Name: Gunjan Jhawar

Subject code: EE 511

Simulation of Scholastic System

USCID: 1857534264

Project: 7

Aim:

Consider a continuous random variable defined by its probability density function, $f(x) = C \sin(x)$ where C is a normalizing constant.

A) Find the distribution function and evaluate C. Compute (analytically) the mean, μ , and variance, $2\ \sigma$.

B) By inverting the distribution function, you can use the rand () function in Matlab to directly generate samples of X. Generate a set of N samples (for N = 25,100,1000). For each case plot the empirical distribution generated by these samples and compare to the true underlying distribution function.

C) For the case of N =100, find the sample mean, m. Use the population variance to find the MSE of the mean estimate. Now suppose the population variance was not known, compute the MSE of the mean estimate based on the sample variance 2 s. How do they compare?

D) For the case of N =100, use the bootstrapping technique to generate M bootstrap samples based on the empirical distribution found in part B and compute the mean and sample variance for each Bootstrap sample. Use M = 50. (If you have time you can try some other values of M)

E) We could try to find the MSE of the sample variance estimate based on the population variance If F doesn't have simple structure, this seems a bit challenging. So, instead, we find an approximation using the bootstrap approach. We can calculate the (population) variance of the empirical distribution.

Software Used:

MATLAB

Procedure:

A. $\int_0^{\pi} f(x)$ dx = 1 so c = ½.

F(x) = $\int_0^{\pi} f(t)$ dt = ½ (1- cosx) ,

μ = π/2 = 1.5708,

σ^2= ½(π^2-4) = 0.4674

B. Using the method of inverting the distribution function, rand() function is used to generate random variables x and y. Then x = arcos (1 – 2y) to directly generate samples of x. Then we compare the theoretical result with generated plot.

C. We know the population variance; the sample mean is 1/n $\sum X$. Expected value can be calculated using E [1/n $\sum X$.] = nμ/n = μ. Similarly, the variance can be calculated as σ^2/n.

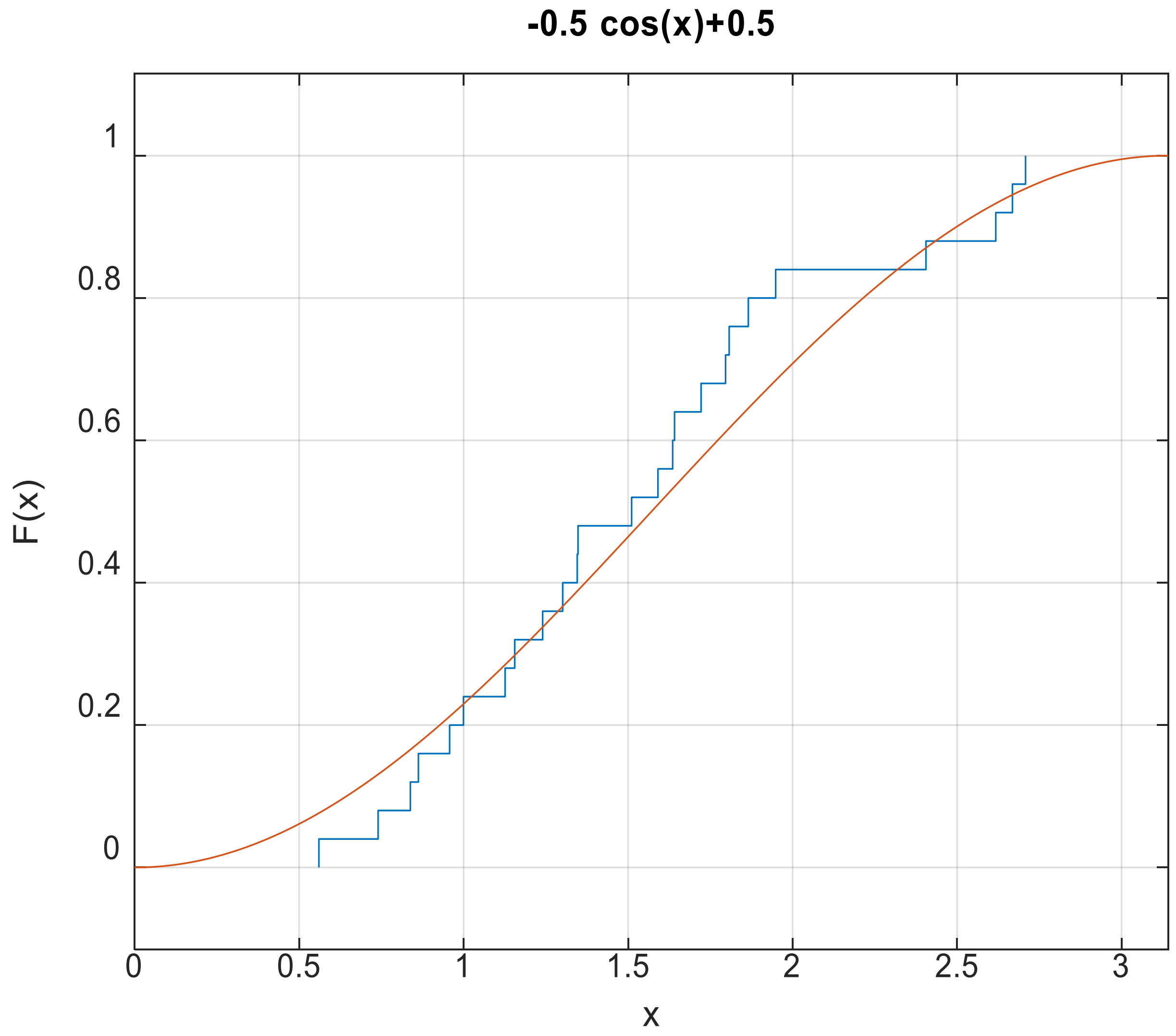
D. We randomly choose 50 samples from the random variable numbers generated before and then calculate the sample mean and sample variance of these 50 samples.

E. The sample randomly chosen in D and the variance, we use MSE equation to estimate mean square estimate based on the population variance.

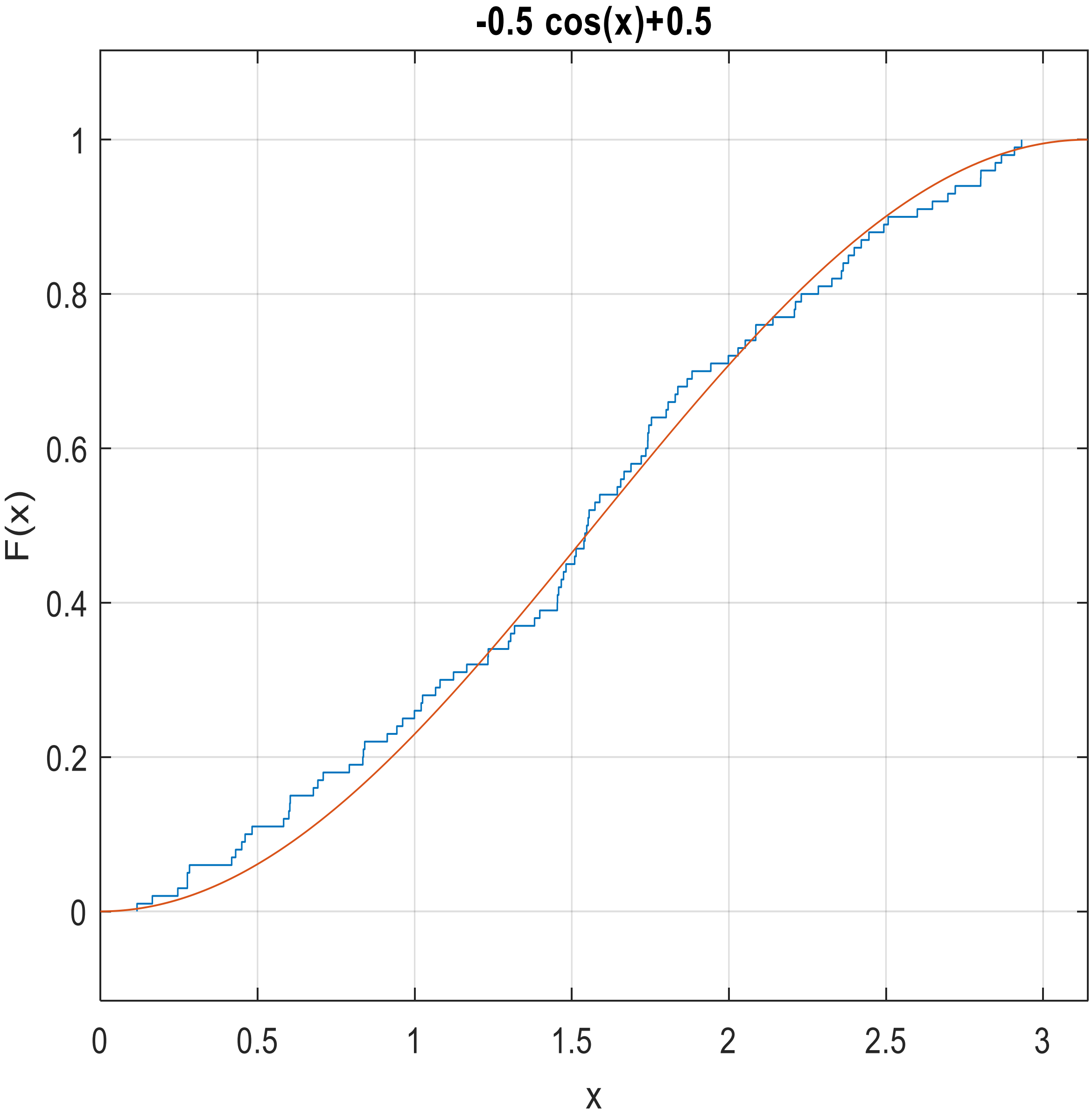
Observation and Analysis:

1. We calculate mean = 1.5708 and var = 0.4675

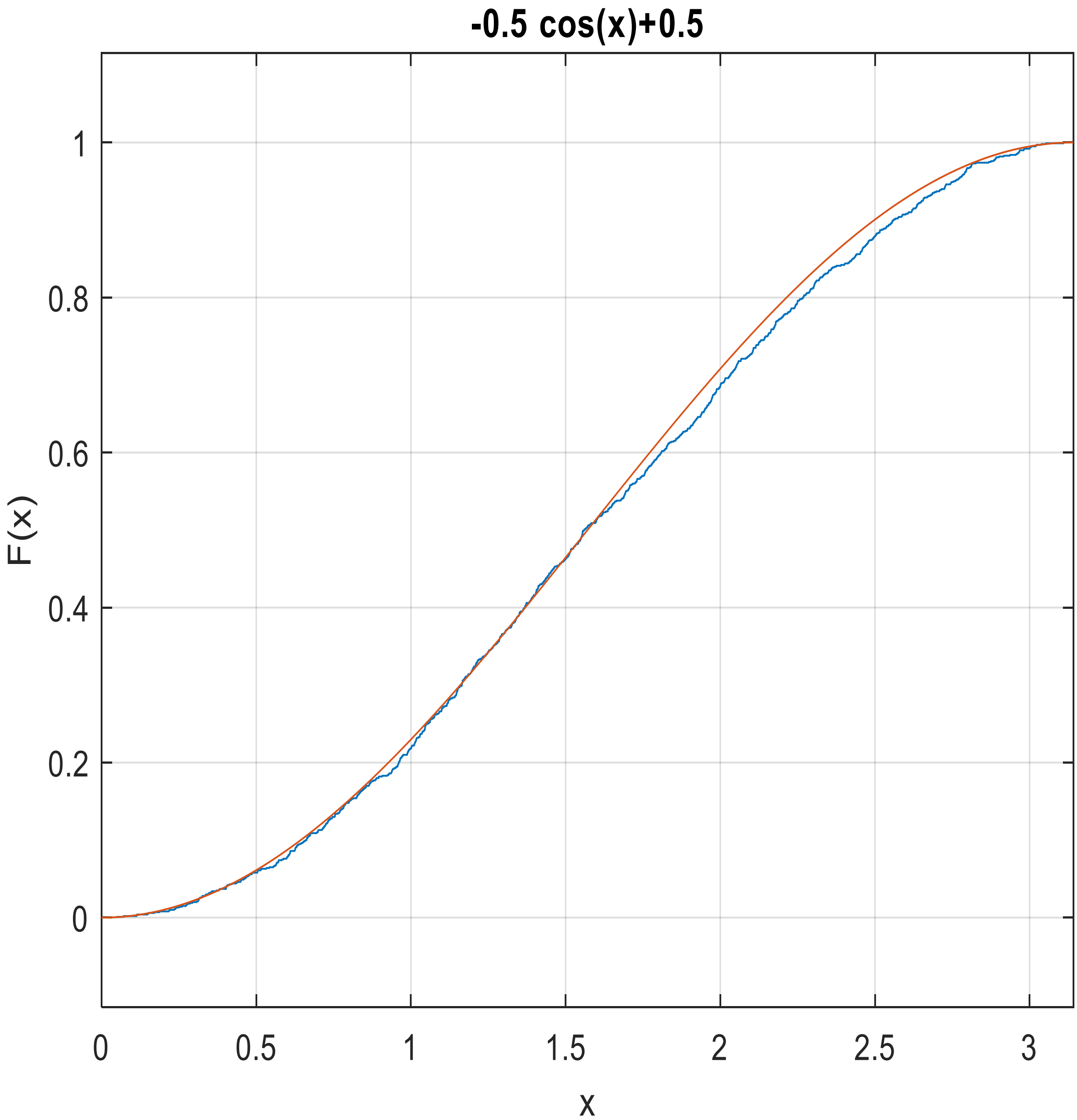
2.
N = 25
Mean = 1.8241
Var = 0.4684



N = 100
Mean = 1.6028
Var = 0.4696



N = 1000
Mean = 1.6070
Var = 0.4555



As sample size increases. The cumulative distribution is closer to the theoretical curve. The estimated mean and variance is also closer to the theoretical value.

3. MSE = 0.0048 and theoretical value based on population variance equals to 0.00467, they are very close.

4. The samples mean and variance are found used boostrapping method:
M = 50.

Variables - mean_sample

mean_sample

10x5 double

	1	2	3	4	5
1	1.6915	1.5352	1.5828	1.4808	1.6818
2	1.6759	1.5132	1.5460	1.6776	1.5107
3	1.6015	1.5877	1.6466	1.6198	1.5945
4	1.5758	1.5805	1.6178	1.4762	1.5001
5	1.4983	1.5852	1.6299	1.5245	1.5627
6	1.5886	1.5240	1.5826	1.5993	1.6587
7	1.4864	1.6286	1.6182	1.4012	1.6173
8	1.5815	1.6855	1.4934	1.5576	1.5382
9	1.5775	1.5755	1.6300	1.6704	1.5085
10	1.7231	1.5179	1.5981	1.3843	1.4966
11					

var_sample

10x5 double

	1	2	3	4	5
1	0.5731	0.4055	0.4062	0.4657	0.5028
2	0.5270	0.4234	0.5060	0.4310	0.4677
3	0.4625	0.4392	0.4902	0.4778	0.4182
4	0.4999	0.5168	0.4827	0.5124	0.3384
5	0.4508	0.4077	0.4023	0.4019	0.4778
6	0.3790	0.4859	0.4952	0.5576	0.5176
7	0.4001	0.4550	0.4319	0.3912	0.4726
8	0.4042	0.4999	0.5065	0.4512	0.4403
9	0.4792	0.4728	0.5198	0.4402	0.3418
10	0.4998	0.4139	0.4444	0.3793	0.4666
11					
12					

5. The result in part c is done used method of boostrapping is valid and does a good estimation. MSE = 0.0044.