Reinforcement Learning Lab Report

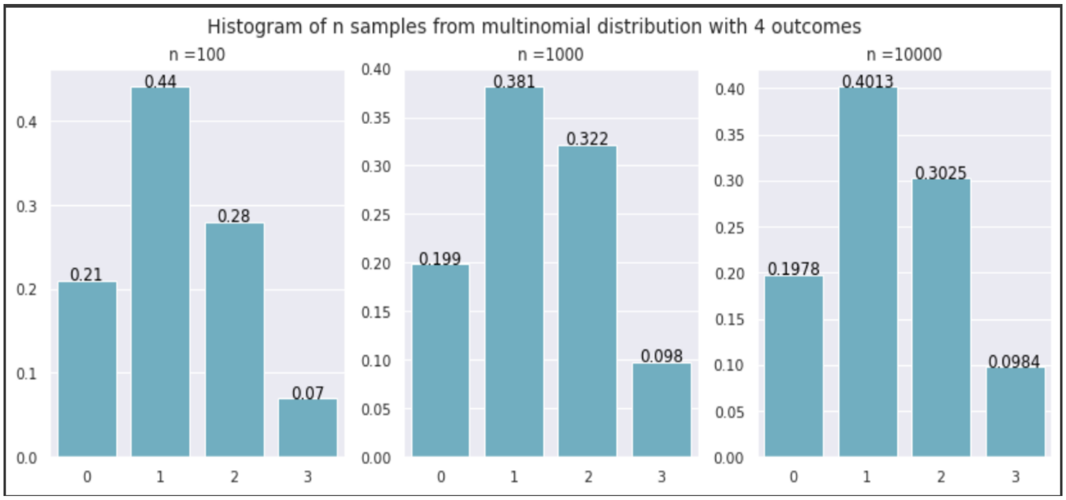
Assignment 1

Jayanth S (201081003)
Praveen Kumar N (201082001)
Rishabh Roy (201082002)

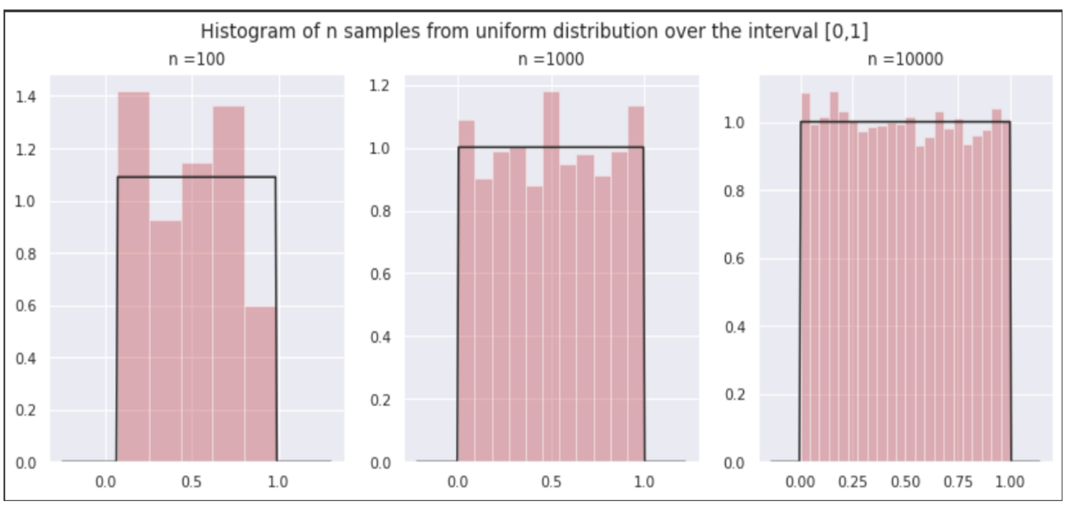
1 Problem 1

Sample points from various distributions and verify if the points are generated according to the respective distribution by plotting a histogram.

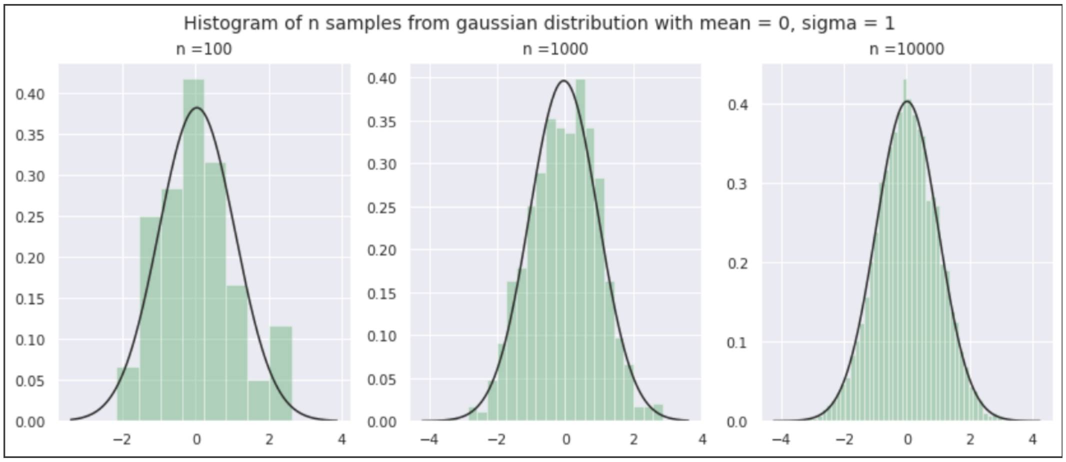
a) Multinomial distribution with probabilities [0.2,0.4,0.3,0.1]:



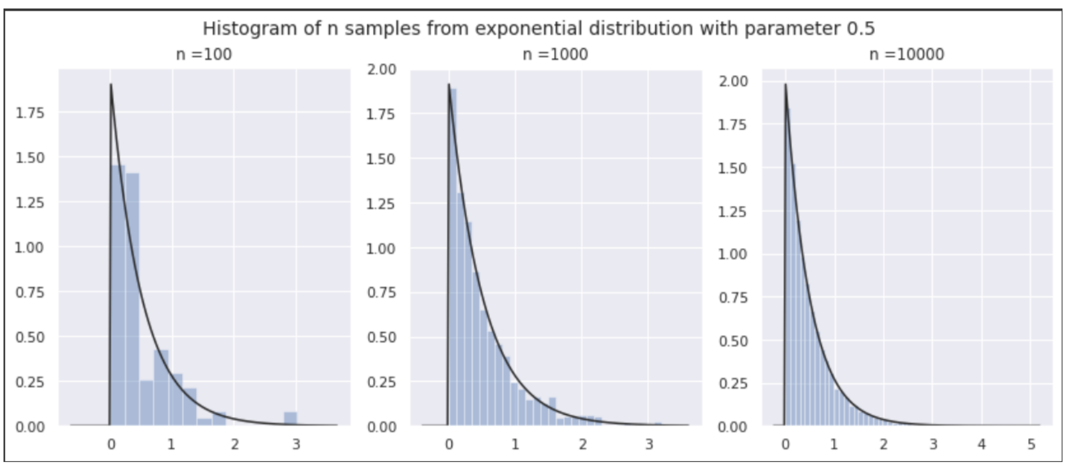
b) Uniform distribution over [0,1]:



c) Normal distribution :



d) Exponential distribution with parameter 0.5:



1.1 Observation :

- As the number of samples n increases, the histogram plot becomes closer to the true distribution as specified by the Law of Large numbers (almost sure convergence or convergence in probability implies convergence in distribution).

2 Problem 2

- We generate the normal random variable samples from uniform random variable using Inverse Transform Method and Box-Muller Transform.
- Inverse Transform Method requires the use of **inverse error function** to generate the normal random variable samples. Whereas Box-Muller transform only requires samples from two independent uniform distributions.

2.1 Inverse Transform Method

The inverse transform sampling method for generating a sample of a Gaussian random variable X with cumulative distribution function,

$$F_X(x) = \frac{1}{2}(1 + \int_0^x e^{\frac{-t^2}{2}} dt) = \frac{1}{2}(1 + erf(x))$$

is as follows :

- Generate the uniform random sample u from uniform random distribution $Unif[0, 1]$.
- Obtain $x = F_X^{-1}(u) = \sqrt{2} erfinv(2u - 1)$. Here $erfinv$ refers to inverse error function which doesn't have a closed form expression (only numerical approximation).
- Also, a sample x of Gaussian random variable, X with mean μ and variance σ^2 is generated as $x = \sqrt{(2\ \sigma^2)} erfinv(2u - 1) + \mu$.

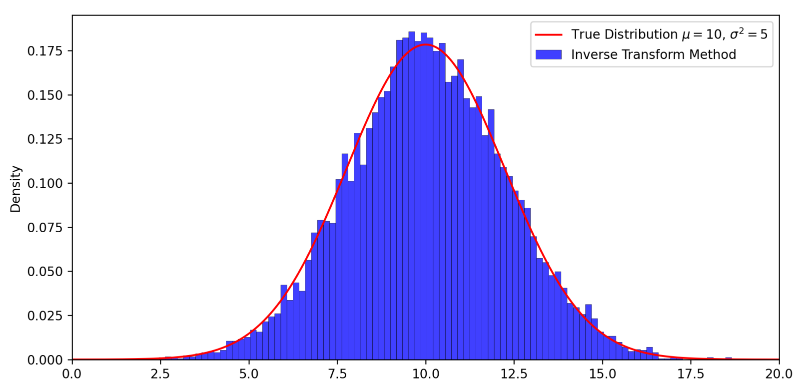


Figure 1: The histogram plot of samples of Gaussian random variable with $\mu = 10$ and $\sigma^2 = 5$ obtained using inverse transform method.

2.2 Box-Muller Method

Box-Muller Method uses samples from two independent standard uniform distribution to generate the Gaussian Random variable. The procedure followed to generate a sample x of Gaussian random variable is:

- Generate the uniform random samples u_1, u_2 from uniform random distribution $Unif[0,1]$ independently.
- Obtain $x = \sqrt{-2log(u_1)}cos(2\pi u_2)$.
- Also, a sample x of Gaussian random variable, X with mean μ and variance σ^2 is generated as $x = \sqrt{(-2\sigma^2)log(u_1)}cos(2\pi u_2) + \mu$.

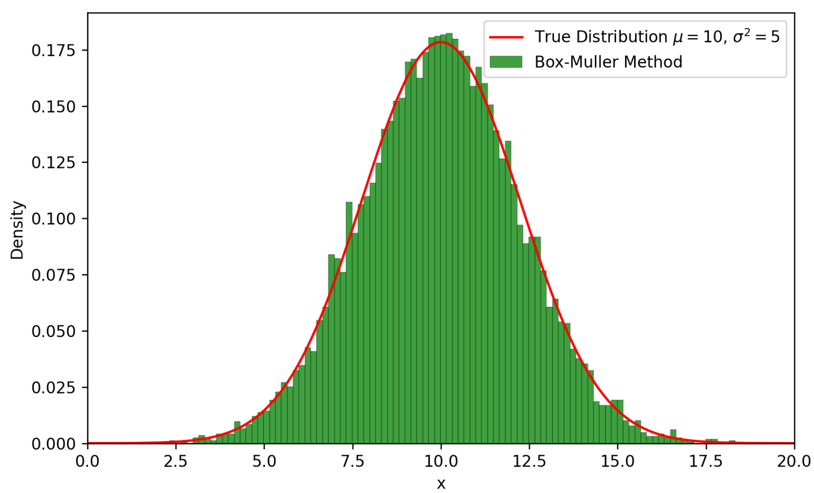


Figure 2: The histogram plot of samples of Gaussian random variable with $\mu = 10$ and $\sigma^2 = 5$ obtained using Box Muller method.

Observations :

- The estimates, $\hat{\mu}$ is **10.051165622671766** and $\hat{\sigma}^2$ is **5.035780448473319** for the samples obtained using inverse transform method.
- The estimates, $\hat{\mu}$ is **9.995409109886058** and $\hat{\sigma}^2$ is **4.973015364343554** for the samples obtained using Box-Muller Algorithm.

3 Problem 3

Find the area under the curve $g(x)$ over an interval $[a, b]$ without using numerical techniques.

Solution :

Given any function $g(x)$ defined over an interval $[a, b]$, we can find the area under the function using the concept of expectation of a continuous random variable.

- First consider the independent variable x to be a uniform random variable over the interval $[a, b]$.
- We know that probability density function (pdf) of a uniform random variable over an interval $[a, b]$ is given by,

$$f_X(x) = \begin{cases} \frac{1}{(b-a)}, a \geq x \leq b \\ 0 \quad , \text{otherwise} \end{cases}$$

- Also we know that expectation of a function of continuous random variable is given by,

$$\mathbb{E}[g(x)] = \int_x g(x) f_X(x) dx$$

- Therefore expectation of a function of uniform random variable can be written as,

$$\mathbb{E}[g(x)] = \int_{x=a}^{x=b} g(x) \times \frac{1}{(b-a)} dx$$

- Now by rearranging the terms in the above equation we can get the area under the function $g(x)$ over the interval $[a, b]$ as,

$$\int_{x=a}^{x=b} g(x) dx = \mathbb{E}[g(x)] \times (b-a)$$

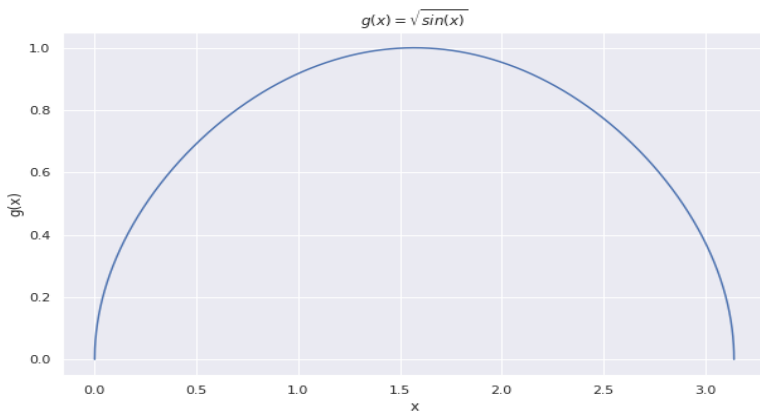
- Since we cannot exactly compute $\mathbb{E}[g(x)]$, we approximate it by the sample mean $\mathrm{S}_n = \frac{1}{n} \sum_{i=1}^{n} g(x_i)$ which will almost surely converge to $\mathbb{E}[g(x)]$ as $n \to \infty$ (by strong law of large numbers). Therefore area under the function $g(x)$ can be approximately obtained as,

$$\int_{x=a}^{x=b} g(x) dx \approx (b-a) \times \frac{1}{n} \sum_{i=1}^{n} g(x_i)$$

- Steps followed to find the area under $g(x)$ in the interval $[a, b]$

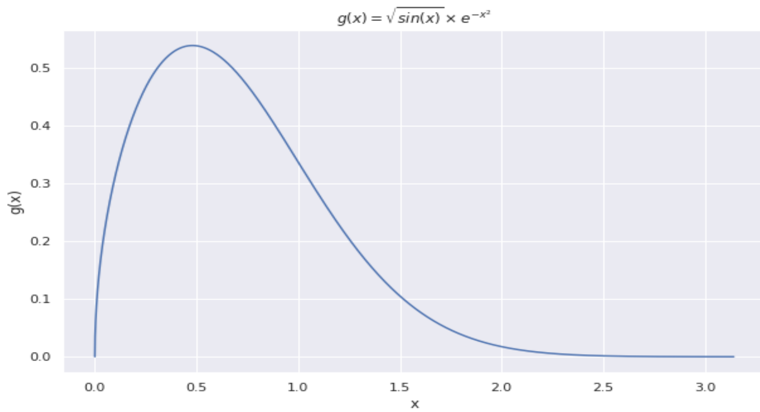
  1. First we generate n samples from uniform distribution in the interval $[a, b]$ and compute $g(x)$ to obtain n samples of $g(x)$.
  2. Next we take the sample mean (S_n) of these n samples to approximate $\mathbb{E}[g(x)]$.
  3. Therefore $\mathrm{S}_n \times (b - a)$ will be approximately equal to the area of the function $g(x)$ and will converge almost surely to the true area as $n \to \infty$ i.e, as $\mathrm{S}_n \to \mathbb{E}[g(x)]$.
  4. Finally we compare the area obtained through this method with the area obtained through integration method in the **scipy** package.

a) $g(x) = \sqrt{sin(x)}$ in the interval $[0, \pi]$:



- Area under $g(x)$ obtained using scipy integration : **2.39628**.
- Area under $g(x)$ obtained using expectation method : **2.39971**.

b) $g(x) = \sqrt{sin(x)} \times e^{-x^2}$ **in the interval** $[0, \pi]$ **:**



- Area under $g(x)$ obtained using scipy integration : **0.57485**.
- Area under $g(x)$ obtained using expectation method : **0.57364**.

4 Problem 4

- We assume that we move from state 7 to state 8 if the outcome of the dice is 2.
- The probability of reaching end state starting from position 0 found using simulation is **0.12514**.
- The probability of reaching end state starting from position 0 found using analytical method is **0.12500**.

Representation of Markov Chain

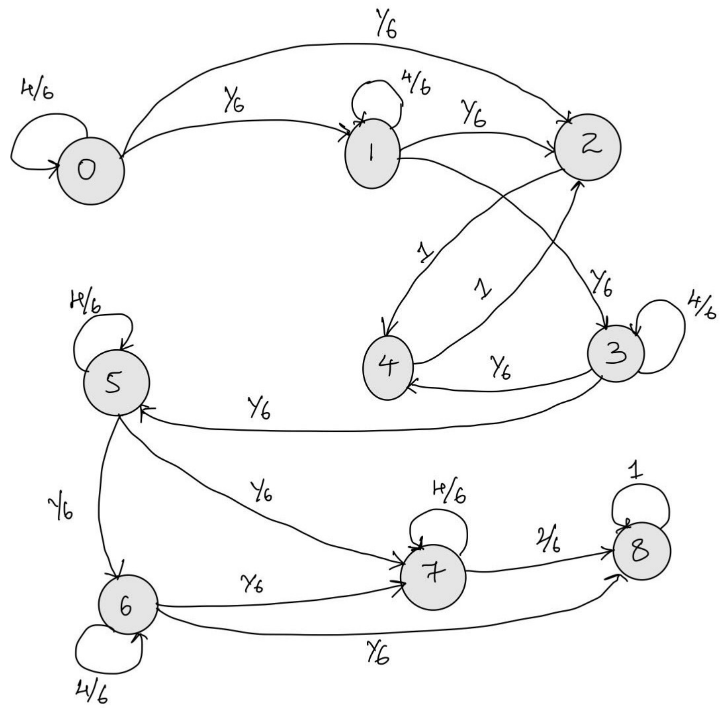


Figure 3: Representation of Snake and Ladder game as a Markov Chain

The corresponding probability transition matrix is given below,

$$P = \begin{bmatrix} 4/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4/6 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4/6 & 2/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Observations :

- The given Markov Chain contains both transient states and recurrent states.

- Since starting form states 2 or 4 we may not be able to reach other states (not accessible), we say that the Markov chain is reducible and hence the Markov chain is not Ergodic.
- Also, steady state distribution doesn't exist because of the transitions between state 2 and 4.

5 References

- How to generate Gaussian samples
- A simple method for numerical integration in python
- scipy.integrate
- PyDTMC github repository