Report: Assignment 1

Ankan Sarkar (210050013)
Soham Joshi (210051004)

August 22, 2022

Contents

1 Introduction

This is the report of Assignment-1 of the course CS215 offered in the autumn semester of '22 in IIT Bombay, by Prof. Suyash Awate. In this report, we will cover the solutions, along with empricial observations and how well they align with the existing theory. We have coded this assignment using MATLAB. The entire code can be accessed in this repository under the folder "code", the graphs and results are given under the folder "results" and this report can be found under the folder "report". Sections $2, \cdots, 6$ of this report correspond to questions $1, \cdots, 5$ of the problem statement. So without further ado, let's start exploring.

2 All about Distributions

Here, we start by dealing with the probability density function (PDF) and the cumulative distribution function (CDF) of Laplace, Gumbel and Cauchy Distributions. In these distributions, we shall plot the PDFs and CDFs and use Riemann-sum to approximately calculate their variance.

2.1 Laplace Distribution

The laplace distribution takes two parameters namely:

1. μ: Location Parameter
2. b: Scale parameter

In our code, mu is referred to by the variable "u", and b by "b". We have used $\mu = 2$ and $b = 2$ as instructed.

The PDF(2.1) has been plotted using the analytic function with the points plotted close enough so that it looks "continuous".

Next, we have used the PDF and applied a Riemann-sum on the PDF from $-e^{15}$ (a negative number sufficiently large in magnitude) to x to compute the CDF(2.1), hence, approximating the equation:

$$CDF(x) := \int_{-\infty}^{x} PDF(x)\, dx \tag{1}$$

Now, we compute variance(2.1) of the laplace distribution. For any distribution with a PDF, say f(x), the expectation is given as :

$$E[X] := \int_{-\infty}^{\infty} x.f(x)\, dx \tag{2}$$

and correspondingly, the variance is given by :

$$Var[X] = E[X^2] - E[X]^2 \tag{3}$$

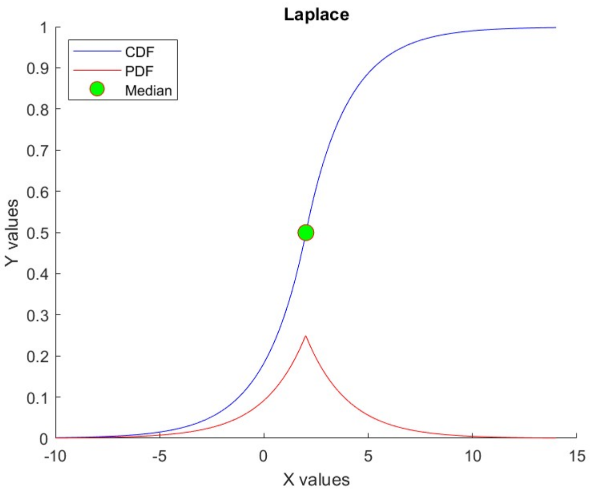
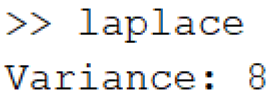


Figure 1: Laplace Distribution

Here, the integrals are calculated using a Riemann-sum, treating $-\infty$ as $-e^{15}$.

Upon calculating the variance analytically, we observe that the value is $2b^2 = 8$, which agrees with our experiment.



```
>> laplace
Variance: 8
```

Figure 2: Laplace Variance

The next two sections deal with the Gumbel and Cauchy distributions with a very similar treatment, hence only the findings and the relevant parameters are mentioned.

2.2 Gumbel Distribution

The Gumbel distribution like the Laplace distribution takes two parameters namely :

1. μ: Location Parameter

2. β: Scale parameter

In our code, mu is referred to by the variable “u”, and β by “b”. We have used $\mu = 1$ and $\beta = 2$ as instructed.

Using a similar procedure as before, we obtain the following distributions

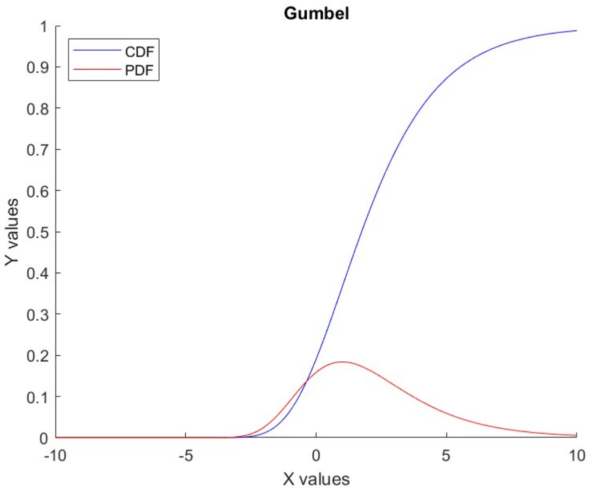
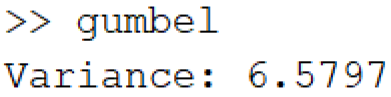


Figure 3: Gumbel Distribution

We now calculate the variance empirically. The value, again, matches with the theoretical value($\frac{\pi^2\beta^2}{6} = 6.5797$) to a great precision.



```
>> gumbel
Variance: 6.5797
```

Figure 4: Gumbel Variance

2.3 Cauchy Distribution

The Cauchy distribution (again!) takes two parameters namely :

1. x_0: Location Parameter
2. γ: Scale parameter

In our code, x_0 is referred to by the variable "x0", and γ by "g". We have used $x_0 = 0$ and $\gamma = 1$ as instructed.

The distribution plots are as follows:

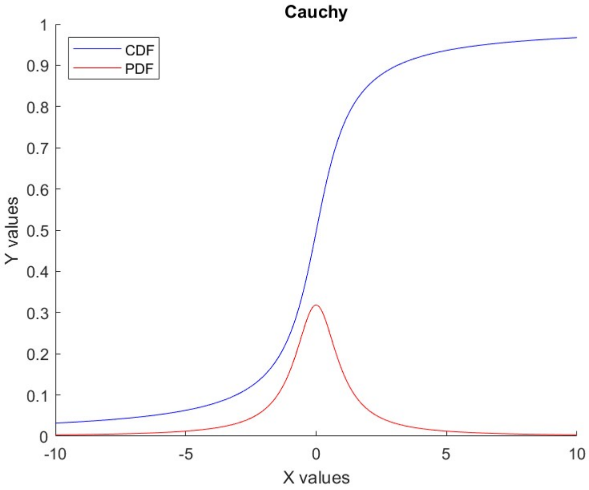


Figure 5: Cauchy Distribution

The theoretical variance for Cauchy distribution is undefined as it diverges. We have made a plot of variance versus log(input size) which shows the same. Hence, we conclude this section.

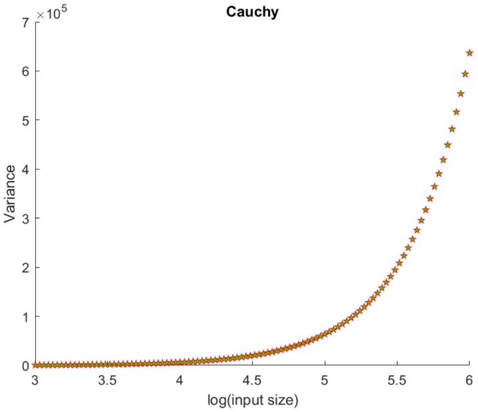


Figure 6: Cauchy Variance

3 Poisson Random Variable

This section deals with Poisson distribution, a probability mass function (PMF). The poisson random variable has the following parameter :

1. λ: Can be interpreted as the average rate of “hits”.

The PMF for the poisson distribution is given by :

$$Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{4}$$

Here, we will calculate the sum of poisson random variables, explore the poisson thinning process and have a look at some comparisons between empirical and theoretical values.

3.1 Sum of Poisson Random Variables

In this subsection, we were instructed to find the empirical estimate of the PMF of sum of two poisson random variables. In order to achieve this, we have used the poissrnd() in-built MATLAB function and used two random variables X and Y with $\lambda_X = 3$ and $\lambda_Y = 4$.

Using this, we have recorded and iterated over $N = 10^6$ instances and stored the results in an array which is used to plot the distribution.

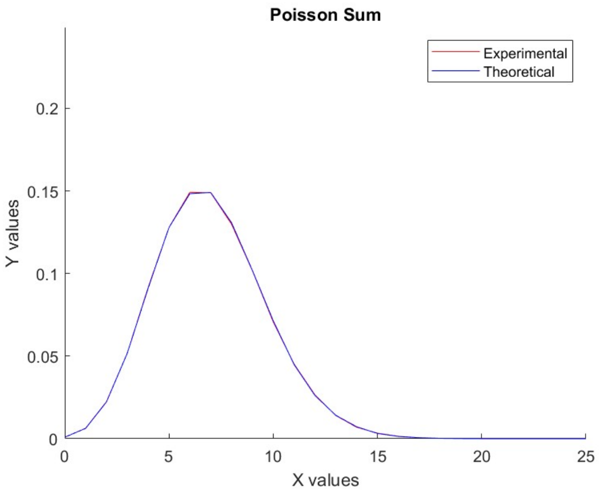
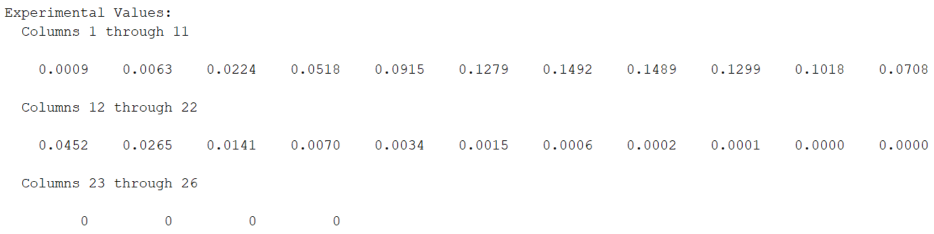


Figure 7: (almost overlapping) Poisson Sum

Now, we have obtained the values of the distribution at $k = \{0, 1, \cdots, 25\}$ from our observations in the array.



```
Experimental Values:
  Columns 1 through 11

    0.0009    0.0063    0.0224    0.0518    0.0915    0.1279    0.1492    0.1489    0.1299    0.1018    0.0708

  Columns 12 through 22

    0.0452    0.0265    0.0141    0.0070    0.0034    0.0015    0.0006    0.0002    0.0001    0.0000    0.0000

  Columns 23 through 26

         0         0         0         0
```

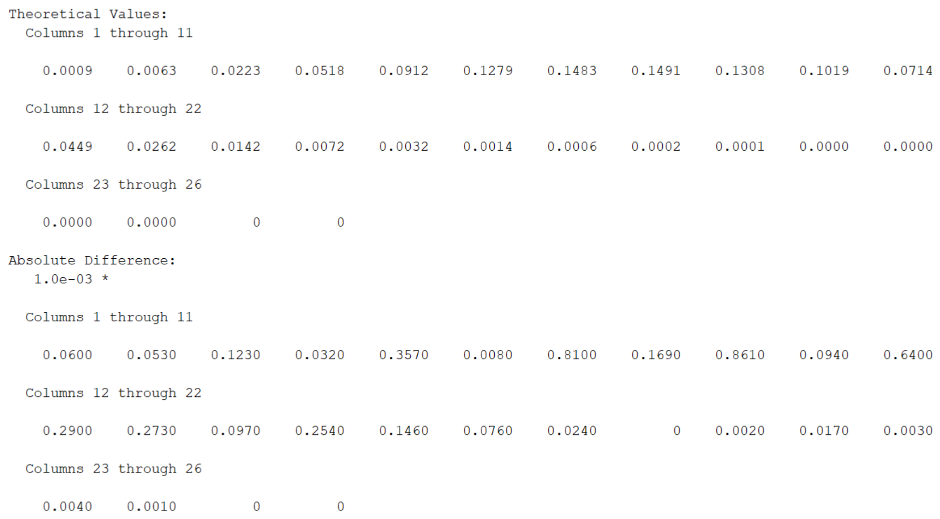
Figure 8: Experimental Poisson Sum Values

Now, let us obtain the same values theoretically and check how well they agree with our empirical data.

Now, upon summing up the random variables, we get the following set of implications. (Complete article)

$$
\begin{aligned}
p_Z(z) &= P(Z = z) \\
&= \sum_{j=0}^{z} P(X = j \ \& \ Y = z - j) \qquad \text{so } X + Y = z \\
&= \sum_{j=0}^{z} P(X = j)P(Y = z - j) \qquad \text{since } X \text{ and } Y \text{ are independent} \\
&= \sum_{j=0}^{z} \frac{e^{-\lambda_1}\lambda_1^j}{j!}\frac{e^{-\lambda_2}\lambda_2^{z-j}}{(z-j)!} \\
&= \sum_{j=0}^{z} \frac{1}{j!(z-j)!} e^{-\lambda_1}\lambda_1^j e^{-\lambda_2}\lambda_2^{z-j} \\
&= \sum_{j=0}^{z} \frac{z!}{j!(z-j)!} \frac{e^{-\lambda_1}\lambda_1^j e^{-\lambda_2}\lambda_2^{z-j}}{z!} \qquad \text{multiply and divide by } z! \\
&= \sum_{j=0}^{z} \binom{z}{j} \frac{e^{-\lambda_1}\lambda_1^j e^{-\lambda_2}\lambda_2^{z-j}}{z!} \qquad \text{using the form of binominal coefficients} \\
&= \frac{e^{-\lambda}}{z!} \sum_{j=0}^{z} \binom{z}{j} \lambda_1^j \lambda_2^{z-j} \qquad \text{factoring out } z! \text{ and } e^{-\lambda_1}e^{-\lambda_2} = e^{-\lambda_1-\lambda_2} = e^{-\lambda} \\
&= \frac{e^{-\lambda}}{z!}(\lambda_1 + \lambda_2)^z \qquad \text{using binomial expansion (in reverse)} \\
&= \frac{e^{-\lambda}\lambda^z}{z!}
\end{aligned}
$$

Hence, we see that in the distribution with two independent poisson variables, the hit rates add up as well. Using this distribution we get the theoretical values as:



```
Theoretical Values:
  Columns 1 through 11

    0.0009    0.0063    0.0223    0.0518    0.0912    0.1279    0.1483    0.1491    0.1308    0.1019    0.0714

  Columns 12 through 22

    0.0449    0.0262    0.0142    0.0072    0.0032    0.0014    0.0006    0.0002    0.0001    0.0000    0.0000

  Columns 23 through 26

    0.0000    0.0000         0         0

Absolute Difference:
   1.0e-03 *

  Columns 1 through 11

    0.0600    0.0530    0.1230    0.0320    0.3570    0.0080    0.8100    0.1690    0.8610    0.0940    0.6400

  Columns 12 through 22

    0.2900    0.2730    0.0970    0.2540    0.1460    0.0760    0.0240         0    0.0020    0.0170    0.0030

  Columns 23 through 26

    0.0040    0.0010         0         0
```

Figure 9: Theoretical Poisson Sum Values

Hence, once again the experimental and theoretical values agree well with the difference being of the order of 10^{-4}.

3.2 Poisson Thinning

In this section, we shall explore the process of Poisson Thinning. While in the poisson process, we measure the probability of a certain number of hits in a unit time interval, in the thinning process we shall observe the probability of getting a number of hits given that the probability of letting the hit pass through is p.

As instructed, we shall use the probability parameter $p = 0.8$, and let us represent the "thinned" random variable by y. We are using the following statement of the thinning theorem.

Suppose that $N \overset{R}{\longleftarrow} Poisson(\lambda)$ and that $X_1, X_2, \cdots$ are independent, identically distributed Bernoulli-p random variables independent of N. Let

$$S_n = \sum_{i=1}^{n} X_i \tag{5}$$

then S_n has the Poisson distribution with mean λp .

Using this definition, we have coded and obtained the empirical estimate of the PMF and values of the empirical probability $\hat{P}(y = k)$ for $k = \{0, 1, \cdots, 25\}$.

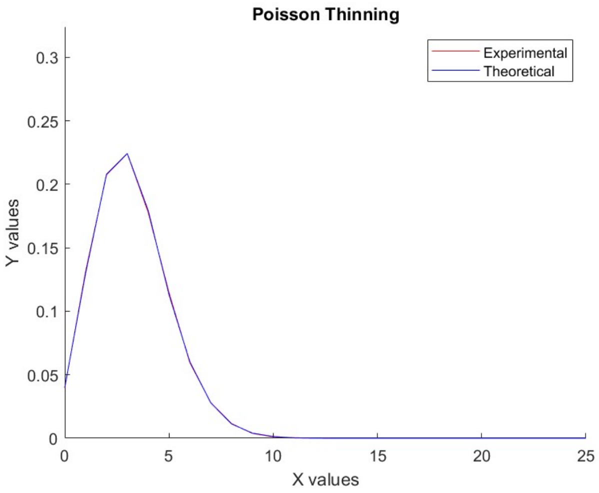
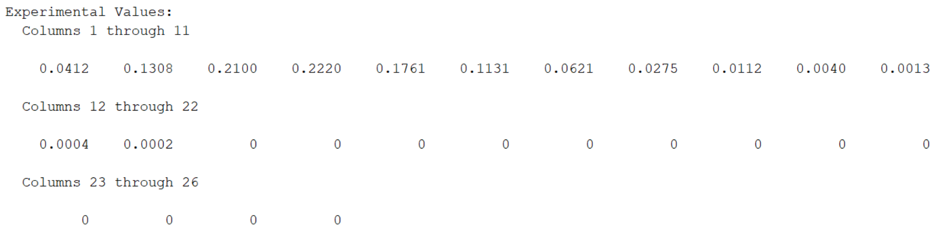


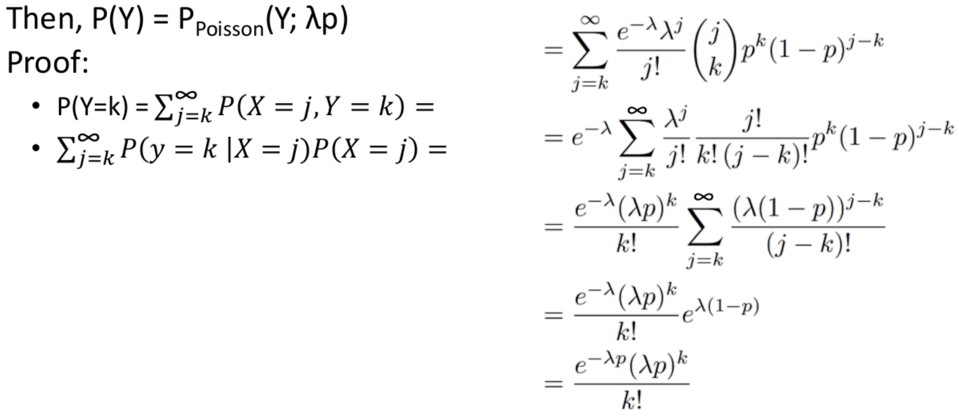
Figure 10: (almost overlapping) Poisson Thinning



```
Experimental Values:
  Columns 1 through 11

    0.0412    0.1308    0.2100    0.2220    0.1761    0.1131    0.0621    0.0275    0.0112    0.0040    0.0013

  Columns 12 through 22

    0.0004    0.0002         0         0         0         0         0         0         0         0         0

  Columns 23 through 26

         0         0         0         0
```

Figure 11: Experimental Poisson Thinning Values

Again, let us derive(taken from slides) a theoretic distribution to check if our observations agree with theory.



Then, P(Y) = P_{Poisson}(Y; λp)

Proof:

- P(Y=k) = $\sum_{j=k}^{\infty} P(X = j, Y = k) =$
- $\sum_{j=k}^{\infty} P(y = k \mid X = j) P(X = j) =$

$$= \sum_{j=k}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!}\binom{j}{k}p^k(1-p)^{j-k}$$

$$= e^{-\lambda}\sum_{j=k}^{\infty} \frac{\lambda^j}{j!}\frac{j!}{k!\,(j-k)!}p^k(1-p)^{j-k}$$

$$= \frac{e^{-\lambda}(\lambda p)^k}{k!}\sum_{j=k}^{\infty}\frac{(\lambda(1-p))^{j-k}}{(j-k)!}$$

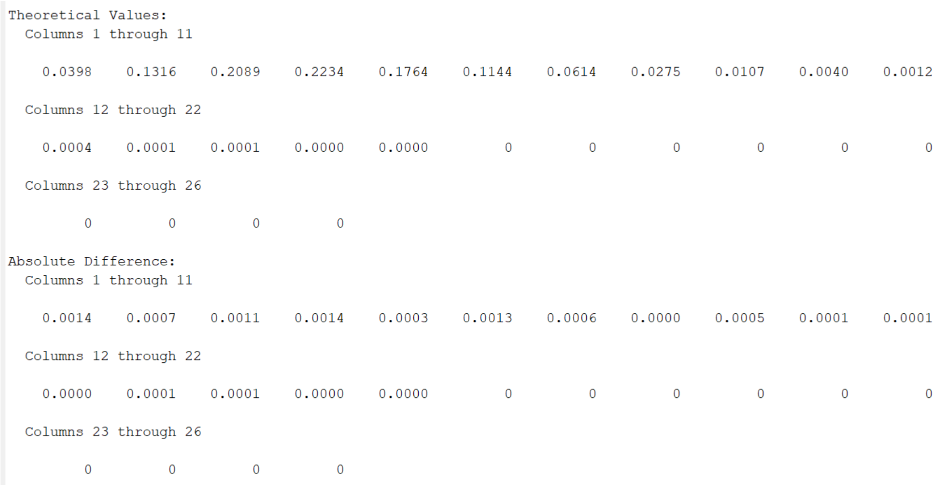
$$= \frac{e^{-\lambda}(\lambda p)^k}{k!}e^{\lambda(1-p)}$$

$$= \frac{e^{-\lambda p}(\lambda p)^k}{k!}$$

Hence, we obtain :

$$P(y = k) = \frac{e^{-\lambda p}(\lambda p)^k}{k!} \tag{6}$$

Now let us compare theory and experiment, i.e. $P(y = k)$ and $\hat{P}(y = k)$ for $k = \{0, 1, \cdots, 25\}$.



```
Theoretical Values:
  Columns 1 through 11

    0.0398    0.1316    0.2089    0.2234    0.1764    0.1144    0.0614    0.0275    0.0107    0.0040    0.0012

  Columns 12 through 22

    0.0004    0.0001    0.0001    0.0000    0.0000         0         0         0         0         0         0

  Columns 23 through 26

         0         0         0         0

Absolute Difference:
  Columns 1 through 11

    0.0014    0.0007    0.0011    0.0014    0.0003    0.0013    0.0006    0.0000    0.0005    0.0001    0.0001

  Columns 12 through 22

    0.0000    0.0001    0.0001    0.0000    0.0000         0         0         0         0         0         0

  Columns 23 through 26

         0         0         0         0
```

Figure 12: Theoretical Poisson Thinning Values

Thus, our empirical data fits with the theoretical data within a margin of 10^{-3}, and we move on to the next section.

4 The Random Walker Problem

This question deals with simulating $N = 10^4$ independent random-walkers each starting from the origin and walking on the real line taking steps of length 10^{-3} either to the left or to the right with equal probability. In this section, we simulate the movement of these random walkers across space-time, analyse their final locations after some time and also have a look at the Law of Large Numbers.

4.1 Space-Time Analysis

Let us start by observing the final position of walkers. We have coded this using an independent Binomial Random Variable for each walker where the value of the random variable, given the number of trials, represents the number of steps taken to the right. Using this, we obtained the following histogram for the final locations of the walkers.

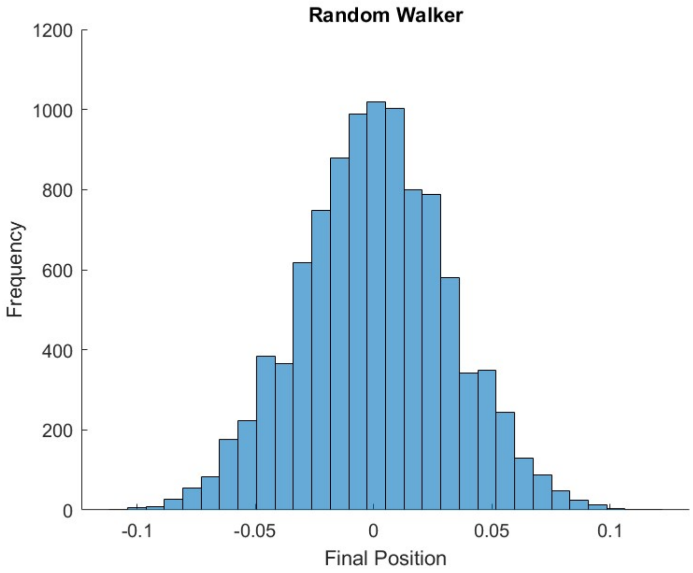


Figure 13: Final Positions of Random Walkers

Now, let us plot the “space-time” curve of the walkers. For this, we have iterated over the number of trials and assigned a Bernoulli Random Variable which determines the side which the walker will step in. We have stored the locations of each walker at a given iteration as a row vector, hence for all the walkers, we have stored each iterated position in a matrix. Simulating this, we get the following plot.

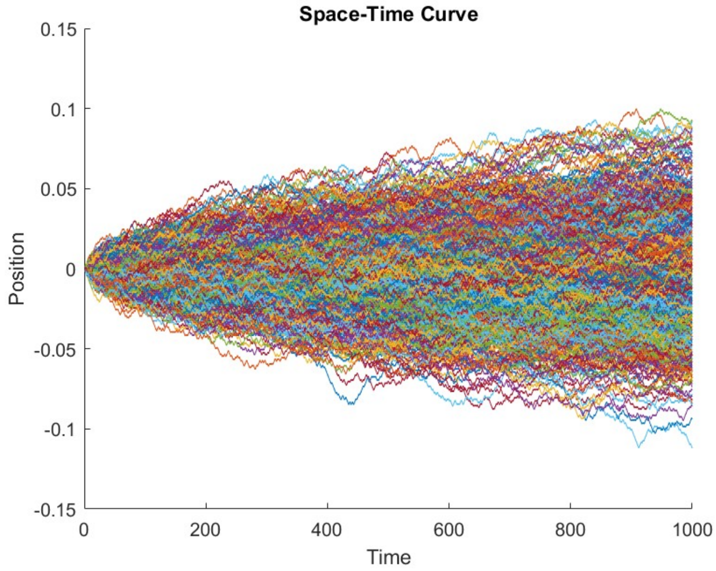


Figure 14: Progressive Positions of Random Walkers

4.2 Convergence Analysis

In this section we show two results :

1. Average of N random variables converges to true mean as $N \rightarrow \infty$
2. Average of variances of N random variables converges to the true variance.

4.2.1 Theoretical Values

Let $X_1, X_2, \cdots X_N$ be N random variables, then, by law of large numbers we know that the random variable that the sample average converges in probability towards the expected value. Hence,

$$\hat{M} = \sum_{i=1}^{N} \frac{X_i}{N} \tag{7}$$

$$\Rightarrow E(\hat{M}) = \mu \tag{8}$$

Hence, we are done proving that the random variable $\hat{M}$ and its expectation converge to the true mean as $N \rightarrow \infty$. Now, proving the second part i.e. variance.

$$\hat{V} = \sum_{i=1}^{N} \frac{(X_i - \hat{M})^2}{N} \tag{9}$$

Hence, opening the brackets and using the previous result, that, $\lim_{x\to\infty} \hat{M} = \mu$

$$E(\hat{V}) = E(\sum_{i=1}^{N} \frac{X_i^2}{N} - 2\frac{X_i \hat{M}}{N} + \frac{\hat{M}^2}{N}) \tag{10}$$

$Y_i = X_i^2$ can be treated as a new random variable, with $E(Y_i) = E(X^2)$. Hence,

$$\Rightarrow \lim_{x\to\infty} E(\hat{V}) = E(X^2) - 2E(X)E(X) + E(X)^2 \Rightarrow \lim_{x\to\infty} E(\hat{V}) = E(X^2) - E(X)E(X) = Var(X) \tag{11}$$

Hence, proved that expected value of $\hat{V}$ converges to the variance of X.

4.2.2 Empirical Values

For this purpose we simulate N random variables, namely $X_1, X_2, \cdots, X_N$ as independent and identically distributed binary variables. We obtain the following plots upon calculating the expectation of the above mentioned scenarios for varied number of random variables.

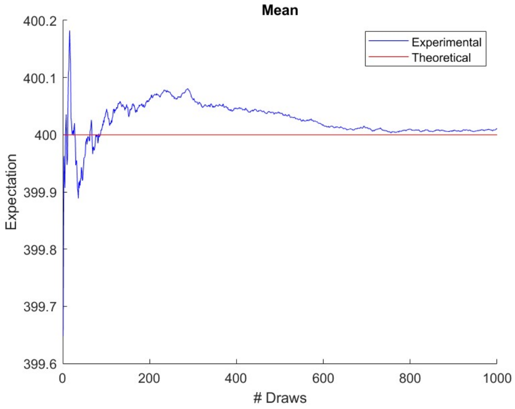


Figure 15: Expectation Values of mean

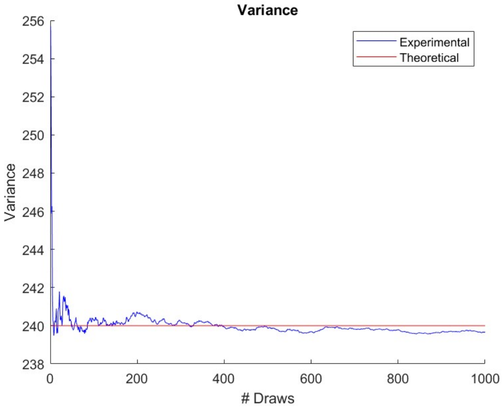
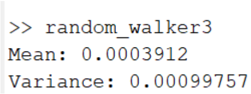


Figure 16: Expectation Values of Variance

As we can see, the empirical expectation eventually converges to the theoretical value.

4.3 Observations

In the random-walker problem, we arrived at the following empirical data through simulations on MATLAB.



```
>> random_walker3
Mean: 0.0003912
Variance: 0.00099757
```

Figure 17: Empirical Mean Variance

But, do these observations agree with theory? Let us derive the expected mean and variance for this purpose.

4.3.1 Expected Mean

For the binomial random variable X modelling the net displacement with step length s and number of steps N,

$$E(X) = \frac{\sum_{i=1}^{N} s \times (2 \times P(X^{'} = 1) - 1)}{N} \tag{12}$$

where, $X^{'}$ is the Bernoulli random variable modelling each step. But, we know that $P(X^{'} = 1)$ is 0.5, hence,

$$E(X) = 0 \tag{13}$$

4.3.2 Expected Variance

For the binomial random variable X modelling the number of steps taken to the right with number of steps N and step length s, let the expected variance of the random variable denoting final position be denoted as V. So, $Y = 2X - N$ models the net number of steps to the left or to the right.

We know that $V = (E(Y^2) - E(Y)^2) \times s^2$. And, $E(Y) = 0$. Hence,

$$V = s^2 \times E(Y^2) \tag{14}$$

$$\Rightarrow V = s^2 \times E(4X^2 - 4NX + N^2) \tag{15}$$

$$\Rightarrow V = s^2 \times (4E(X^2) - 4NE(X) + N^2) \tag{16}$$

We know that the variance of a binomial random variable is npq where n is the number of trials, and p, q are the respective probabilities of success and failure. Also, we know that $Var(X) = E(X^2) - E(X)^2$ and $E(X) = N/2$.

Putting these facts together, we get:

$$\Rightarrow V = s^2 \times (4N \times 0.5 \times 0.5 + 4E(X)^2 - 4NE(X) + N^2) \tag{17}$$

$$\Rightarrow V = Ns^2 \tag{18}$$

Hence,

$$\boxed{V = Ns^2}$$

where, s = step length and N = number of steps

4.3.3 Comparison

Now, let us compare our theoretical expected values and values obtained via code(4.3). Theoretically, we obtain

$$\boxed{E = 0}$$

$$\boxed{V = Ns^2 = 10^{-6} * 10^3 = 10^{-3}}$$

Hence we can see, yet again, that the experiment and theory align, with the error in both the mean and variance being of the order of 10^{-4}

5 PDF Based Random Number Generation

Up until now, we have been generating random values with known random generating functions and known PDFs. However, this question deals with the design of a random generator for a known PDF, that is, the only random generator we have access to is the uniform random number generator taking values between 0 and 1 and from this we have to construct a new random number generator. Moreover, we shall be dealing with the CDF of this PDF and explore the Central Limit Theorem.

5.1 Designing a Random Number Generator

A PDF is given to us as follows :

$$P_X(x) = \begin{cases} 0 & |x| > 1 \\ |x| & \text{otherwise} \end{cases} \tag{19}$$

First, we designed a random variable that gives the above PDF. For this purpose, we created a large array that takes all values from $-N$ to N, where N is a large number, using the fact there there exists a bijection $\mathbb{N} \to \mathbb{Z}$. Then, we divided all elements of this large array by N to approximately obtain the set $[0, 1]$. Hence, choosing an element of the array randomly is equivalent to choosing an element in the set $[0, 1]$. Now for the final trick, we have created the array such that the frequency of the elements increases linearly with their absolute value, hence, we obtain a random variable that satisfies the conditions given in the problem!

The histogram of the PDF and CDF of this random variable are:

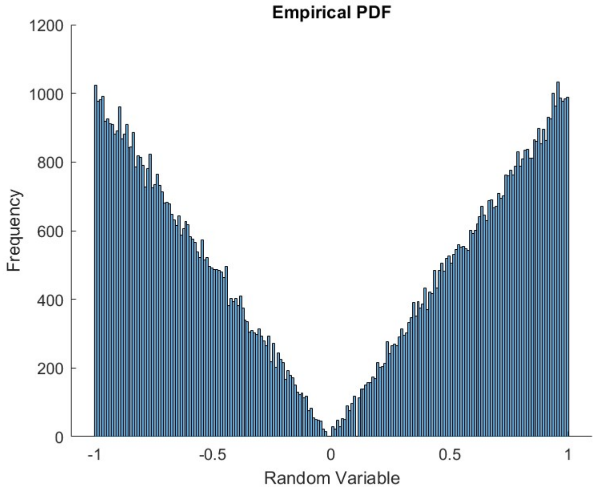


Figure 18: Histogram PDF

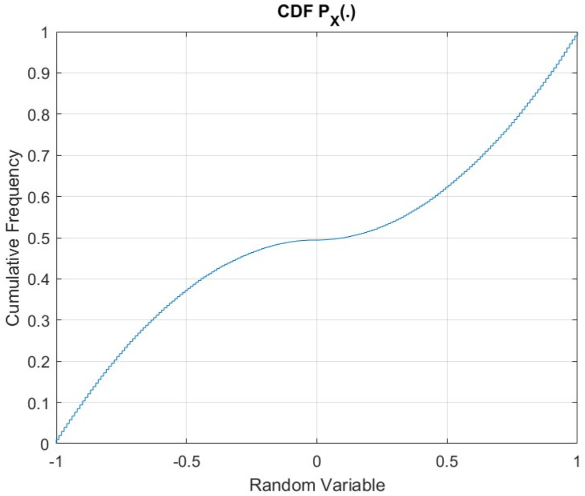


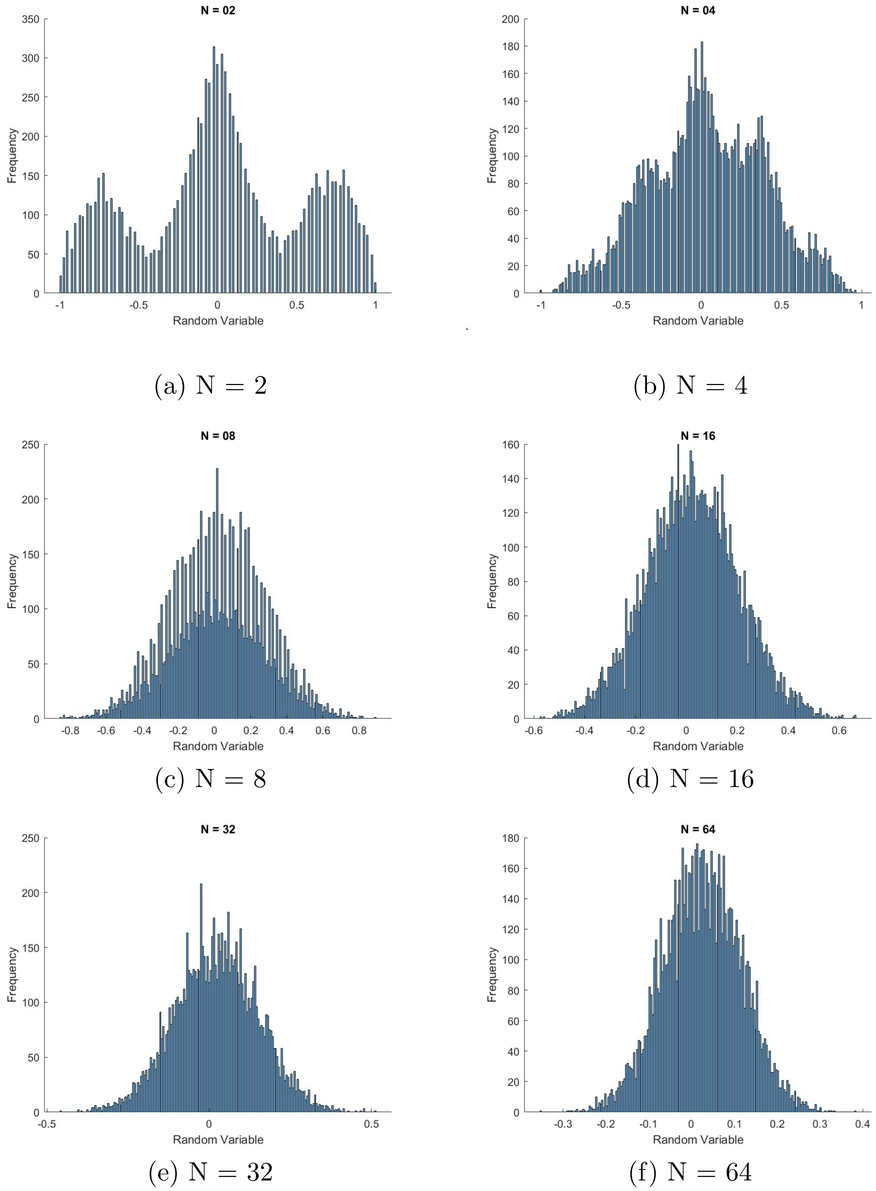
Figure 19: CDF

5.2 Multiple Random Variables

Now, that we have a random variable which models the desired probability distribution, we can generate a new random variable as follows:

$$Y_N = \frac{\sum_{i=1}^{N} X_i}{N} \tag{20}$$

Now, this new random variable has a distribution of it's own, which varies as we vary N. Hence, we will get a different PDF for each value of N for this random variable Y_N. As instructed, we have taken the values of $N = \{2, 4, 8, 16, 32, 64\}$. The graphs are given below(20).



(a) N = 2

(b) N = 4

(c) N = 8

(d) N = 16

(e) N = 32

(f) N = 64

Figure 20: PDFs of Y_N

Upon observing these graphs, we can see that with increasing values of N, the graphs tend more and more towards a Gaussian Distribution, hence verifying the Central Limit Theorem. Similarly, we get the CDFs for different values of N as follows:

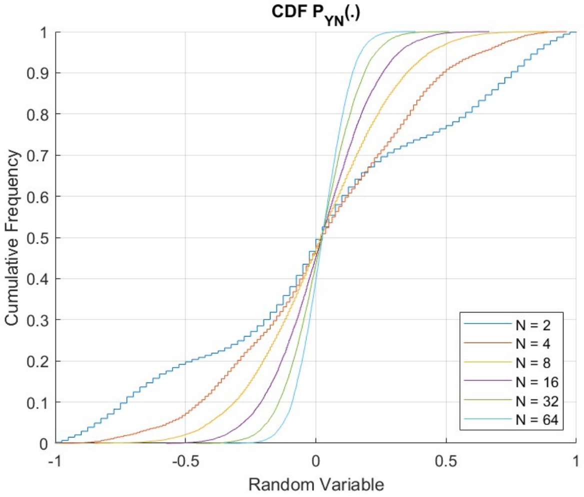


Figure 21: CDFs of Y_N

Hence, we are done with this analysis.

6 Analysis of Error

This question deals with the generation of datasets, boxplots and the analysis of distribution of errors. We generate dataset of sizes $N = \{5, 10, 20, 40, 60, 80, 100, 500, 10^3, 10^4\}$ and for each N we repeat the following $M = 100$ times:

1. Generate data
2. Compute average $\hat{\mu}$
3. Measure error $|\hat{\mu} - \mu_{true}|$

6.1 Plotting Errors

In order to plot the errors across all $M = 100$ iterations, we store the values of errors in an array, and then plot that array using a “box-and-whisker” plot, commonly called a boxplot.

Hence, in totality, we will have multiple vertical boxplots stacked side by side since we are plotting multiple values of N simultaneously. We obtain the boxplots for:

1. Uniform Distribution
2. Gaussian Distribution: $\mu = 0$ and $\sigma^2 = 1$.

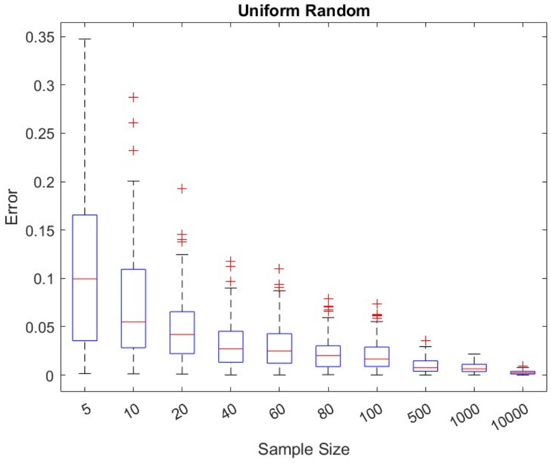


Figure 22: Errors in Uniform Distribution

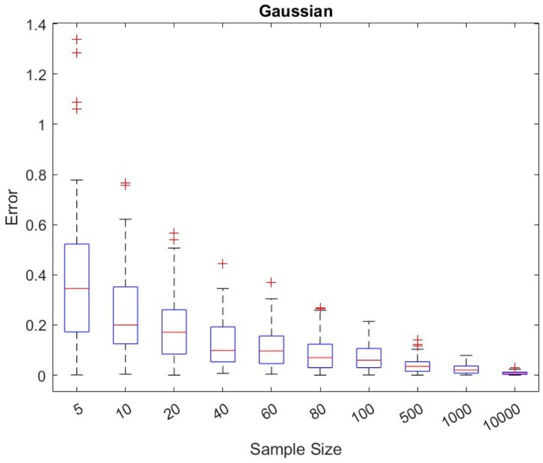


Figure 23: Errors in Gaussian Distribution

6.2 Interpretation of Boxplots

Both the plots show a common trend. As the value of N increases, the distribution of the error, i.e., the distribution of $|\hat{\mu} - \mu_{true}|$:

1. has a “narrower” spread.
2. has median and quartiles all tending to 0

From these observations we can conclude the following :

1. The Law of Large Numbers holds, i.e., the variance of the average of N random variables tends to 0 as $N \to \infty$
2. The expectation of the average of N random variables tends to the true mean as $N \to \infty$

Hence, we have completed the plotting and analysis of errors in both the cases.

7 References

[1] Class Slides, *autumn semester*, Suyash P. Awate, IIT Bombay, 2022.

[2] prob1805, Purdue, 2014.

[3] Steven Lalley, *notes*, University of Chicago.

[4] CS109 , lecture handouts, Stanford University.