ECE 514 Random Process - Project I Report

Author: Nazneen Percy Kotwal ID: 200205655

1. Simulation of Random Variables:

A. Random variables were simulated using Acceptance-Rejection method and the MATLAB Routines:

- Normal with mean = 0 and variance = 1
- Uniform on [0, 1]
- Exponential with parameter 1

MATLAB has in-built routines to generate uniform, normal and exponential variates. The Matlab routines are: rand, randn and exprnd respectively. We can also generate these variates using a technique known as the Acceptance-Rejection method. The Normal variates for this project were generated by using a uniform distribution as the envelope distribution. The magnitude of uniform distribution is kept at $1/\sqrt{2\pi}$ which is the maximum value of standard normal probability density function $N(0,1)$. Exponential and uniform distribution are generated in a similar way by ensuring the envelope distribution completely encompasses the distribution function to be estimated fast by suitably scaling it, but kept close enough for the algorithm to converge. For each of the three distributions, random variables are generated in sets of $T = 100,\ T = 1000\ and\ T = 10000$. Histograms obtained by Acceptance-Rejection method and MATLAB routines are shown in Figure 1- 9 below.

1. Normal with mean = 0 and variance = 1

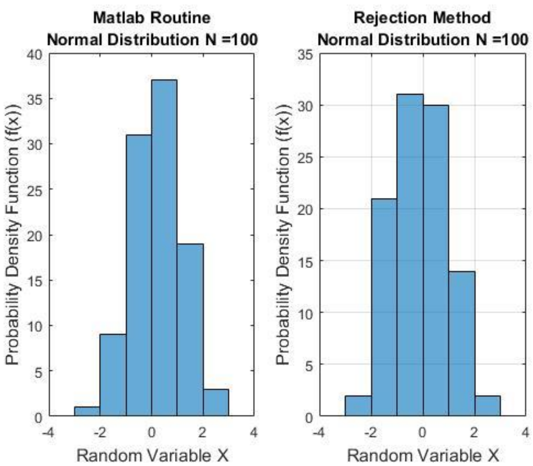

Fig. 1. Histogram for Normal distribution for 100 Variates

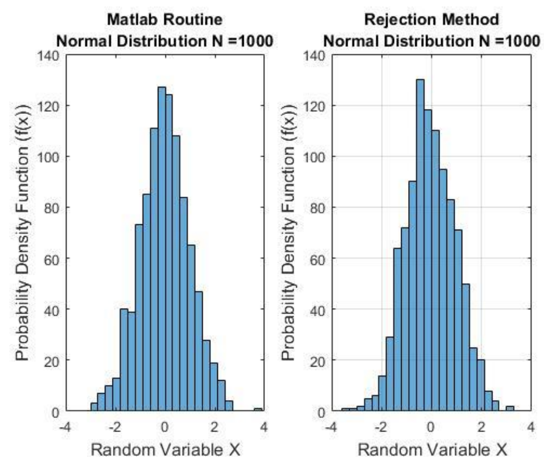

Fig. 2. Histogram for Normal distribution for 1000 Variates

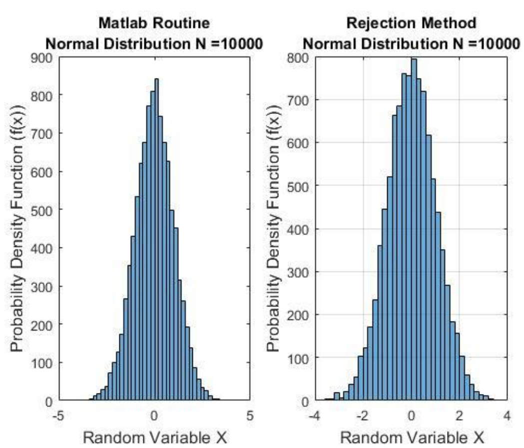

Fig. 3. Histogram for Normal distribution for 10000 Variates

Table 1. Mean of Normal Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	-0.1198	0.0418	-0.1198	0.0418
1000	-0.0161	-0.0178	-0.0161	-0.0178
10000	-0.0057	-0.0121	-0.0057	-0.0121

Table 2. Variance of Normal Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	1.0804	1.0521	0.0804	0.0521
1000	0.9535	0.9530	-0.0465	-0.047
10000	0.9981	0.9817	-0.0019	-0.0183

2. Uniform on [0, 1]

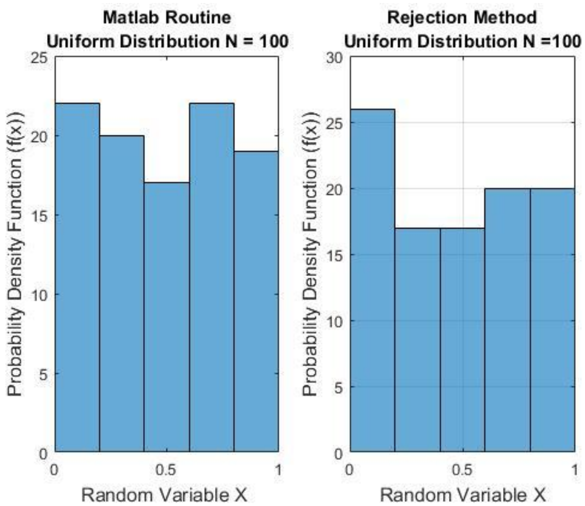

Fig. 4. Histogram for Uniform distribution for 100 Variates

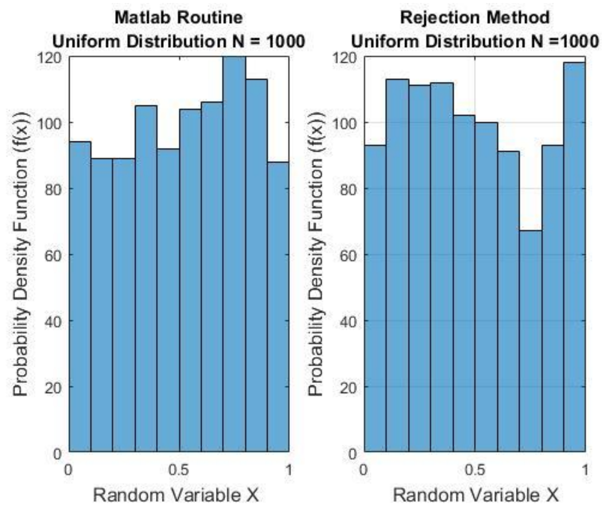

Fig. 5. Histogram for Uniform distribution for 1000 Variates

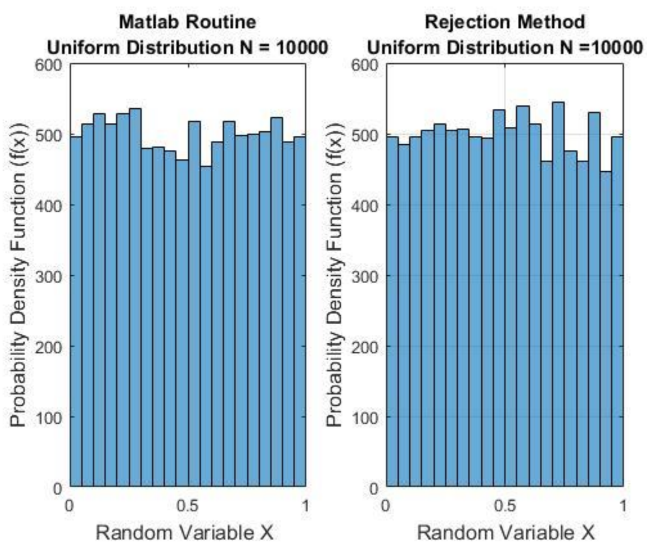

Fig. 6. Histogram for Uniform distribution for 10000 Variates

Table 3. Mean of Uniform Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	0.4742	0.5740	-0.0258	0.074
1000	0.4947	0.5072	-0.0053	0.0072
10000	0.4978	0.4983	-0.0022	-0.0017

Table 4. Variance of Uniform Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	0.0951	0.0841	0.01177	0.00077
1000	0.0817	0.0828	-0.00163	-0.00053
10000	0.0849	0.0834	0.00157	7E-05

3. Exponential with parameter 1

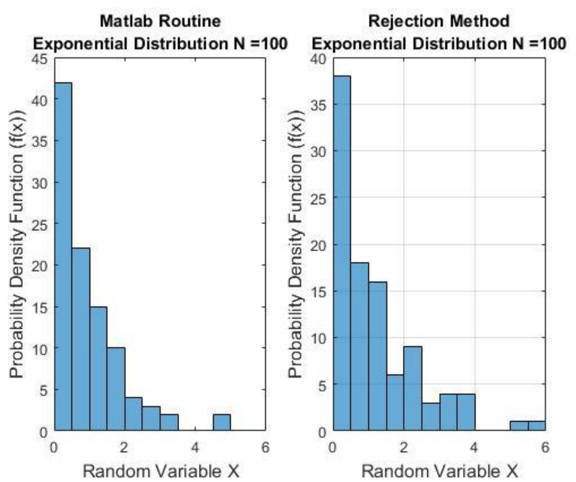

Fig. 7. Histogram for Exponential distribution for 100 Variates

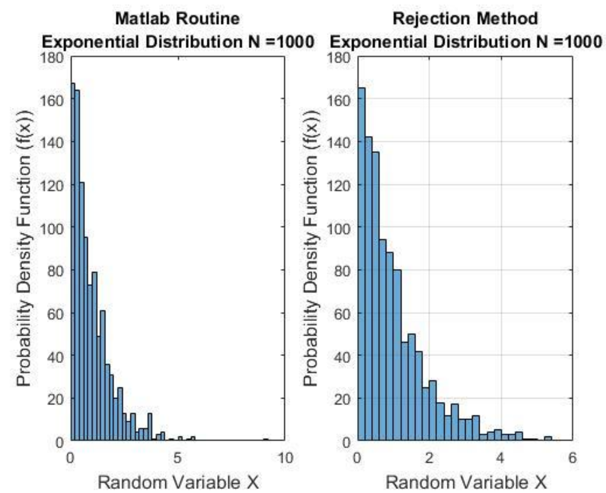

Fig. 8. Histogram for Exponential distribution for 1000 Variates

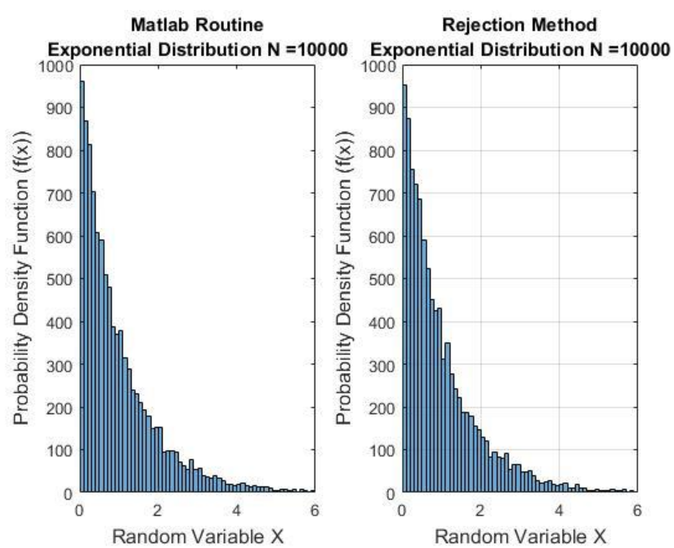

Fig. 9. Histogram for Exponential distribution for 10000 Variates

Table 5. Mean of Exponential Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	1.1755	1.0349	0.1755	0.0349
1000	1.0169	1.0098	0.0169	0.0098
10000	0.9769	1.0066	-0.0231	0.0066

Table 6. Variance of Exponential Distribution

Number of Variates (T)	Rejection Method	MATLAB Routine	Deviation in Rejection Method	Deviation in MATLAB Routine
100	0.8152	1.0810	-0.1848	0.0810
1000	0.9873	0.9368	-0.0127	-0.0632
10000	0.9144	1.0296	-0.0856	0.0296

B. Computing Mean and Variance

Mean and Variance is calculated for the Normal, Uniform and Exponential distributions with different number (100, 1000 and 10000) of variates. The results are expressed in a tabular manner above. Table 1-6 shows the mean as well as the variance of the random variates. Rows represent the number of samples whose mean/variance is taken, and column represents the method of sample generation.

Theoretical values of mean and variance for different random variables under consideration are as follows:

- Normal random variable: Mean = 0 and variance = 1
- Uniform random variable: Mean = 0:5 and variance = $\frac{1}{12} = 0.08333$
- Exponential random variable: Mean = 1 and variance = 1

C. Comparison with Theoretical Values:

For each type of distribution and method of variate generation, namely Matlab Routines and Acceptance-Rejection Method, simulation is carried out three times, corresponding to three different number of samples: 100, 1000 and 10000 variates.

Comparison of Mean:

From table 1, we can see that mean of normal random variates is close to the theoretical value. Maximum absolute deviation observed is only 0.1198 from the theoretical value of zero mean. Moreover, as the number of samples increases from 100 to 10,000, the approximation is improved for both methods of generation. For 10,000 variates, the maximum deviation observed is only 0.0057 which is approximately twenty times better than the deviation from the mean observed for 100 variates. This observation is consistent with the law of large number.

Similar trend is observed in case of uniform and exponentially distributed samples. From table 3 and 5 we see that the worst-case deviation for uniformly distributed variates is 0.1755 and for exponentially distributed variates is 0.1848. This deviation or error is reduced when sample size is increased. For 10,000 samples, the worst-case error or deviation in uniformly distributed samples is 0.0022 and for exponential it is 0.0066.

Comparison of Variance:

For the same set of samples as above, variance is calculated and compared with theoretical value of the distribution. We can see from Table 2 that the maximum error in variance of normally distributed samples is 0.0804. This error corresponds to the case of 100 variate. As the number of samples is increased, the deviation from the theoretical value of the mean decreases. We see that the maximum error corresponding to 10,000 variates in this case is only 0.0019. Similar trend is observed for uniform and exponentially distributed samples. For uniformly distributed samples, the worst-case deviation is 0.01177 which occurs for 100 variates. This error reduces to 0.00157 when number of variates are 10, 000. Worst case error for exponentially distributed sample is 0.1848 which is for the case of 100 variates. This value reduces to 0.08856 when number of samples is increased to 10,000.

D. Possible Reasons for Deviation of the Mean and Variance from the Theoretical Values:

These variations we observed in the above cases is due to the number of samples taken into consideration. We can see that as n tends to infinitely many samples, the mean and variance approaches values that are close to the theoretical value.

The deviation can also be due to errors in the sample generation procedures. The Acceptance-Rejection method accepts a sample by generating a uniformly generated random number and comparing this with the ratio of the sample's value at $f(x)$, the function to be estimated and C g(x) , the envelope distribution $g(x)$ scaled by factor C so that C g(x) encompasses the envelope of $f(x)$. Because the acceptance of a sample is based on a random number generation, there exist possibility, that a wrong sample is accepted. Hence causing the deviation in estimated parameters. As the number of samples increases, the contribution by the wrongly accepted samples reduce and so the approximation is improved.

2. Transforming Random Variables

We define a random variable Y such that $Y_i = \frac{\sum_i Xi}{T}$ $i = 1 \ldots T$. In the above section we generated samples that are Normal, Uniform and Exponentially distributed. Moreover, the sample size was changed in three steps, i.e, $T = 100, T = 1000 \; and \; T = 10{,}000$. Hence, we have three different population for random variable Y, corresponding to each method and sample size. The nine histograms corresponding to Y for the three-different distribution and sample size are shown in Fig. 10-12

The closest pdf that matches the distribution of Y for the case of all the above nine cases is the Normal distribution. The mean of the distribution corresponds to the mean of the distribution of X that was used to generate the samples of Y and variance of the Normal distribution is equal to $\sigma\sqrt{n}$, where σ is the standard deviation for the distribution of X.

It is observed that the results obtained in this section align with the Central Limit Theorem which states that as the sample size increases, the sampling distribution of the mean, Y, can be approximated by a normal distribution with mean μ and standard deviation $\sigma\sqrt{n}$ where: μ is the population mean, σ is the population standard deviation n is the sample size

In other words, if we repeatedly take independent random samples of size n from any population, then when n is large, the distribution of the sample means will approach a normal distribution.

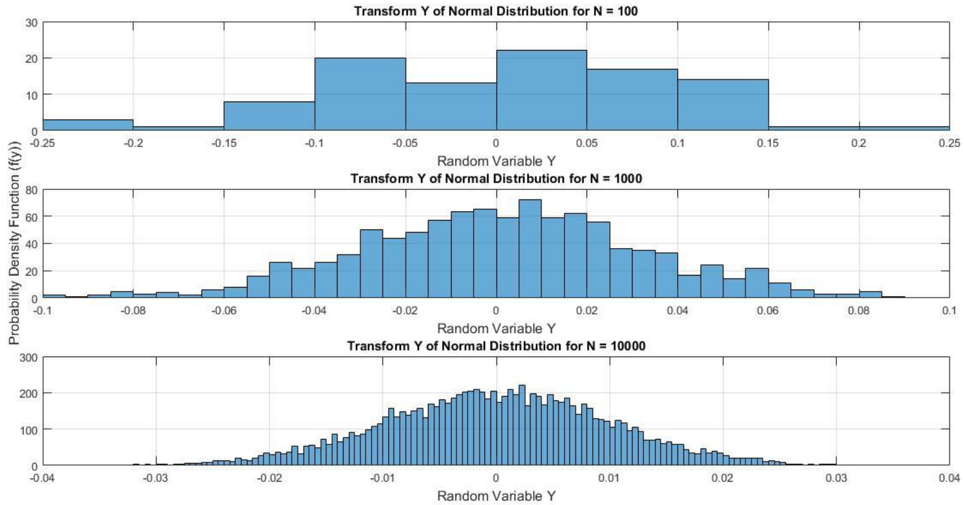

Fig. 10: Histogram for distribution of Y from Normal variates for 100, 1000 and 10000 samples

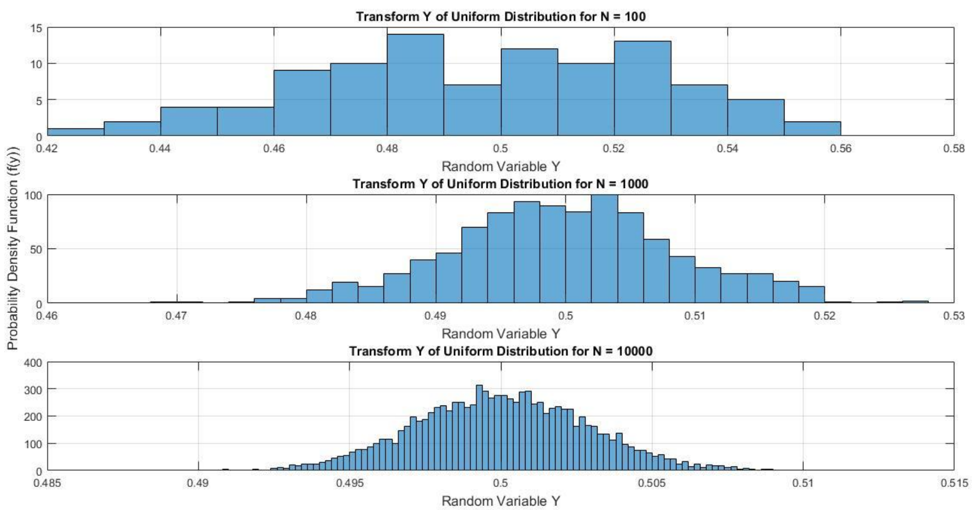

Fig. 11: Histogram for distribution of Y from Uniform variates for 100, 1000 and 10000 samples

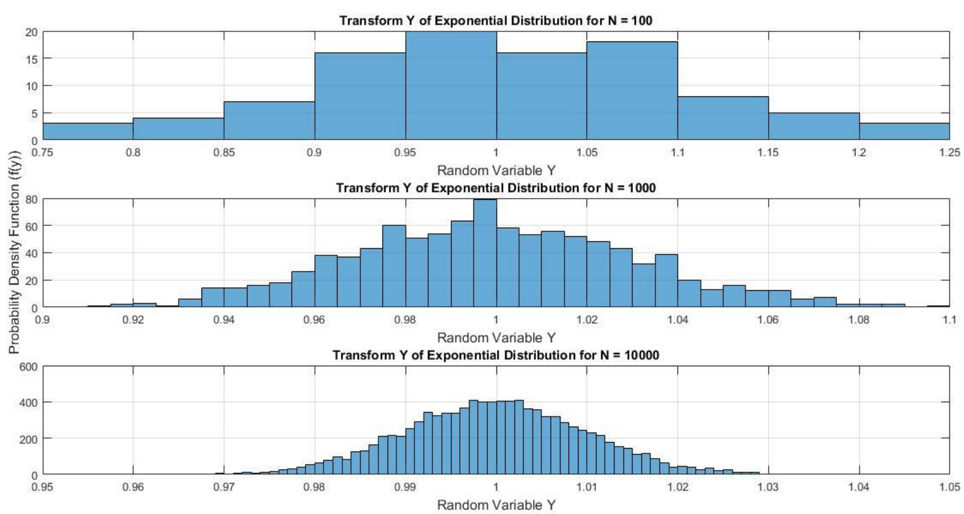

Fig. 12: Histogram for distribution of Y from Exponential variates for 100, 1000 and 10000 samples

3. Convergence of Random Variables

Based on the paper "Understanding Convergence Concepts: A Visual-Minded and Graphical Simulation-Based Approach", a demo using MATLAB is generated to answer the following questions:

1. $Y_T \xrightarrow{P} 0$
2. $Y_T \xrightarrow{A.S} 0$
3. $Y_T \xrightarrow{M.S} 0$
4. $Y_T \xrightarrow{L} X$

As the in the above sections, this demo is created for proving convergence of Y generate using three different distribution schemes namely, Normal, Uniform and Exponential. Following the paper closely the sample size T for the distribution of Y is chosen to be a value of $T = 2000$ and the number of realizations is chosen to be $M = 500$. The convergence of the distribution of Y is consistent with the law of large numbers which states that for sufficiently large n of random variables; the normalized mean Mn with a high probability takes values close to the mean of the distribution from which you sample.

It is observed that in the case of normal distribution Y converge to the mean value of the distribution of X, i.e $E[X] = 0$. It should be noted that in the case of uniform and exponential distributions $Y_T \xrightarrow{P} E[X] = 0.5$ and $Y_T \xrightarrow{P} E[X] = 1$ respectively. Therefore, to study the convergence in probability of a random variable Y_T to $E[X]$, we can define the random variable $Z_T = Y_T - E[X]$ and study the convergence in probability of Z_T to the constant 0. This remark is also valid for almost sure convergence and convergence in r^{th} mean.

The screenshot of the GUI display in Fig. 13 depicts the different convergence modes of the distribution Y generated using Normal samples. The In-Probability convergence (blue curve), the Almost Sure convergence (red curve) and the Mean Square convergence to the value zero is clearly observed in the centered graph. The right most graph depicts the convergence in distribution of Y to the normal distribution which has a mean value of zero. This is consistent with the expected result. The M = 500 realizations for Y from Normal variates is plotted in the first plot shown in Fig. 13.

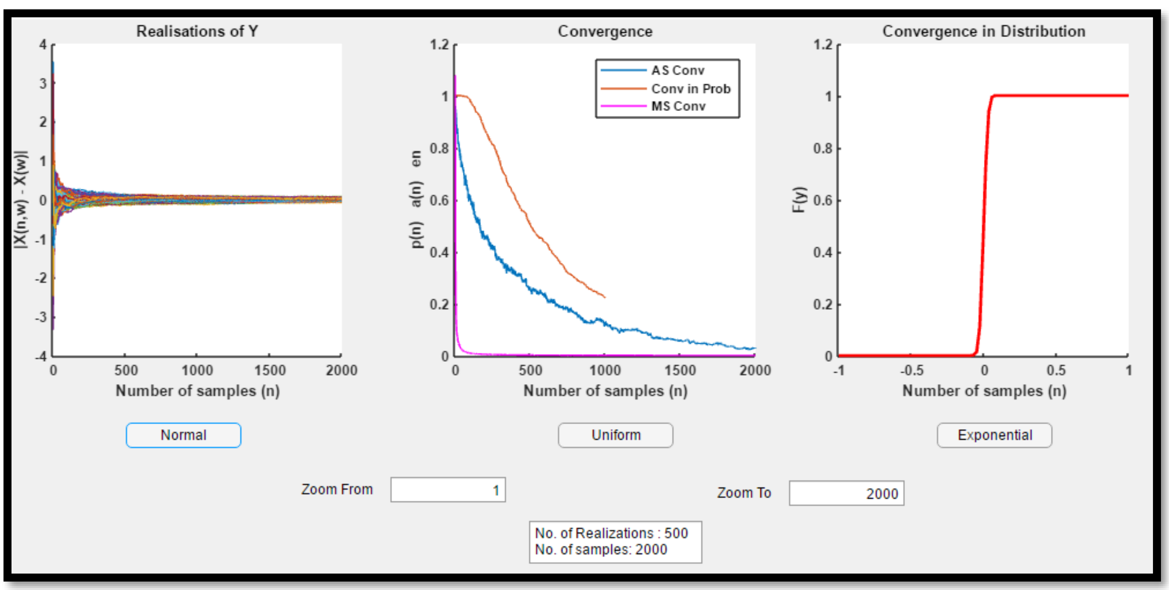

Fig. 13: Screen-shot of GUI to establish different Convergence Modes of Y generated from Normal variates.

The screenshot of the GUI display in Fig. 14 depicts the different convergence modes of the distribution Y generated using Uniform samples. It should be noted that as mentioned before, the In-Probability convergence (blue curve), the Almost Sure convergence (red curve) and the Mean Square convergence (green curve) depicted in the centered graph is for $Z_T = Y_T - E[X]$ where $E[X] = 0.5$. The convergence of Z_T to the constant 0 is observed. The right most graph depicts the convergence in distribution of Y to the normal distribution which has a mean value of $E[X] = 0.5$. The M = 500 realizations for Y from Uniform variates is plotted in the first plot shown in Fig. 14. It is clearly observed that Y_T converges to the value 0.5, which in turn implies the convergence of $Z_T = Y_T - 0.5$ to 0.

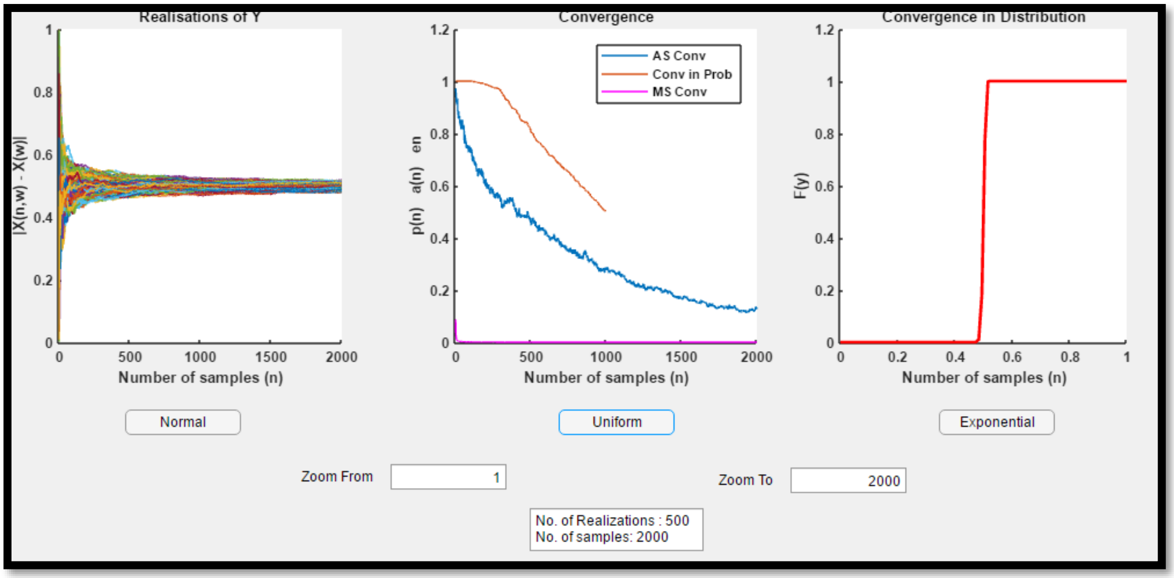

Fig. 14: Screen-shot of GUI to establish different Convergence Modes of Y generated from Uniform variates.

The screenshot of the GUI display in Fig. 15 depicts the different convergence modes of the distribution Y generated using Exponential samples. It should be noted that as mentioned before, the In-Probability convergence (blue curve), the Almost Sure convergence (red curve) and the Mean Square convergence (green curve) depicted in the centered graph is for $Z_T = Y_T - E[X]$ where $E[X] = 1$. The convergence of Z_T to the constant 0 is observed. The right most graph depicts the convergence in distribution of Y to the normal distribution which has a mean value of $E[X] = 1$. The M = 500 realizations for Y from Exponential variates is plotted in the first plot shown in Fig. 15. It is clearly observed that Y_T converges to the value 1, which in turn implies the convergence of $Z_T = Y_T - 1$ to 0.

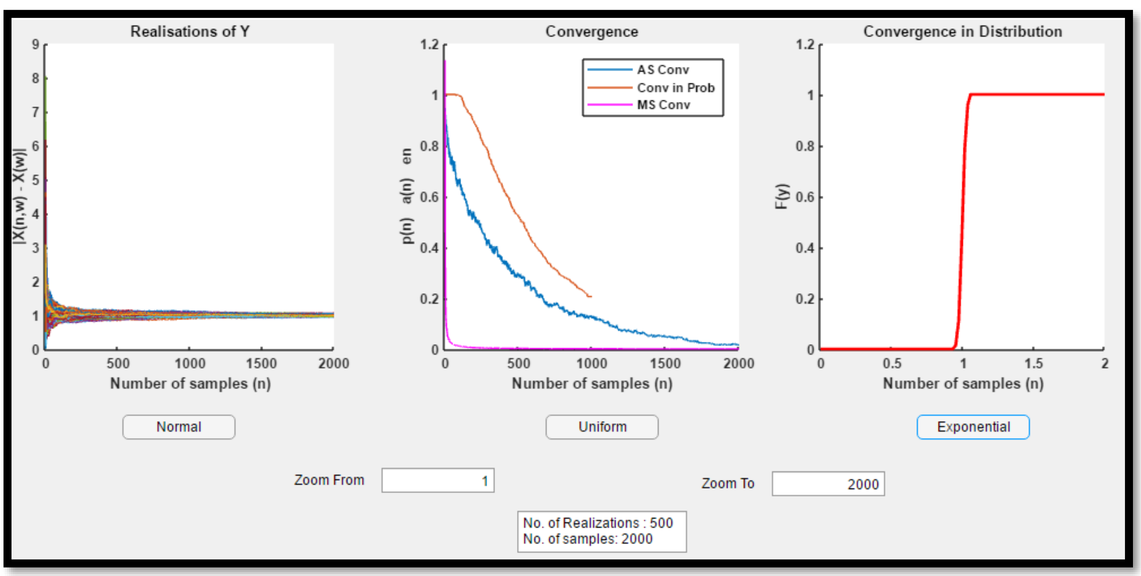

Fig. 15: Screen-shot of GUI to establish different Convergence Modes of Y generated from Exponential variates.