Project 0: Coin Tossing Experiment

EE 511 – Section: Tuesday 5pm

Name: Darshan Patil

Student ID: 9575227834

1.

i. Problem Statement

Let $X \sim U(0,1)$, evaluate the mean, μ, and variance, σ_X^2.

ii. Theoretical Exploration

Generally,

Mean or expected value

$$\mu = E[X] = \begin{cases} \sum_{x=0}^{x} x p_x(x) & \mathit{for\ \ a\ discrete\ RV} \\ \int_0^x x f_x(x) dx & \mathit{for\ a\ continuous\ RV} \end{cases}$$

Variance

$$VAR(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

For $X \sim U\ (0,1)$, we have,

$$\mu = E[X] = \int_0^1 x f_x(x) dx = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

$$VAR(X) = \sigma^2 = E[X^2] - E[X]^2 = \int_0^1 x^2 f_x(x) dx - \left[\frac{1}{2}\right]^2 = \frac{x^3}{3} \Big|_0^1 - \frac{1}{4} = \frac{1}{12}$$

$$\sigma = \sqrt{VAR(X)} = \sqrt{\frac{1}{12}} = 0.2887$$

iii. Simulation Methodology

Simulation steps:

- Start the Trial X ~ U (0,1)
- Assign Value 1 to function $f(x)$
- Define function to be integrated to calculate mean i.e. $x * f(x)$
- Calculate mean which is equal to integral of $x * f(x)$ where $x = 0\ to\ x = 1$

- Define function to be integrated to calculate Variance i.e. $x^2 * f(x)$
- Calculate Variance which is equal to (integral of $x^2 * f(x)\ where\ x = 0\ to\ x = 1) - (0.5)^2$ i.e. $Variance = E[X^2] - E[X]^2$
- Calculate Standard Deviation which is equal to root of Variance.

iv. Results

Mean: 0.500000
Variance: 0.083333
Standard Deviation: 0.288675

Discussion of Results
We have successfully calculated the mean, variance, and standard deviation for $X{\sim}U\ (0{,}1)$

v. References

- https://www.tutorialspoint.com/matlab/
- https://www.mathworks.com/help/matlab/ref/integral.html#btdd9x5

vi. Source Code

```
%Start of the Programming code
clear all;                          % clear and close matlab desktop, workspace
close all;
clc;
F = 1;                              % function f(x) equal to 1
fun1=@(x)(x*F);                     % function to integrate x * f(x)
Mean=integral(fun1,0,1);            % Calculate mean

fun2=@(x)(x.^2*F);                  % function to integrate x^2*f(x)
Variance=integral(fun2,0,1)-(0.5).^2; % Calculating Variance=E[x^2]-E[x]^2

SD=sqrt(Variance);                  % Calculating Standard Deviation

fprintf('Mean: %f\nVariance:%f\nStandard Deviation:%f\n\n', Mean, Variance, SD);        % Display
```

2.

i. Problem Statement

Generate a sequence of N=100 random numbers between [0,1] and compute the sample mean $m = \frac{1}{N}\sum_{i=0}^{N} X_i$ and sample variance $s^2 = \frac{\sum_{i=1}^{N}(X_i - m)^2}{N-1}$ and compare to μ and σ^2. Also estimate the (sample) variance of the sample mean (based on the Central Limit Theorem). Repeat for $N = 10{,}000$.

ii. Theoretical Exploration

Generally,

Mean or expected value

$$\mu = E[X] = \begin{cases} \sum_{x=0}^{x} x p_x(x) & for \ \ a \ discrete \ RV \\ \int_0^x x f_x(x) dx & for \ a \ continuous \ RV \end{cases}$$

Variance

$$VAR(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

For $X \sim U\ (0,1)$, we have,

$$\mu = E[X] = \int_0^1 x f_x(x) dx = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

$$VAR(X) = \sigma^2 = E[X^2] - E[X]^2 = \int_0^1 x^2 f_x(x) dx - \left[\frac{1}{2}\right]^2 = \frac{x^3}{3}\Big|_0^1 - \frac{1}{4} = \frac{1}{12}$$

$$\sigma = \sqrt{VAR(X)} = \sqrt{\frac{1}{12}} = 0.2887$$

When we run a simulation, we collect a series of N observations $\{X_i\}$ (typically the values $\{X_i\}$ are independent identically distributed and estimate these statistics using data computed from the sample data.

Sample mean $m \ or \ \bar{X}$: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{N}$

Sample variance s^2: $s^2 = \frac{\sum (X_i - m)^2}{N-1}$

Standard Deviation s: $s = \sqrt{s^2}$

iii. Simulation Methodology

Simulation steps:

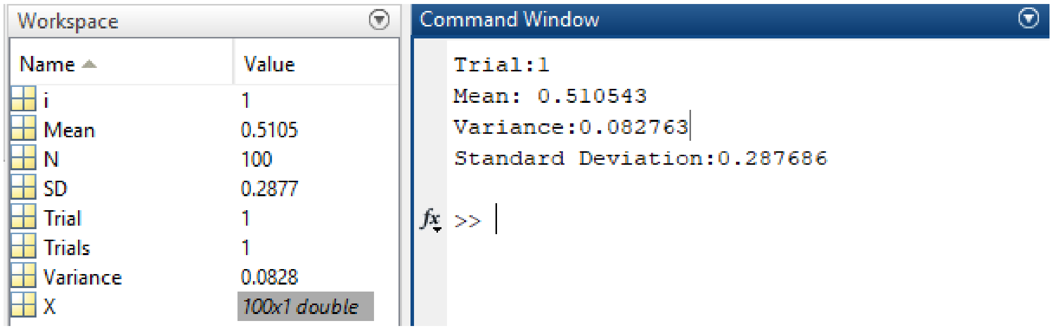
1. Assign the number of Random Number to be generated to N i.e.100
2. Assign the Number of Trial to be performed. i.e. Trials = 1
3. Generate 100 random number using the MATLAB built-in function, rand (N, Trials)
4. Store the Trial numbers in the matrix
5. Calculate Mean using the MATLAB built-in function, mean(X,1)
6. Calculate Standard Deviation using the MATLAB built-in function, std (X,1)
7. Calculate Variance which is equal to square of Standard Deviation
8. Display the Sample Mean, Sample Variance, and Standard Deviation

iv. Experiments and Results

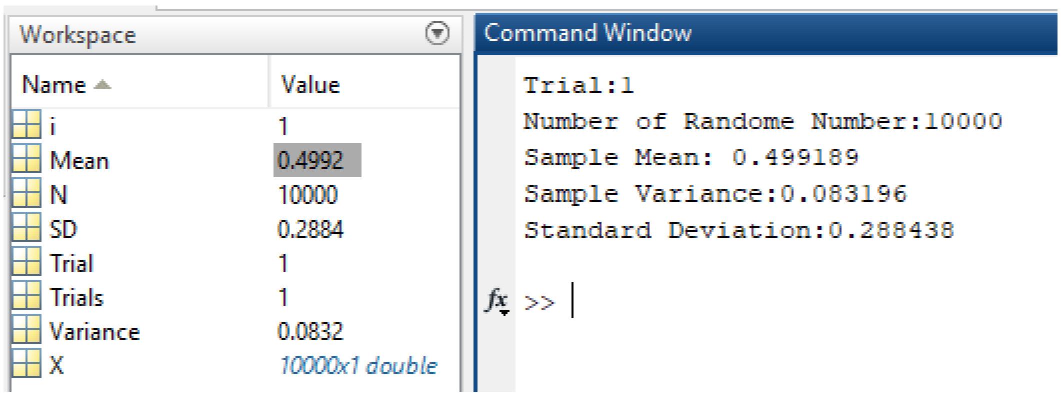
For N=100,

We got the below random number's,

0.6181	0.609802	0.233653	0.325806	0.731407
0.932183	0.166891	0.456425	0.095949	0.781374
0.835088	0.188092	0.384567	0.747534	0.367286
0.895424	0.094629	0.538601	0.748509	0.744868
0.582519	0.323186	0.991704	0.543299	0.892267
0.582747	0.769597	0.75522	0.338132	0.242603
0.854926	0.234118	0.980455	0.832334	0.129597
0.034866	0.740365	0.234783	0.552572	0.225068
0.88542	0.692818	0.528559	0.957543	0.350014
0.407731	0.824078	0.051436	0.892833	0.287085
0.036382	0.827978	0.756875	0.356504	0.927488
0.746148	0.293368	0.60198	0.546402	0.051314
0.154829	0.309369	0.857169	0.346682	0.592667
0.143908	0.52303	0.988277	0.622803	0.162899
0.605959	0.325299	0.929484	0.796625	0.838406
0.254481	0.831843	0.409515	0.745875	0.167561
0.324154	0.810295	0.000341	0.125536	0.502201
0.401791	0.556998	0.540878	0.822394	0.999329
0.406373	0.262964	0.207731	0.025151	0.355407
0.386191	0.680566	0.219284	0.414429	0.047078



For N= 10,000



Therefore,

We calculated – Sample Mean $m = 0.510543$, Sample Variance $s^2 = 0.082763$

We have – $\mu = 0.5$ *and* $\sigma^2 = 0.08333$

Discussion of results

We can observe that the Sample Mean m and Sample Variance s^2 closely matches with the theoretical values μ and σ^2 .

v. References

- https://www.tutorialspoint.com/matlab/
- https://www.mathworks.com/help/matlab/

vi. Source Code

```
%start of the programming code
clear all;              % clear and close matlab desktop, workspace
clc;
N = 100;                % Assigning sequence of random number
Trials = 5;             % Assigning Number of Trial
X= rand(N,Trials);      % Sequence of 100 random numbers between 0 to 1
Trial = find(X,Trials); % Storing the Trial number
Trial = Trial';         % Converting Trial number Matrix from Row to
                          Columns
Mean = mean(X,1);       % Calculate mean using Matlab built-in
                          function mean
SD = std(X,1);          % Calculate Standard Deviation using matlab
                          built-in function std
Variance = SD.^2;       % Calculate variance which is square of
                          Standard Deviation


for i=1:Trials
   fprintf('Trial:%d\nMean: %f\nVariance:%f\nStandard   ...Deviation:%f\n\n',
     Trial(1,i),Mean(1,i), Variance(1,i), SD(1,i)); %Display
end
```

3.

i. Problem Statement

The Central Limit Theorem says that $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{N} \rightarrow N(\mu, \frac{\sigma^2}{n})$.

Repeat the experiment in (2 with $N = 100$) 50 times to generate a set of sample means $\{m_j, j = 1 \dots 50\}$. Do they appear to be approximately normally distributed values with mean μ and variance $\frac{\sigma^2}{n}$?

ii. Theoretical Exploration

$$\mathit{Central\ limit\ theorem}$$

$$As\ n \rightarrow \infty: \left(\frac{\sum_{i=1}^{n} X_i}{n}\right) \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$$

The Central Limit Theorem (CLT) states that the sample mean X of an adequately large sample ($n \rightarrow \infty$) from a discrete RVs with mean m and variance s^2 will follow a normal distribution with parameters m = μ and $s^2 = \sigma^2$.

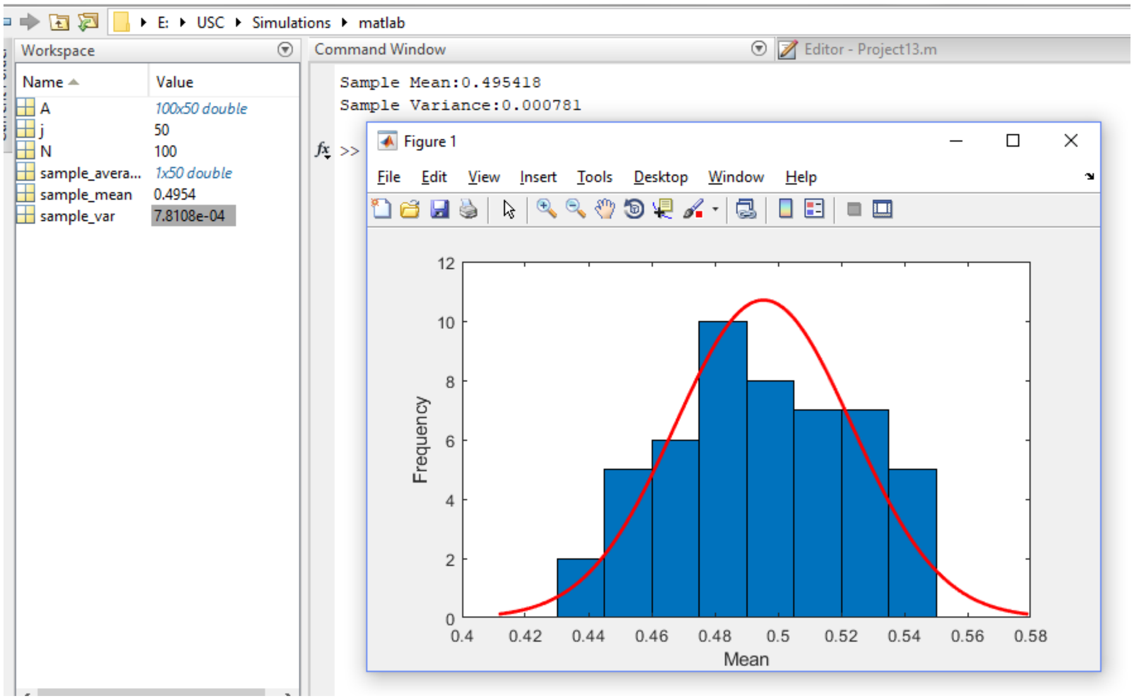
iii. Simulation Methodology

Simulation steps:

1. Assign the number of Random Number to be generated to N i.e.100
2. Assign the Number of Trial to be performed. i.e. j = 50
3. Generate 100 random number using the MATLAB built-in function, rand (N, j)
4. Store column of 1's om sample_averages
5. Calculate Sample Mean by equating (sample_averages * A)/N
6. Calculate Mean of Sample means using built-in Matlab function
7. Calculate Variance using the in-built MATLAB function var(sample_averages);
8. Display the histogram showing Mean vs Frequency Graph
9. Display the Sample Mean and Sample Variance and compare with the theoretical values.

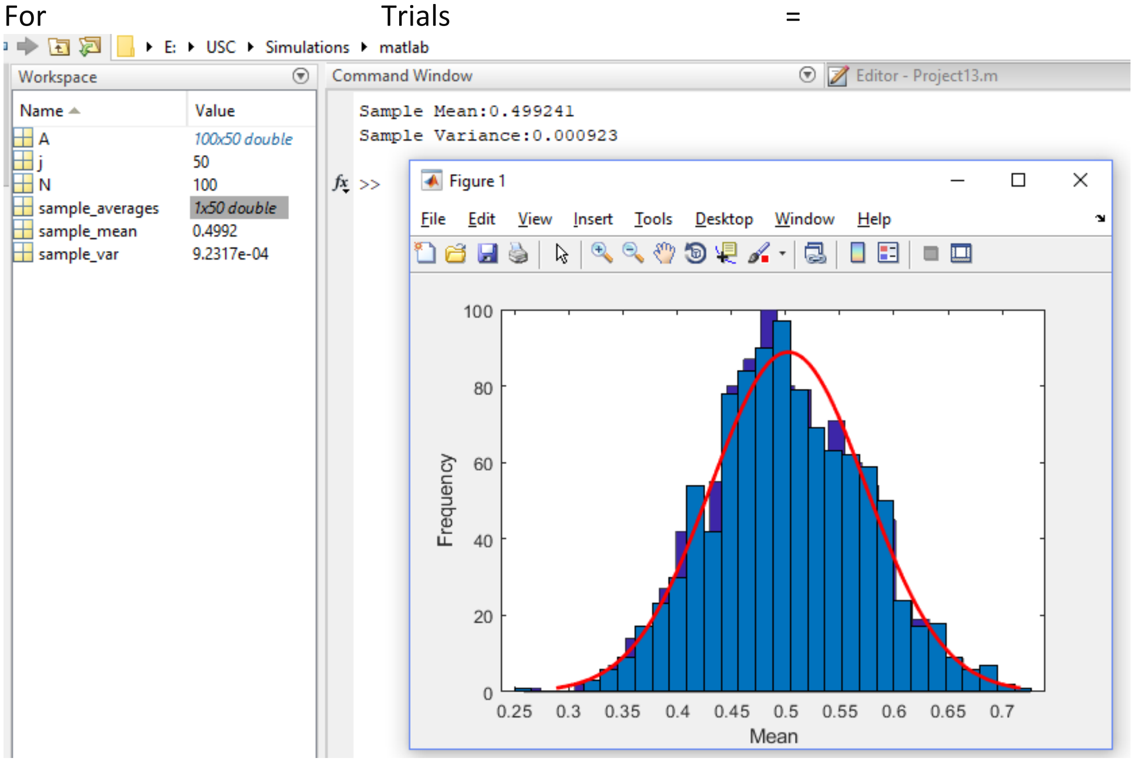
iv. Experiments and Results

For Trials = 50,



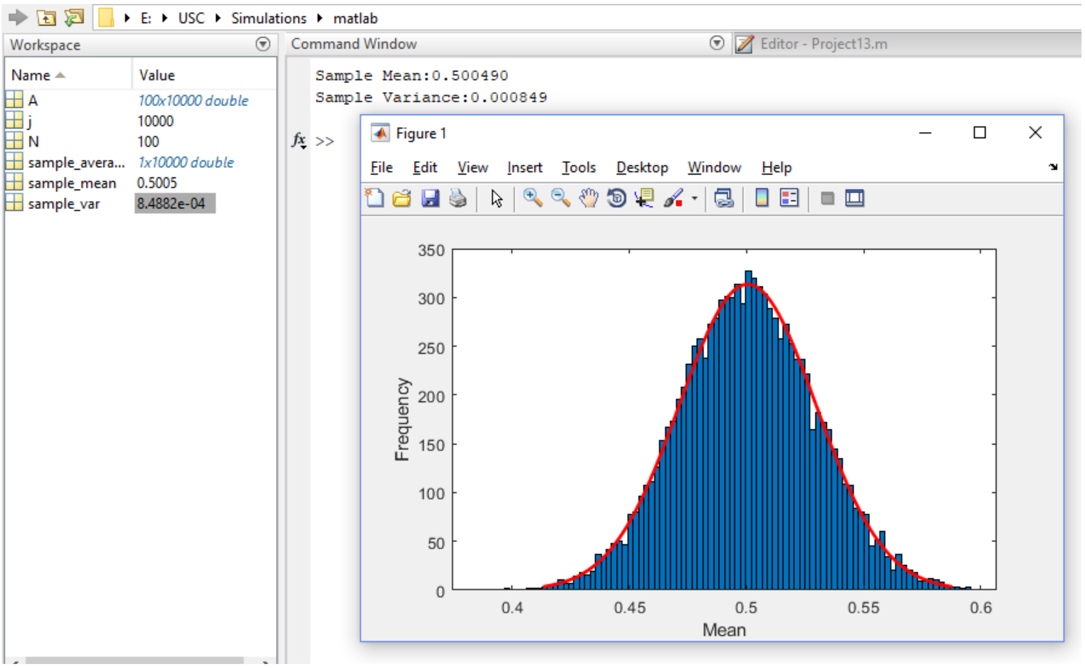
Sample Mean: 0.495418, Sample Variance: 0.000781

For Trials = 500,



Sample Mean: 0.499241, Sample Variance: 0.000923

For Trials = 10,000



Sample Mean:0.500490, Sample Variance: 0.000849

Discussion of Results:

As predicted by Central Limit Theorem, the experimental values i.e. sample mean m, and sample variance closely matched those prescribed by theory. 500 and 10,000 Trials were performed to ensure a sufficiently large sample of sample averages to closely resemble a normal distribution. As we can observe increasing the number of trials would be expected to provide experimental values for sample mean, and sample variance that even more closely matches the theoretical values.

v. References

- https://www.tutorialspoint.com/matlab/
- https://www.mathworks.com/help/matlab/

vi. Source Code

```
%start of the programming code
clear all;                                  % clear and close matlab desktop, workspace
N=100;                                      % Assigning sequence of random number
j=50;                                       % 50 times to generate a set of sample means
A=rand(N,j);                                % generate random variable and store in A
sample_averages = ones(1,N);                % store all one in a row
sample_averages =(sample_averages*A)/N;     % computing average mean for each  trial
sample_mean = mean(sample_averages);        % Mean of all 50 trials
sample_var = var(sample_averages);          % Variance of 50 trials
figure(1);                                  % create a figure
histfit(sample_averages);                   % plot sample means vs frequency
xlabel('Mean');                             % X Label
ylabel('Frequency');                        % Y Label
```

4.

i. Problem Statement

We want to check whether there is any dependency between X_i and X_{i+1}
Generate a sequence of $N+1$ random numbers that are $\sim U\ (0,1)$ for $N = 1{,}000$
Compute

$$Z = \left[\frac{\sum\ X_i X_{i+1}}{N}\right] - \left[\frac{\sum_{i=1}^{N} X_i}{N}\right]\left[\frac{\sum_{j=2}^{N+1} X_j}{N}\right]$$

Comment on what you expect and what you find.

ii. Theoretical Exploration

For a random number generator with a set of data $\{X_i\}$ we might want to look for dependency between samples i.e. we would look at $\{X_i, X_{i+1}\}$ pair.
$\therefore COV(X, X_{i+1}) = E[X_i . X_{i+1}] - E[X_i]E[X_{i+1}]$

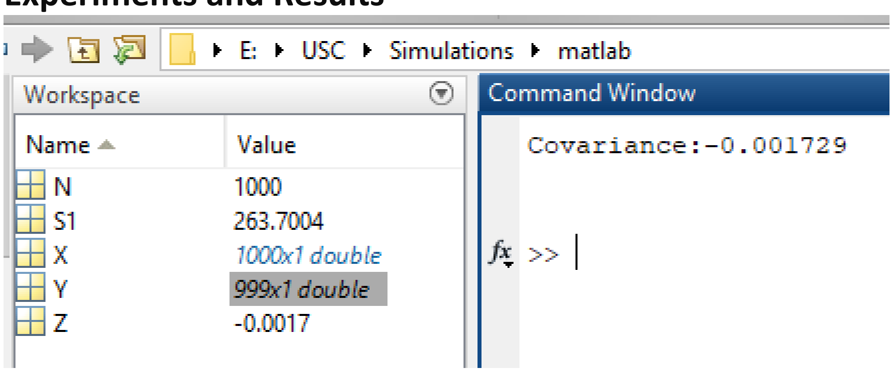
We expect that the value of $Cov(X_i, X_{i+1})$ to be negative since the variable $X_i aand\ X_{i+1}$ are independent.

iii. Simulation Methodology

Simulation steps:

1. Assign the number of Random Number to be generated to N i.e.1000
2. Generate random number's using the MATLAB built-in function rand(N, 1)
3. Store the random numbers in X_i
4. Store the X_{i+1} in another matrix Y
5. Calculate the summation for $X_i X_{i+1}$ divide by N to get $E[X_i X_{i+1}]$
6. Calculate the Covariance using the equation $E[X_i X_{i+1}] - E[X_i]E[X_{i+1}]$

iv. Experiments and Results



Discussion of results:
We find that the $X_i\ and\ X_{i+1}\ are\ independent\ since\ Covariance\ is\ negative.$

v. References

- https://www.tutorialspoint.com/matlab/
- https://www.mathworks.com/help/matlab/

vi. Source Code

```
%start of the programming code
clear all;       % clear and close matlab desktop, workspace
close all;
clc;
N=1000;                 % Assign N=1000
X=rand(N,1);            % Generate N random number and store in X
Y=X(2:end,1);           % Store Xi+1 in Y
S1=Y'*X(1:N-1);         % summation for 1st equation from 0 to N
Y=[Y;rand(1)];          % Add N+1 random number to Y
Z=S1/N-(sum(X)/N)*(sum(Y)/N); % Calculate Covariance
fprintf('Covariance:%f\n\n\n',Z); %display
```