}^{r} e^{-r^2/2} r dr$$

using change of variables as $r^2/2 = s$ we have $rdr = ds$ and the limit $r = 0 \implies s = 0$ and $r = r \implies s = r^2/2$. The integration will be

$$G(R) = \int_{s=0}^{r^2/2} e^{-s} = 1 - e^{-r^2/2}$$

This R distribution can be generated as

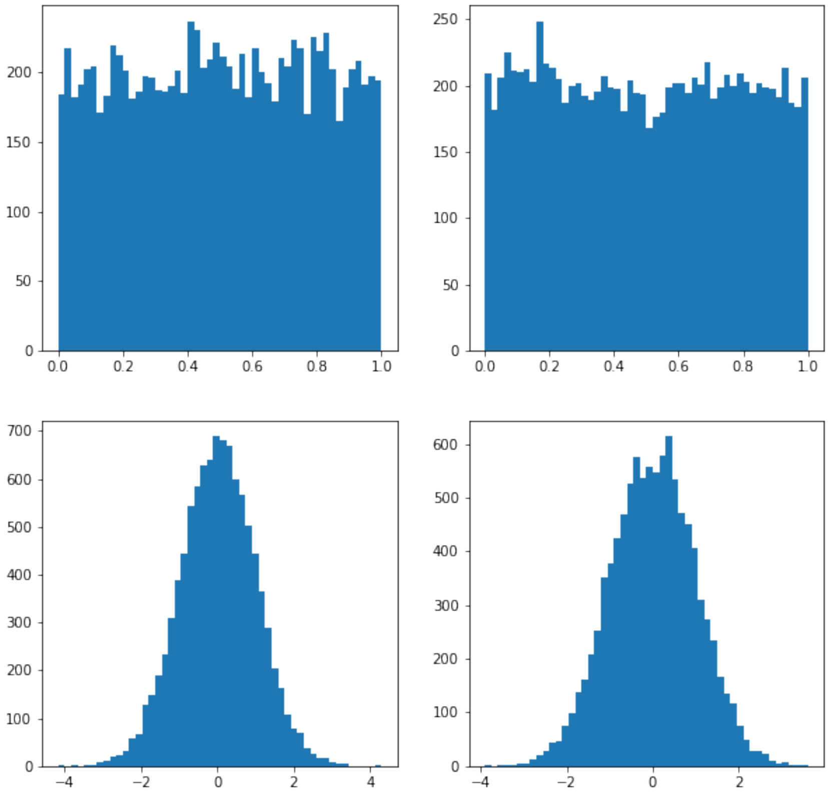
$$G(R) = 1 - e^{-R^2/2} = 1 - U_2$$

. The solution of above will give us $R = \sqrt{-2 \ln U_2}$. Hence the independent standard normal distribution can be generated as

- $X = R\cos\Theta = \sqrt{-2\ln U_2} \cos 2\pi U_1$
- $Y = R\sin\Theta = \sqrt{-2\ln U_2} \sin 2\pi U_1$

2.5 Results for Box-Muller Method

For samples n=10000



[1]

[1]**Please Refer here for more about Box Muller**

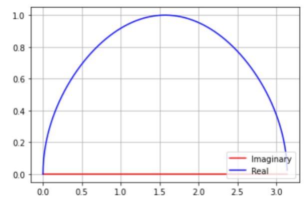
3 Problem 3

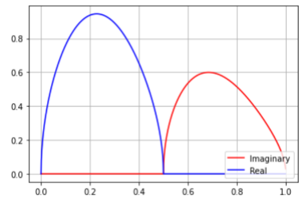
3.1 Aim

Find the area of the given function using the concepts of probability.

3.2 Theory

1. $\sqrt{\sin x}$ in the interval $(0, \pi)$.
2. $\sqrt{\sin x} e^{-x^2}$ in the interval $(0, \pi)$.





To estimate an integral to function $f(x)$ i.e, $\int_a^b f(x)$. This can be writen as $E[g(x)] = \int_a^b \frac{f(x)p(x)}{p(x)}$ where $p(x)$ is distribution function and $g(x) = \frac{f(x)}{p(x)}$
If $p(x)$ is an Uniform Distribution then $p(x) = \frac{1}{b-a}$ for x in interval (a,b) $E[g(x)] = \frac{b-a}{N} \sum_{i=1}^{N} g(x_i)$

However to solve $\sqrt{\sin x}\, e^{-x^2}$ in the interval $(0, \pi)$ We could also use notion of normal distribution $\mathcal{N}\{0, \frac{1}{2}\}$ on x such that

$$E[\sqrt{\sin x}] = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \sqrt{\sin x}\, e^{-x^2} dx$$

. To get probabilities from normal distribution with mean μ and variance σ^2 using a standard normal distribution $\mathcal{N}\{0, 1\}$ we use a transformation from $x \in \mathcal{N}\{0, 1\} \rightarrow z \in \mathcal{N}\{\mu, \sigma^2\}$ such that $z = \frac{x-\mu}{\sigma}$. As x is in interval $(0, \pi)$, We introduce a *factor* in equation that defends the definition of expectation i.e, Expectation of function$f(x)$ whose parameter x is distributed according to a known distribution is fraction of area under the function in the total area under distribution in interval.

$$E\left[\sqrt{\sin z}\right] = \frac{1}{factor * \sqrt{\pi}} \int_0^{\pi} \sqrt{\sin z}\, e^{-z^2} dz$$

To get the fraction of area occupied by $\sqrt{sinx}$ in area of normally distribution

$$factor = \frac{1}{\sqrt{\pi}} \int_0^{\pi} e^{-x^2} dx = F_x\left(\frac{\pi - 0}{\frac{1}{\sqrt{2}}}\right) - F_x(0)$$

So,

$$\int_0^{\pi} \sqrt{\sin x}\, e^{-x^2} dx = factor * \sqrt{\pi} * E\left[\sqrt{\sin x}\right]$$

3.3 Results

- $f1 = \sqrt{\sin x}$ in the interval $(0, \pi)$.

- $f2 = \sqrt{\sin x}e^{-x^2}$ in the interval $(0, \pi)$.

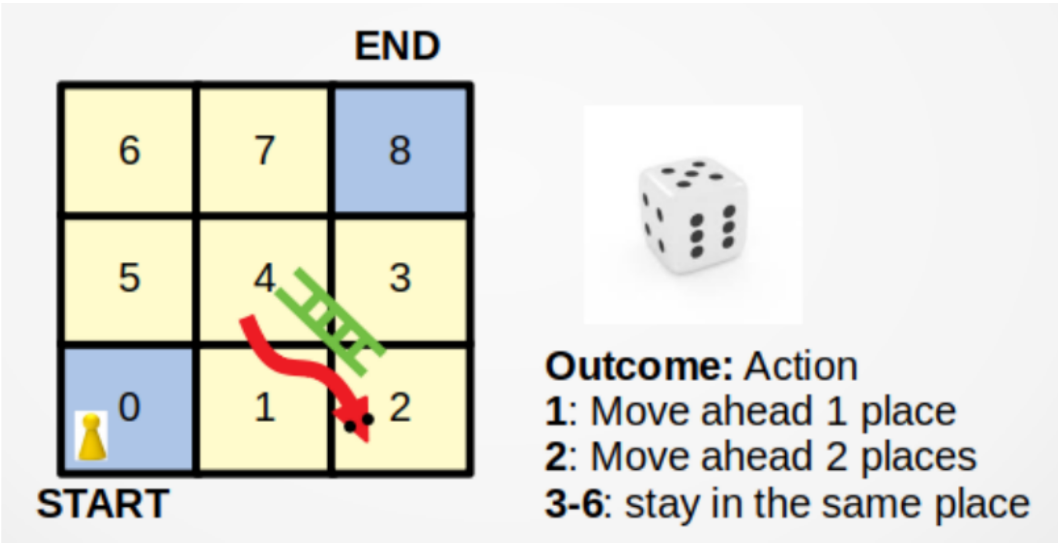
  For samples n=10000

- Area of $f1$: 2.3957490419965475
- Area of $f2$ using uniform distribution : 0.5714130902788518
- Area of $f2$ using normal distribution : 0.5754776996298201

4 Problem 4

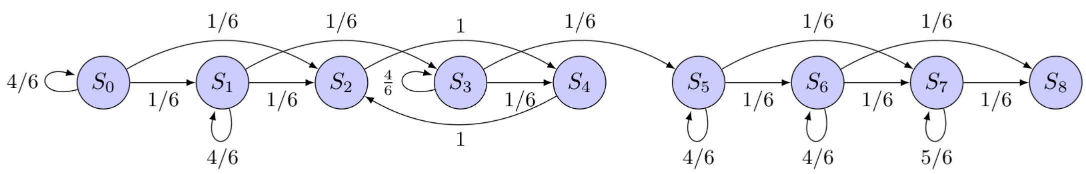
4.1 Aim

To Find the probability of finishing a peculiar Snake and Ladder game using Markov chain and probability concept.



4.2 Theory

State Transition Diagram.



Taking states representing tiles on Snake and Ladder game in range 0 to 8, We get the following P matrix where P_{ij} is probability to go from State-i to State-j

$$P = \begin{bmatrix} \frac{4}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The notion of P^N: Taking Transition probability matrix P ;P^N represents probability of transition from state i to state j in exactly N steps. This can be seen as follows

$$p_{ij}^{(2)} = p_{i1}^{(1)} * p_{1j}^{(1)} + p_{i2}^{(1)} * p_{2j}^{(1)} + ... + p_{in}^{(1)} * p_{nj}^{(1)}$$

Here $p_{ij}^{(N)}$ is probability of going from state i to state j in N transitions.We can see this happening in P^N
Here we are interested in going from State-0 to State-8 over N transitions. Our initial state probability then will be

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2}$$

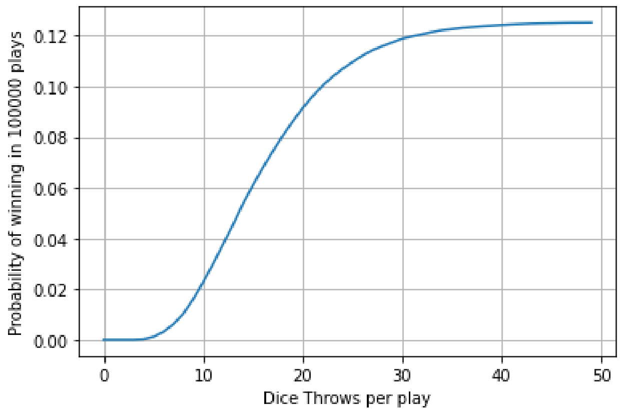
So our final state probability vector is calculated as

$$\Pi_N = S.P^N$$

and the probability of winning can be given by $\Pi_N[8]$ i.e., the last element of the vector.

4.3 Results:

Simulation Results
No. of games played = 1000000
No. of game won is = 12590
Thus Probability of winning is 0.1259.



Comparison of Simulation and Analytical Results.
By Analytical we got Probability of winning as 0.1246.