EE 511 – Project #4 **Muthulakshmi Chandrasekaran**

October 16,2017 **USC ID 4486180242**

1.1 Problem Statement

Part A – Sum of Uniform Random Variables

To define $N = \min\{n: \sum_{i=1}^{n} U_i > 1\}$, where $\{U_i\}$ are independent and identically distributed Uniform Random Variables in the range (0,1). To find $\hat{m} = E[N]$, an estimator of the mean. To guess the true value of $E[N]$.

Part B – Minima of Uniform Random Variables

To define $N = \min\{\, n: U_1 \ \le U_2 \le \cdots \le U_{n-1} \ > U_n \,\}$

where $\{U_i\}$ are independent and identically distributed Uniform Random Variables in the range (0,1). Here nth term is the first term less than its predecessor. To find $\hat{m} = E[N]$, an estimator of the mean. To guess the true value of $E[N\}]$.

Part C – Maxima of Uniform Random Variables

Considering a sequence of independent and identically distributed Uniform Random Variables in the range (0,1) , $\{U_i\}$. If $U_i > max_{i=1:j-1}\{U_i\}$, then U_j is a record.

To define X_i as a random variable from $(i-1)th$ record to the $i^{\ th}$ record. To obtain a histogram of X_2 and X_3. To compute the sample means. To find the analytical expression for $P(X_2 = k\,)$.

1.2 Mathematical Basis

A Random Variable is a mapping from the set of outcomes to the set of numbers. A Uniform Random Variable has each outcome equally likely, and values are uniformly distributed throughout some interval. Consider a uniformly distributed random variable distributed as $\{\, x_1, x_2, x_3, \ldots., x_M\}$ on the interval $[a, b]$. The probability of each outcome is given as

$$Probablity = \frac{1}{M}$$

The mean of a Uniform distribution in the range $[a, b]$ is given as

$$Mean = \frac{a+b}{2}$$

Variance of the uniform random variable distributed in the range $[a, b]$ is given as

$$Variance = \frac{(b-a)^2}{12}$$

This project is mainly based on the generation of Random Variables. Random numbers can be generated in many ways – direct methods, sampling, inversion method, rejection method etc. The topic of interest is the generation of Discrete Random Variables.

- Direct methods directly use the definition of the distribution.
- In sampling method, the actual data is samples at different points to yield a set of discrete data.
- Inversion methods are based on the observation that continuous cumulative distribution functions range uniformly over the interval (0,1). If U is a uniform random number on (0,1), then using $X = F^{-1}(U)$ generates a random number X from a continuous distribution with specified cdf F.
- Acceptance-rejection methods begin with uniform random numbers, but require an additional random number generator. To generate random numbers from a distribution with probability mass $P_p(X = i) = p_i$, given a distribution with probability mass $P_q(X = i) = q_i$ .The steps are as follows :
  i. Choose a probability mass function P_q.
  ii. Find a constant c such that $\frac{p_i}{q_i} \leq c$ for all i .
  iii. Generate a uniform random number u.
  iv. Generate a random number v from P_q.-
  v. If $c * u \leq \frac{p_v}{q_v}$, accepts and returns v.
  vi. Otherwise, rejects v and goes to step 3.

  For efficiency, and the scalar c should be small. The expected number of iterations to produce a single random number is c.

In this project, inversion method is used to check the validity of the computed value.

1.3 Simulation in Matlab

Generation of random numbers, finding the sum of the generated random variables, calculating the maxima and minima of random numbers is simulated in Matlab.

1.3.1 Part A : Sum of Uniform Random Variables.

Random numbers are generated in Matlab, and their sum is calculated, according to the given formula $N = \min\{n: \sum_{i=1}^{n} U_i > 1\}$.

Description of the code :

First the number of random sequences to be generated is obtained from the user, here 100. In each sequence, there are a fixed number of uniformly generated random values in the range $(0,1)$ – say 1000. Then for each sequence, the number of terms whose sum exceed 1, n is calculated. The actual mean is calculated by the in – built $mean$ function. The expected mean is calculated from the inverse transform method.

Observations :

The function $rand$ generates uniformly distributed random numbers. The function $rand(M,N)$ generates a MxN matrix of uniformly distributed random numbers. An example of random number generation is shown in Figure 1.

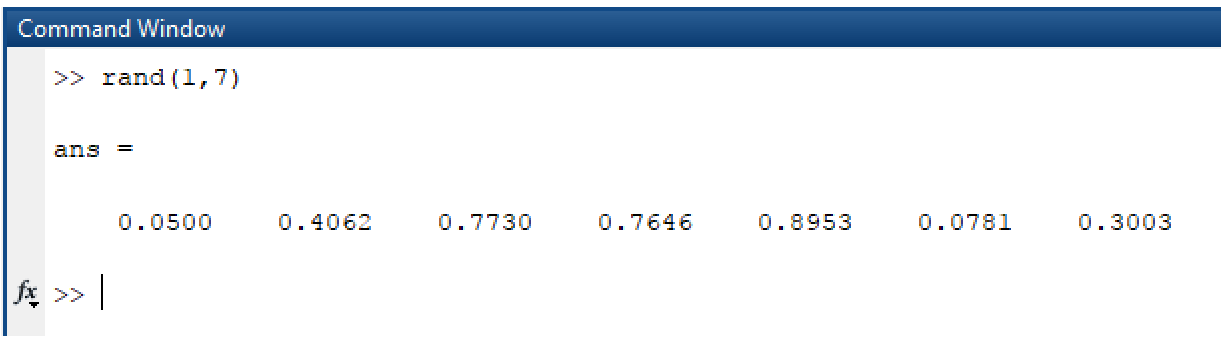


Figure 1 Generation of random numbers in Matlab using the function $rand$

The number of iterations is obtained from the user as in Figure 2.

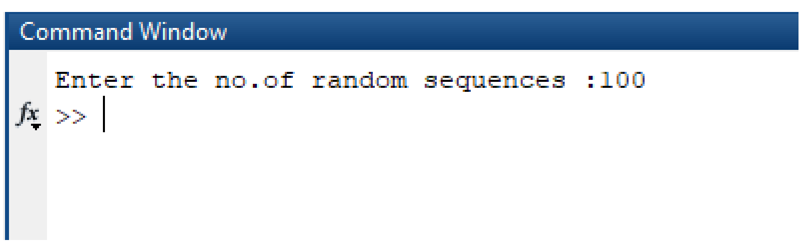


Figure 2 Window showing the number of iterations, obtained from the user

Some sample values for n is shown in Figure 3, in the Command Window.

```
Command Window
Enter the no.of random sequences :100

n =

  Columns 1 through 22

     2     4     5     3     2     3     6     3     3     3     2     3     2     2     3     2     4     2     2     3     2

  Columns 23 through 44

     3     3     3     2     3     2     3     5     2     2     2     2     4     2     2     4     3     2     2     2     4

  Columns 45 through 66

     2     5     2     2     5     3     3     3     2     3     2     2     2     2     3     3     2     2     2     3     4

  Columns 67 through 88

     2     2     2     2     2     2     3     3     3     2     2     2     2     2     2     2     2     3     2     3     2

  Columns 89 through 100

     3     4     3     2     5     3     4     3     2     2     2     3
```

Figure 3 Command Window output of all n values

Using the mean function in Matlab, the estimate of the mean is found, as in Figure 4.

```
Command Window
Enter the no.of random sequences :100
Computed value of Mean :
    2.6200
```

Figure 4 Command Window showing the computed mean value.

The value of n for each iteration is recorded. The histogram plot of the recorded observations is shown in Figure 5.

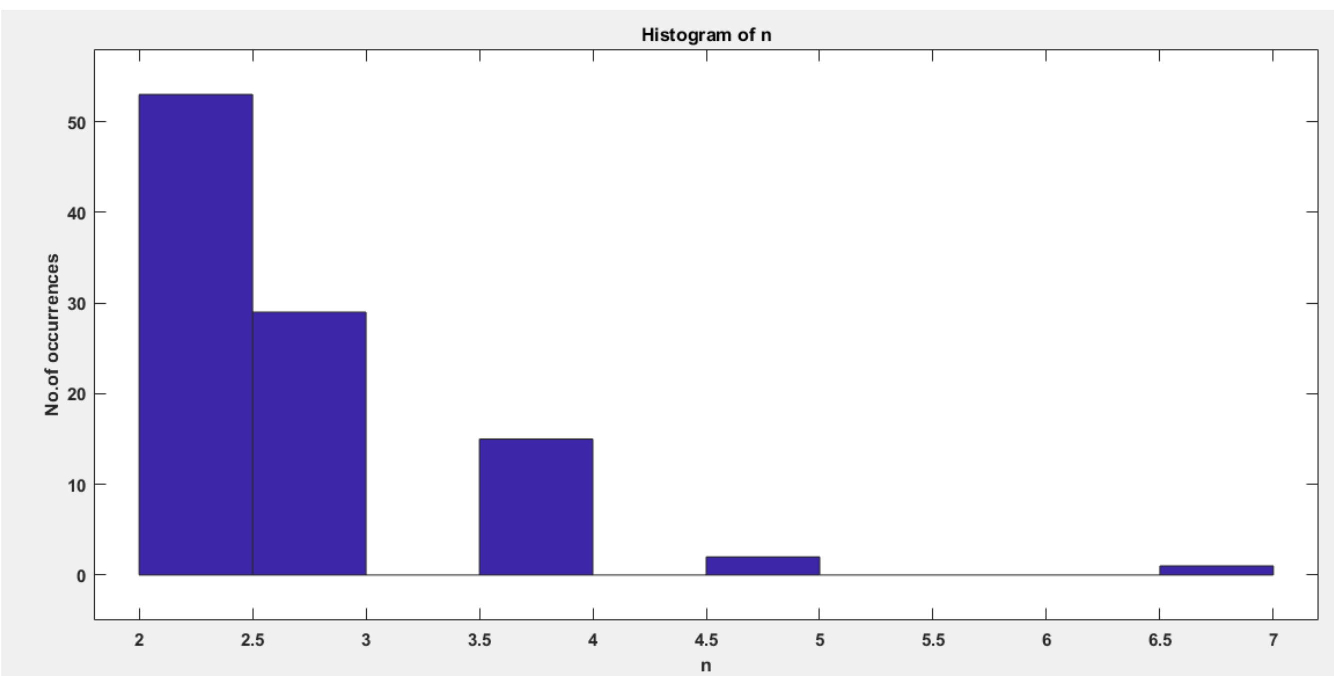


Figure 5 Histogram of n

The mean value of the output is computed to be 2.62, as shown in Figure 4 or approximately 3, since it is a discrete distribution. Inverse transform method is done to obtain the estimate of the mean to be around 2.71, as shown in Figure 6.

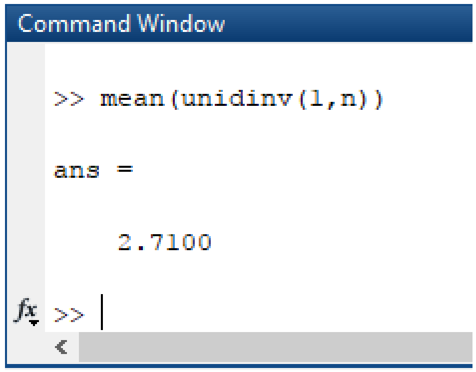


Figure 6 Mean obtained

Result :

The estimate of the mean is done by using the in – built mean function. The expected value of the mean is approximately between 2 and 3, as observed from Figure 3.

1.3.2 Part B : Minima of Uniform Random Variables

A minimum function is defined on a set of iid Uniform Random Variables. The mean is estimated.

Description of the code :

First the number of random sequences to be generated is obtained from the user, here 100. In each sequence, there are a fixed number of uniformly generated random values in the range $(0,1)$ – say 1000. Then for each sequence, the number of terms that satisfy the minimum criterion , $N = \min\{ n: U_1 \le U_2 \le \cdots \le U_{n-1} > U_n \}$ is calculated. Here nth term is the first term less than its predecessor. The actual mean is calculated by the in – built $mean$ function. The expected mean is calculated from the inverse transform method.

Observations :

The sequence of random numbers is generated using the randn function in Matlab. The number of iterations is obtained as an input from the user. The Sample output for one iteration and a 10 – point random sequence is shown in Figure 7.

```
Command Window
Enter the no.of random sequences :1

x =

    0.1807    0.9785    0.8626    0.2769    0.3890    0.6079    0.5057    0.1980    0.5703    0.9005


n =

     3


val =

    0.8626

Computed value of Mean :
     3

>>
```

Figure 7 Sample Output for one iteration – 10 point random sequence

The experiment is done with random sequences having 1000 samples, and repeated for 100 iterations(input from the user). Output mean is calculated using the inbuilt $mean$ command. The output mean is shown in Figure 8.

```
Command Window
Enter the no.of random sequences :100
Computed value of Mean :
    3.0200

>>
```

Figure 8 Output mean for 100 iterations of random sequences with 1000 samples each

The histogram of the values obtained is shown in Figure 9.

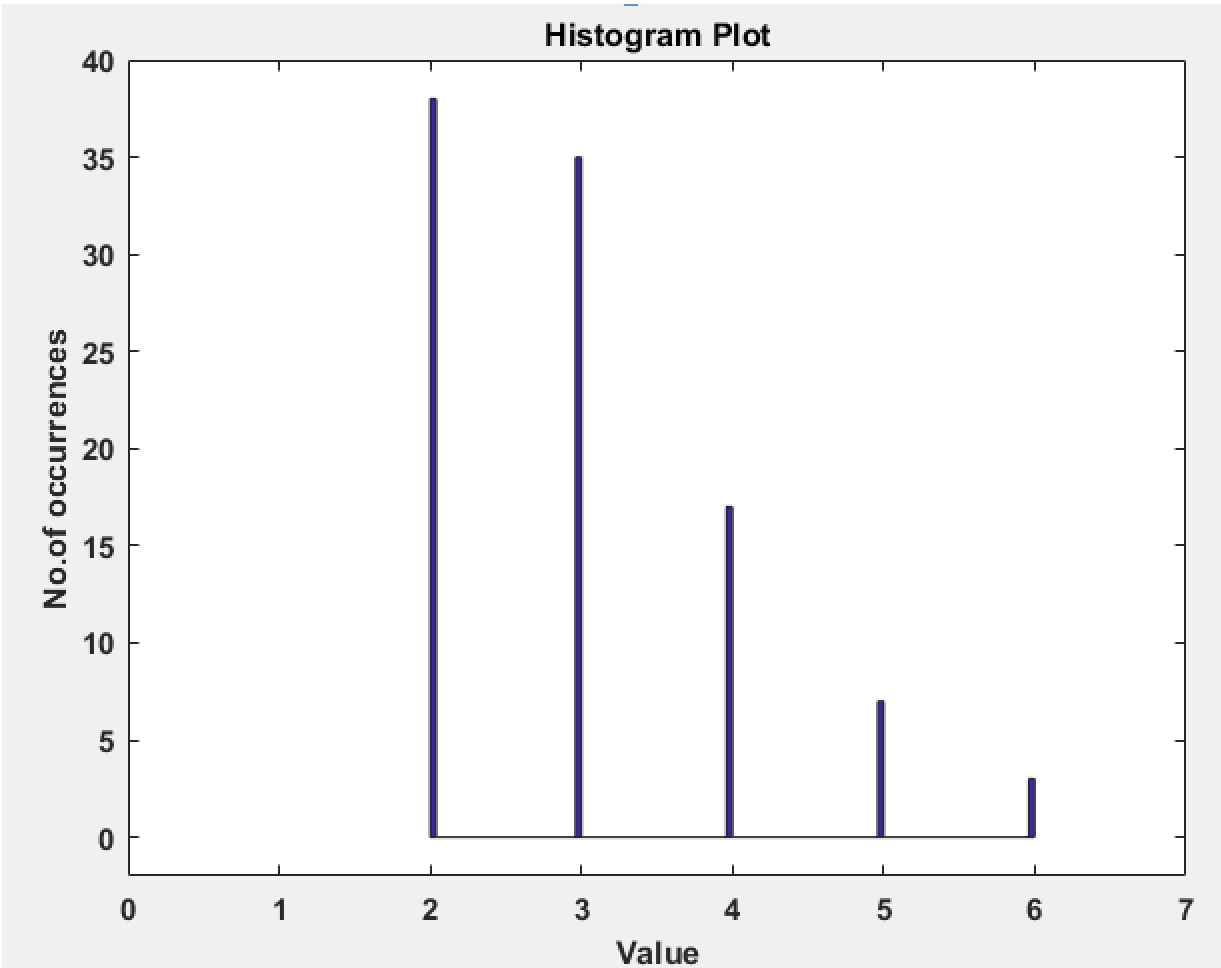


Figure 9 Histogram of the values obtained.

The value is verified using the inverse transform method, by using the in – built command $unidinv$.

Result :

The minima of the random Variables is calculated. The satisfying criterion for minima of the Uniform Random Variables is met. The mean is also calculated, and by inverse transform, the mean is found to be approximately the same.

1.3.3 Part C : Maxima of Uniform Random Variables

Uniform Random Variables are generated based on a specific criterion. Then sample mean is calculated, and histograms are plotted.

Description of the code :

A sequence of independent and identically distributed Uniform Random Variables in the range (0,1) is generated using the $rand$ function. Then individual records are created using the specified criterion: If $U_i > max_{i=1:j-1}\{U_i\}$, then U_j is a record.

Then another random variable is created. X_i as a random variable from $(i-1)th$ record to the i^{th} record. Probability histogram of X_2 and X_3 are computed . Sample mean is computed. $P(X_2 = k\ )$ is also found.

Observations :

The sequence of random numbers is generated using the randn function in Matlab. The number of iterations is obtained as an input from the user. The random variables X2 and X3 are calculated. The Sample output for one iteration and a 10 – point random sequence is shown in Figure 10.

```
Command Window
Enter the no.of random sequences :1

x =

    0.4867    0.5651    0.1412    0.7752    0.9843    0.8268    0.0165    0.3300    0.2242    0.7566


X2 =

     1


X3 =

     2

>>
```

Figure 10 Sample Output for one iteration – 10 point random sequence

This experiment is repeated for 100 iterations. All values of X2 and X3 are stored in separate arrays. Figure 11 shows the histogram plot of X2 and Figure 12 shows the histogram plot of X3.

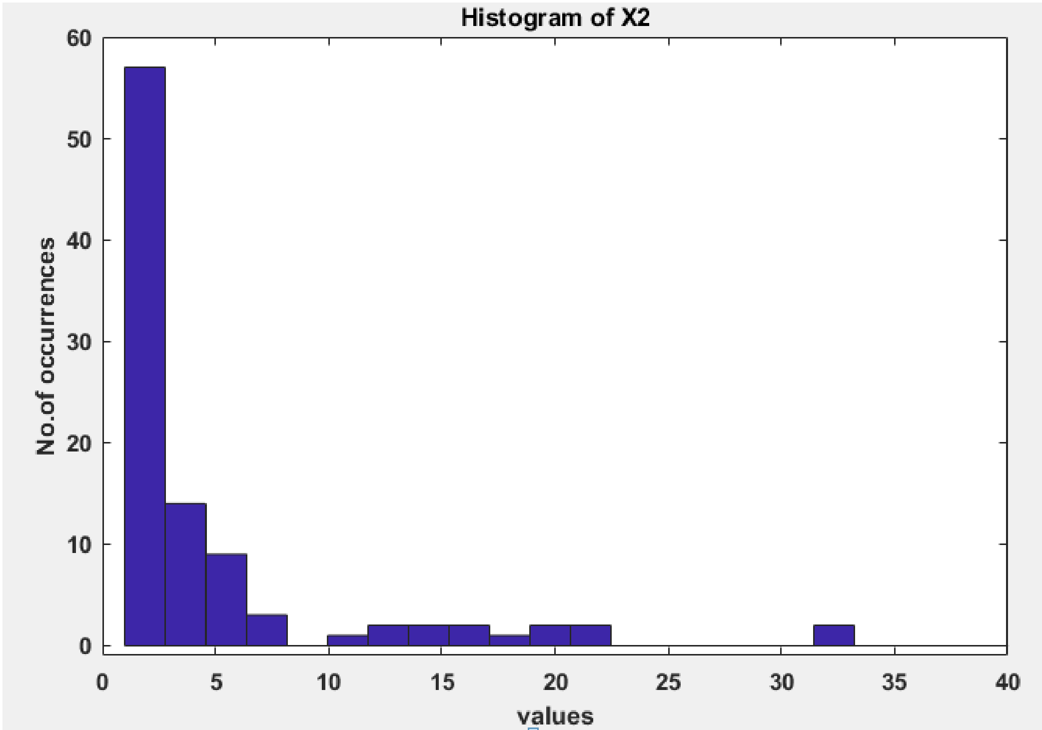


Figure 11 Histogram of X2

From Figure 11, it is observed that the value of 1 has the highest probability of occurrence. So, out of 1000 iterations performed, the value 1 occurred more number of times, compared to all other values.

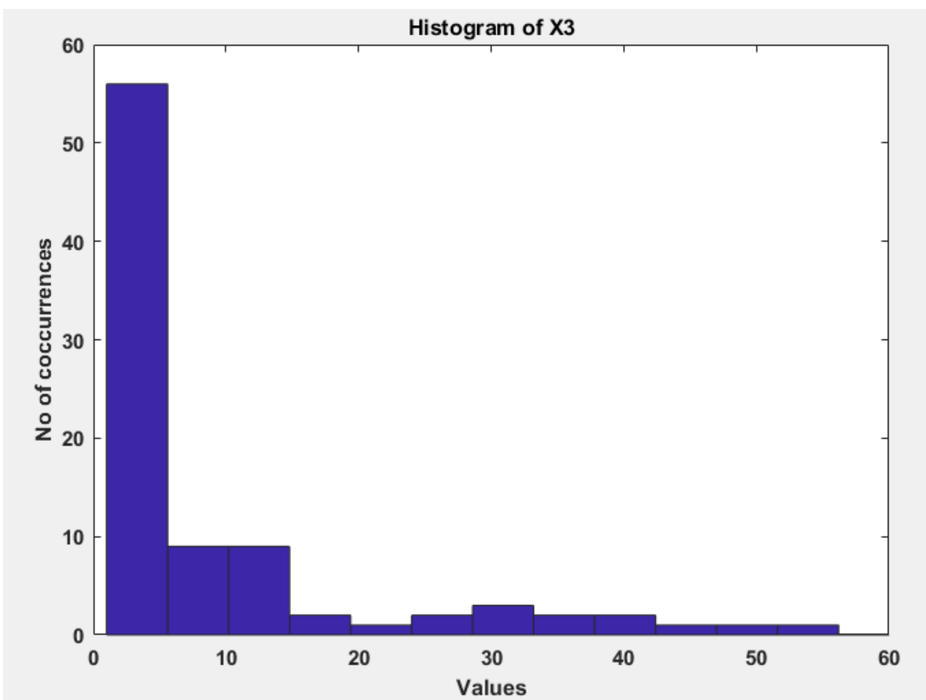


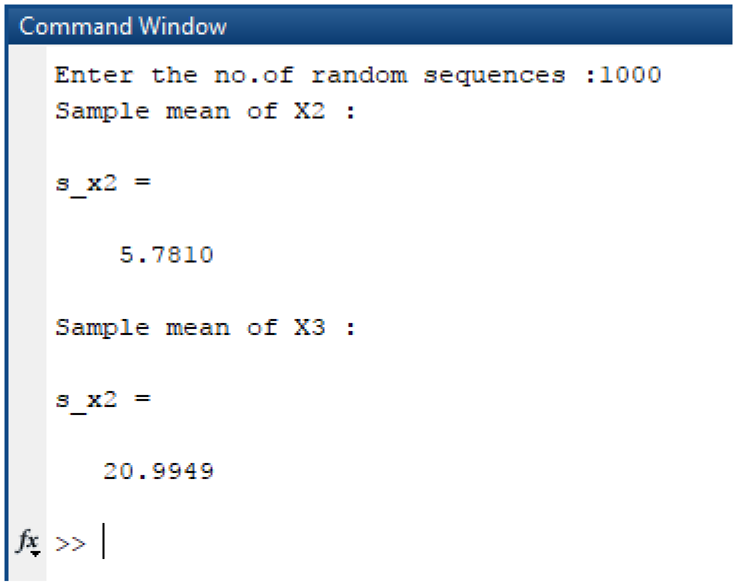
Figure 12 Histogram of X3

From Figure 12, it is observed that the value of 1 has the highest probability of occurrence. So, out of 1000 iterations performed, value 1 occurred more number of times, compared to all other values.

Sample mean of each of X2 and X3 is calculated. Sample mean $\bar{x}$ of a sequence of M samples, say $x_1, x_2, x_3, \ldots, x_M$ is given by

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_M}{M}$$

Sample mean of X2 and X3 is shown in Figure 13.



```
Command Window
Enter the no.of random sequences :1000
Sample mean of X2 :

s_x2 =

    5.7810

Sample mean of X3 :

s_x2 =

   20.9949

>>
```

Figure 13 Sample mean of X2 and X3

To find $P(X_2 = k)$, the first random variable that is generated is also to be considered. $P(X_2 = k)$ now becomes $P(X_2 = k|U_1)$. Considering that the first random variable U_1 has already occurred, the value of k differs.

If k=1, $P(X_2 = 1) = \frac{No.of\ occurrences\ of\ 1\ in\ the\ array\ X2}{Total\ number\ of\ occurrences} = \frac{56}{100} = 0.56$

If k=2, $P(X_2 = 2) = \frac{No.of\ occurrences\ of\ 2\ in\ the\ array\ X2}{Total\ number\ of\ occurrences} = \frac{20}{100} = 0.2$

This is shown in Figure 14.

```
Command Window
Enter the no.of random sequences :100
Sample mean of X2 :

s_x2 =

    2.9600

Sample mean of X3 :

s_x2 =

   15.2626

No.of occurrences of k=1 :

c1 =

    56

No.of occurrences of k=2 :

c2 =

    20

fx >>
```

Figure 14 Finding $P(X_2 = k\ )$

From the above, it is observed that $P(X_2 = k\ )$ follows a Geometric distribution.

<u>Result :</u>

The maxima of random variables is computed, based on the specific criterion of records. Sample mean for two records is calculated. It is observed that $P(X_2 = k\ )$ follows a Geometric distribution.

1.4 Results and Conclusion

1. Random sequences are generated using the rand function. The sum of the generated random sequence is found. The minimum of the uniform sum distribution function is calculated. Mean is computed.
2. The random numbers are generated for each iteration of the total 1000 iterations, and in each case, mean is calculated. The minima of the random variables is computed.
3. The maxima of the Uniform Random Variables is also computed. Another random variable is also generated. The sample means are calculated. It is found that probability distribution is geometric.

1.5 Reference

1. Sheldon M Ross, “Simulation”, 5th edition
2. www.mathworks.com

1.6 Matlab Codes

Part A – Sum of Uniform Random Variables

```
%Uniform Random Variables
clc;clear all;close all;
pts=1000; % Length of each Random sequence
trials = input('Enter the no.of random sequences :'); % No.of iterations
n=[];
for i=1:1:trials
    x=rand(1,pts); %Generate each random sequence x~(0,1)
    sum=0;j=0;
    %Check if the sum of first j numbers is >1
    while(sum<=1)
        j=j+1;
        sum=sum+x(j);
    end
    %Store the value of all such j in an array n, and values in array val
    n=[n j];
    %Estimating the mean

end
n
 %Calculation of expected value of n
mean_n=mean(n);
disp('Computed value of Mean :');
disp(mean_n)
hist(n)
```

Part B – Minima of the Uniform Random Variables

```
%Minima of Uniform Random Variables
clc;clear all;close all;
pts=1000; % Length of each Random sequence
trials = input('Enter the no.of random sequences :'); % No.of iterations

n=[];
val=[];
for i=1:1:trials
    x=rand(1,pts); %Generate each random sequence x~(0,1)
    1sum=0;j=1;
    %Check if the term j is less than  term j-1
    while(x(j)<x(j+1))
        j=j+1;
    end
    %Store the value of all such j in an array n, and values in array val
    n=[n j+1];
    val=[val x(j+1)];
end
```

```
 %Calculation of expected value of n
mean_n=mean(n);
disp('Computed value of Mean :');
disp(mean_n)
hist(n)
```

Part C – Maxima of the Uniform Random Variables

```
%Minima of Uniform Random Variables
clc;clear all;close all;
pts=1000; % Length of each Random sequence
trials = input('Enter the no.of random sequences :'); % No.of iterations

n=[];X2=[];X3=[];

 for i=1:1:trials
     x=rand(1,pts); %Generate each random sequence x~(0,1)
    sum=0;j=1;
    val=[];
    %Compute the records in the sequence
    for k=1:1:pts
        y=x(1:k);
        max_no=max(y);
        val=[val max_no];
    end
    v=[]; v=[1];
    %find the values of the difference in records
    for t=2:1:length(val)
        if val(t-1)==val(t)
        else
            v=[v t];
        end
    end
    %Generate each RV Xi= distance of (i-1)th record to ith record
    X=[];X(1)=v(1);
    for m=2:1:length(v)
        b=v(m)-v(m-1);
        % X2 takes the second record
        if m==2
           X2=[X2 b];
        end
        %X3 takes the third record
        if m==3
            X3=[X3 b];
        end
    end
 end
%Computation of sample mean of X2
s1=0;c1=0;c2=0;
for i=1:length(X2)
    s1=s1+X2(i);
    if X2(i)==1
        c1=c1+1;
    else if X2(i)==2
            c2=c2+1;
        end
    end

end
```

```
disp('Sample mean of X2 :');
s_x2=s1/length(X2)

%Computation of sample mean of X3
s1=0;
for i=1:length(X3)
    s1=s1+X3(i);
end
disp('Sample mean of X3 :');
s_x2=s1/length(X3)

%Finding P(X2=k)
%k=1
disp('No.of occurrences of k=1 :');c1
%k=2
disp('No.of occurrences of k=2 :');c2
```