Name: Deepika Kanade
Assignment 1- EE 511

Q 1.a)

Task: Writing a routine to simulate a fair Bernoulli trial and generating a histogram for 100 simulated Bernoulli trials.

Code:-

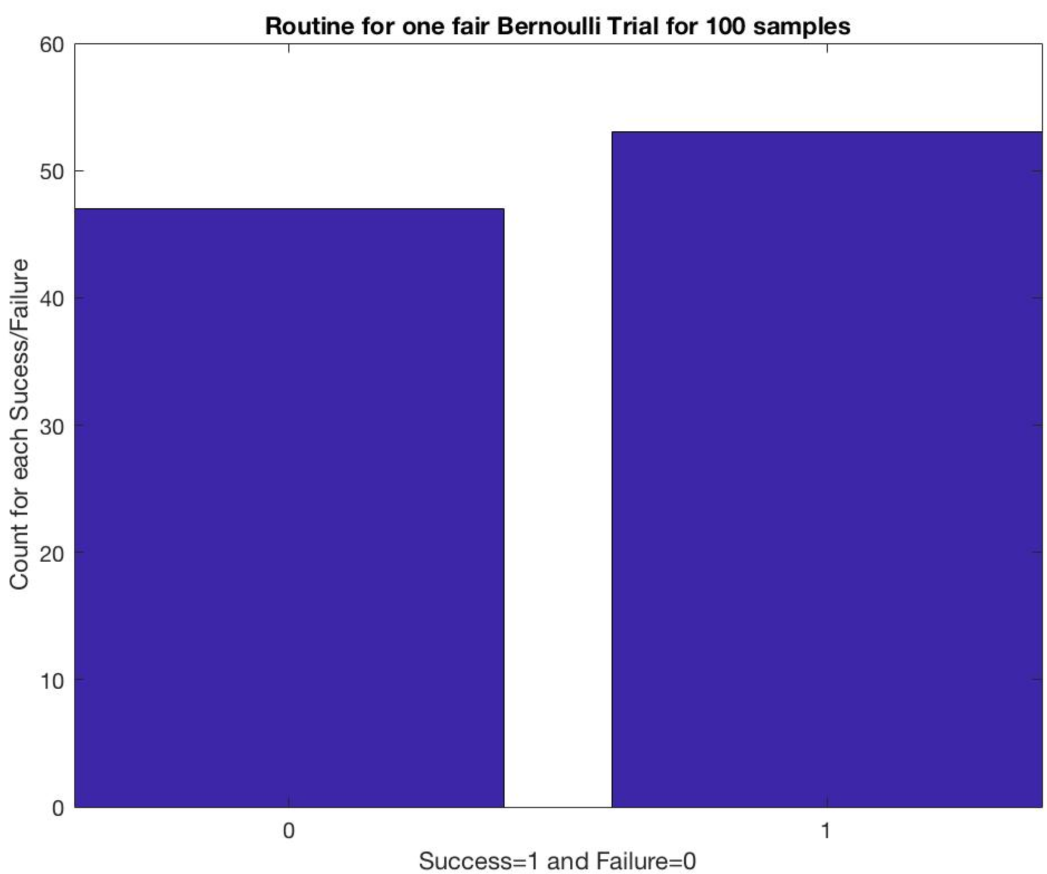
```
clc;
close all;

%Initialization of variables
size = 10;
x = rand(size);   %To generate 100 uniform random numbers
A=2;
count = zeros(size*size,1);
count_final = zeros(A,1);
count1=0;
count0=0;

%Generate a fair Bernoulli trial for 100 samples
for i=1:100
  if(x(i)>=0.5)
    count(i)=1;
    count1=count1+1;
  else
    count(i)=0;
    count0=count0+1;
  end
end

%count successes for all samples
for j = 1:size*size
  for i = 1:A
    if count(j) == i-1
      count_final(i,1) = count_final(i,1)+1;
    end
  end
end

disp("No. of Successes are: ");
disp(count1)
disp("No. of Failures are: ");
```

```
disp(count0);

%Plotting Histogram of the data
bar(0:1,count_final);
xlabel("Success=1 and Failure=0");
ylabel("Count for each Sucess/Failure");
title("Routine for one fair Bernoulli Trial for 100 samples");
```

Output:-

```
No. of Successes are:
    53

No. of Failures are:
    47
```

Histogram for 100 samples



A routine to calculate the successes of 100 Bernoulli trials was implemented in Matlab and the histogram was plotted. As is visible from the histogram, the number of successes is 53 which is labelled as 'l' in the routine, whereas the number of failures is 47 which is labelled as '0' in the routine. The total number of successes and failure combined is 100 which is equal to the random samples generated. The distribution of the 100 simulations is a Bernoulli distribution.

Q 1.b)

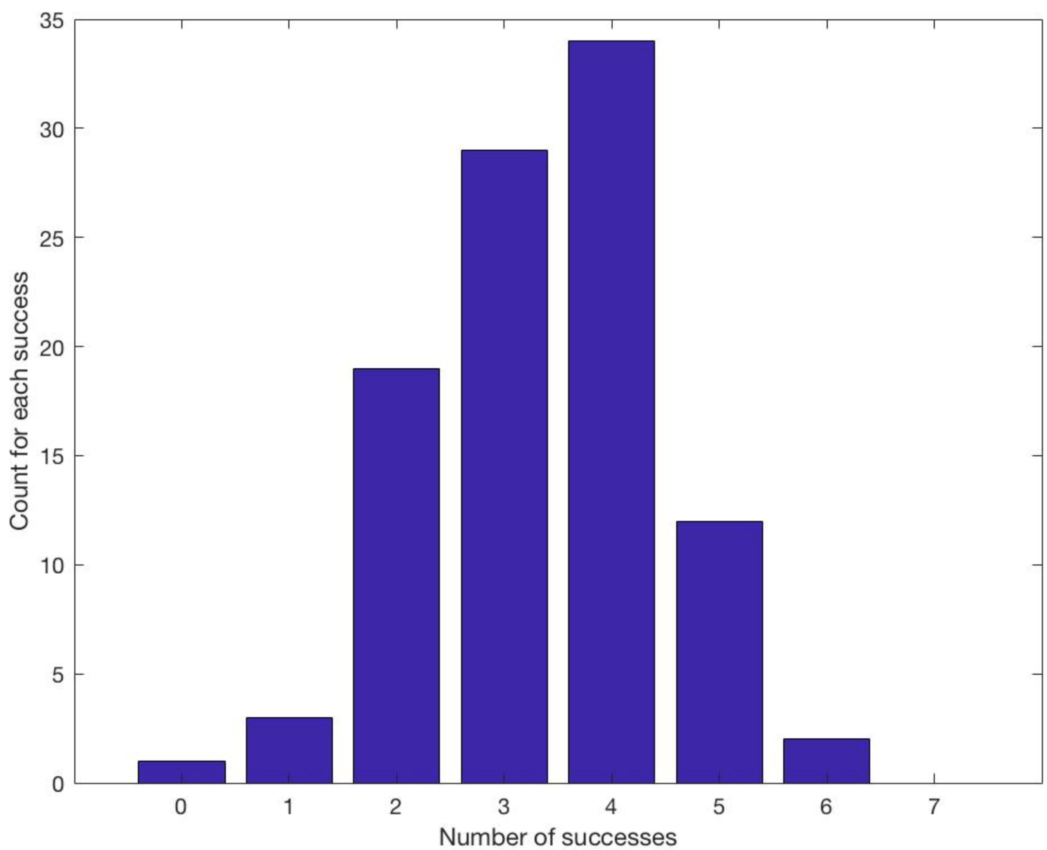
Task: Writing a routine to count the number of successes in 7 Bernoulli trials and generate a histogram for 100 such samples.

Code :-

```
clc;

%Initialization of variables
bern = 7;
prob = 0.5;
samp = 100;
temp1=zeros(samp);
temp2=zeros(bern+1,1);

%Generate Random Numbers and count successes in 7 trials for 100 samples
for j = 1:samp
   count=0;
   for i = 1:bern
      S(i) = rand(1);
      if S(i)>prob
      count = count + 1;
      end
   end
   temp1(j) = count;
end

%count successes for all samples
for j = 1:bern+1
   for i = 1:samp
     if temp1(i) == j-1
        temp2(j,1) = temp2(j,1)+1;
     end
   end
end

%Potting Histogram
```

```
figure(1);
bar(0:bern,temp2);
xlabel('Number of successes');
ylabel('Count for each success');
```

Output :-



A routine to calculate the number of successes in 7 fair Bernoulli trials was implemented and the code was run for 100 such samples. A histogram of the number of successes possible and the count for each success was computed which resembles a normal(Gaussian) curve. As in this code the Bernoulli trials are tested for 100 samples, the output does not resemble an exact Gaussian curve, but if the number of samples is increased beyond 100, the histogram will get closer to the ideal Gaussian curve. As is visible from the histogram, the most number of successes encountered in 7 Bernoulli trials is 4 and the count for the occurrence of four successes is 33. The distribution for the 7 Bernoulli trials is a Binomial distribution but when the random variable is simulated 100 times, it starts getting close to a normal curve.

Q 1.c)

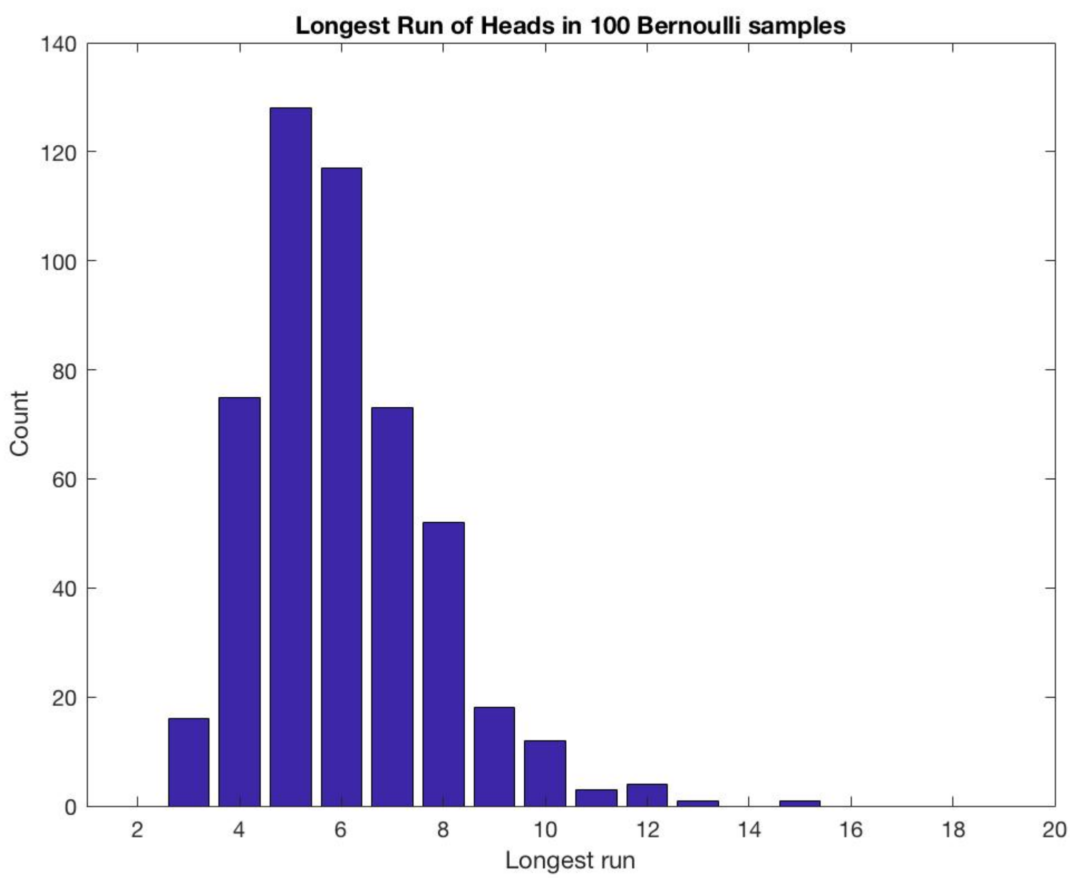
Task: Writing a routine to count the longest run of heads in 100 Bernoulli samples taking into consideration 500 such simulations for the Random Variable.

Code: -

```
clc;

%Initialization of variables
size = 500;
prob = 0.5;
samp = 100;
sample = zeros(samp,1);
count = zeros(size,1);
count_final = zeros(samp,1);
max = 0;
temp= 0;

%Generate Random samples and find longest run of heads in all the trials
for i = 1:size
  max = 0;
  temp = 0;
  for j = 1:samp
    A(j) = rand(1);
     if A(j)>prob
        A(j) = 1;
     else
        A(j) = 0;
     end
  end
  for j = 1:samp-1
     if A(j) == 1 && A(j+1) == 1
        temp = temp+ 1;
     else
        temp = 0;
     end
     if max < temp
        max = temp;
     end
  end
  count(i) = max;
end

%Count the number of longest run of Heads for all samples
```

```
for i = 1:samp+1
   for j= 1:size
      if count(j) == i-1
         count_final(i) = count_final(i) + 1;
      end
   end
end

%Plotting Histogram
figure(1);
bar(1:samp,count_final');
xlim([1,20]);
xlabel('Longest run');
ylabel('Count');
title('Longest Run of Heads in 100 Bernoulli samples');
```

Output :-



A routine to count the longest run of heads in 100 Bernoulli samples was implemented wherein the process was repeated 500 times. The histogram for the count of longest run of heads and the count of such runs in 500 simulations was plotted and it was observed that this resembles a Poisson curve. Hence, the distribution is Poisson distribution with the mean=4 approximately.

Q 2)

Task: Writing a routine to count the sum of number of successes in k=5,10,30,50 Bernoulli trials and generate a histogram for 300 such samples.

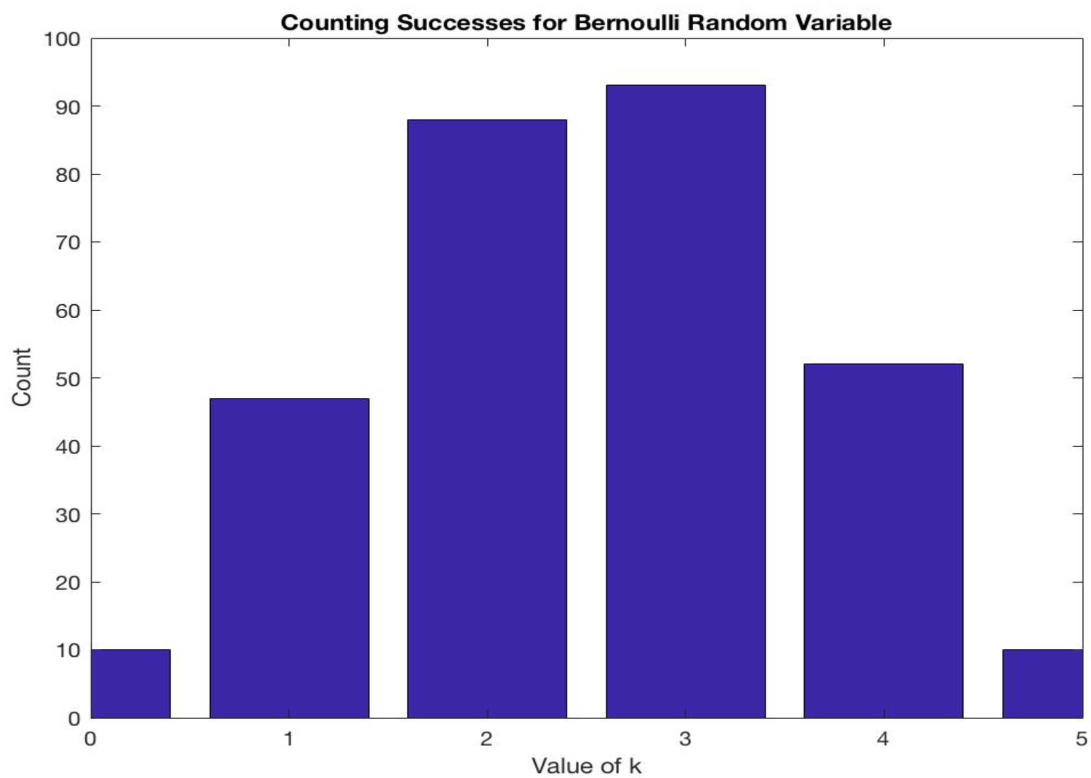
Code :-

```
clc;

%Initialization of variables
k = 50;    %k = 5,10,30,50
prob = 0.5;
samp = 300;
count1 = zeros(samp);
count_total = zeros(k+1,1);

%Generate random numbers and count number of successes in all the trials
for j = 1:samp
   count = 0;
   for i = 1:k
      A(i) = rand(1);
      if A(i) > prob
      count = count + 1;
      end
   end
   count1(j) = count;
end

%Count of Successes in all the samples
for j = 1:k+1
   for i = 1:samp
     if count1(i) == j-1
        count_total(j,1) = count_total(j,1)+1;
     end
   end
end
```

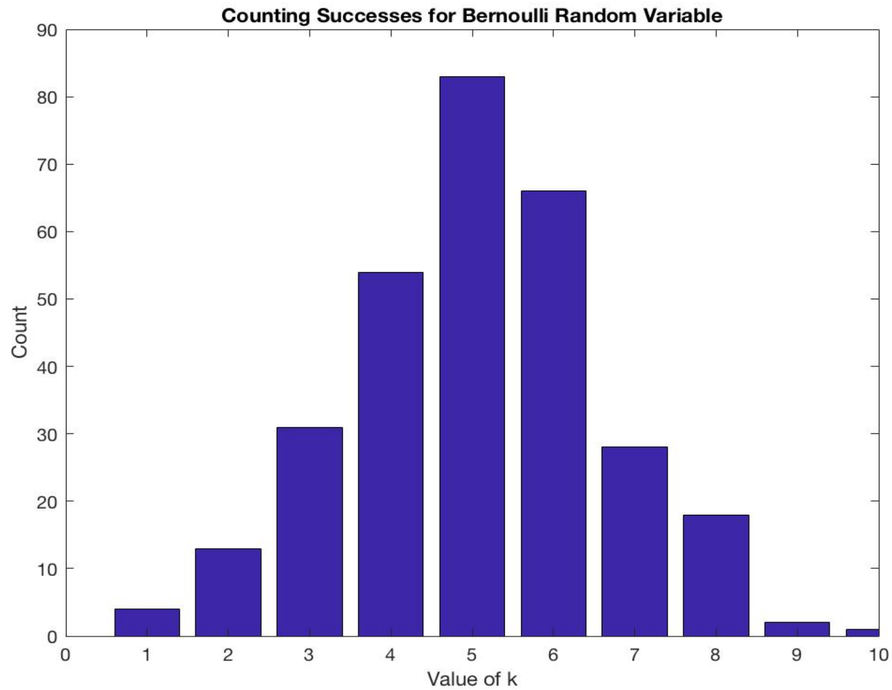
```
%Plotting histogram
figure(1);
bar(0:k,count_total);
xlim([0,k]);
xlabel('Value of k');
ylabel('Count');
title('Counting Successes for Bernoulli Random Variable');
```

Output :-

K=5

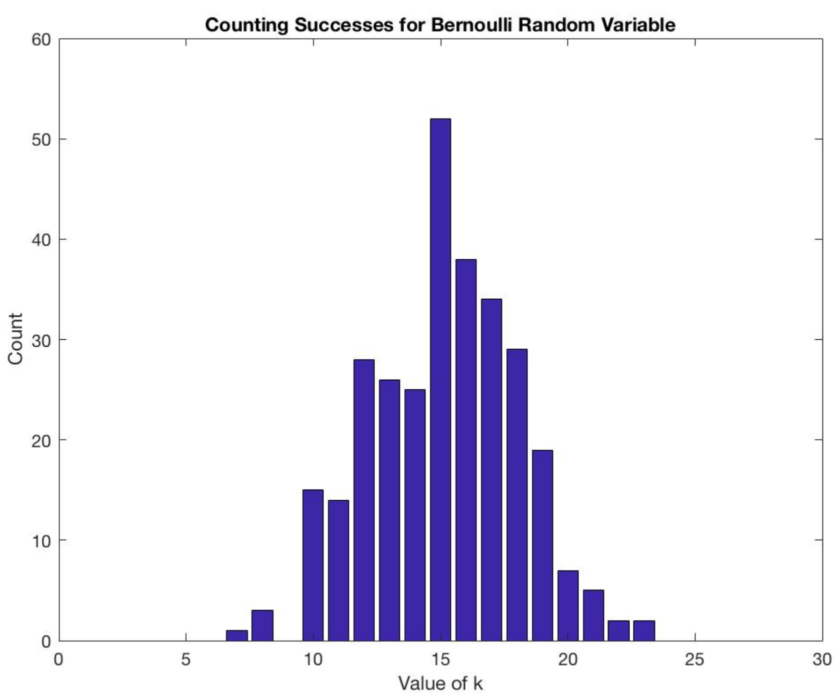


K=10



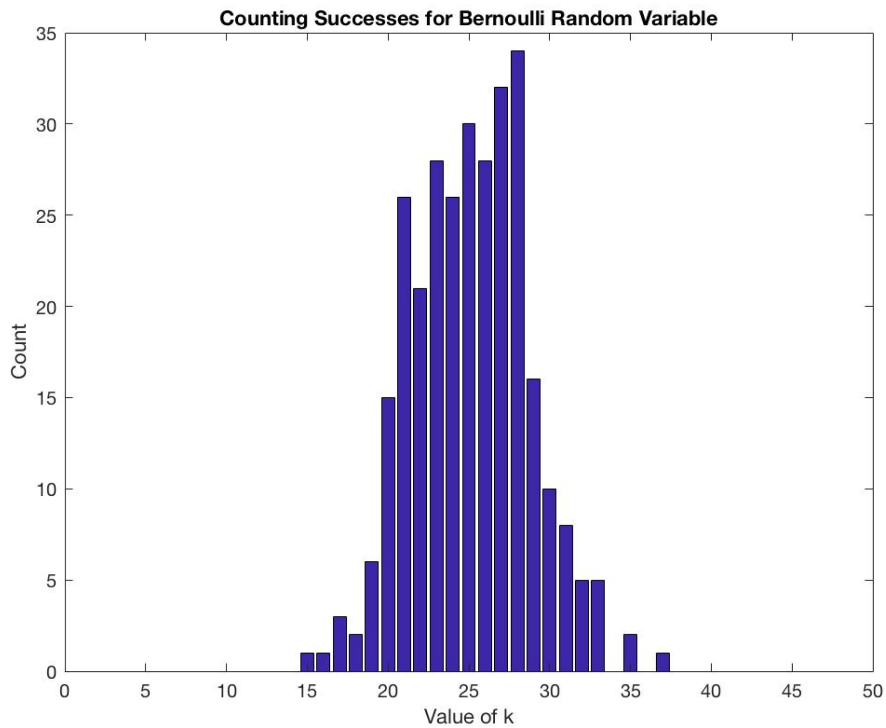
Counting Successes for Bernoulli Random Variable
Count
Value of k
0
10
20
30
40
50
60
70
80
90
0
1
2
3
4
5
6
7
8
9
10

K=30



Counting Successes for Bernoulli Random Variable
Count
Value of k
0
10
20
30
40
50
60
0
5
10
15
20
25
30

K=50



A routine to calculate the sum of number of successes in k=5,10,30,50 samples was implemented, and the code was run for 300 such samples. A histogram of the number of successes possible and the count for each success was computed which resembles a normal(Gaussian) curve.

As the value of K is increased from 5 to 50, the distribution gets closer to the Gaussian distribution as the variance starts reducing and a lot of data gets accumulated near the mean. As seen from the histograms, for K=5 the histogram is spread out and the variance is high. However, for K=30 and 50 a large amount of data is centered near the means 15 and 25 respectively. This happens as a result of the central limit theorem, as even though the distribution of each Bernoulli trial where k=5,10,30,50 is a binomial random variable, the distribution generated over 300 such simulations is a Gaussian distribution.

Q 3)

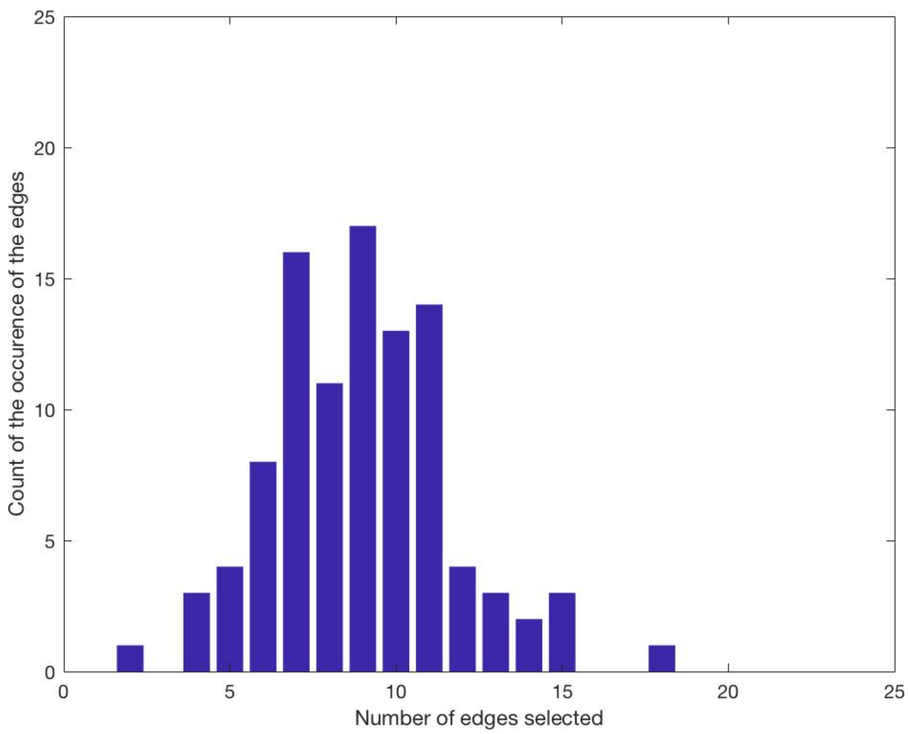
Task: Writing a routine to select an edge with probability p=0.05 out of all possible edges and generating a histogram for 100,500,1000 such samples.

Code :-

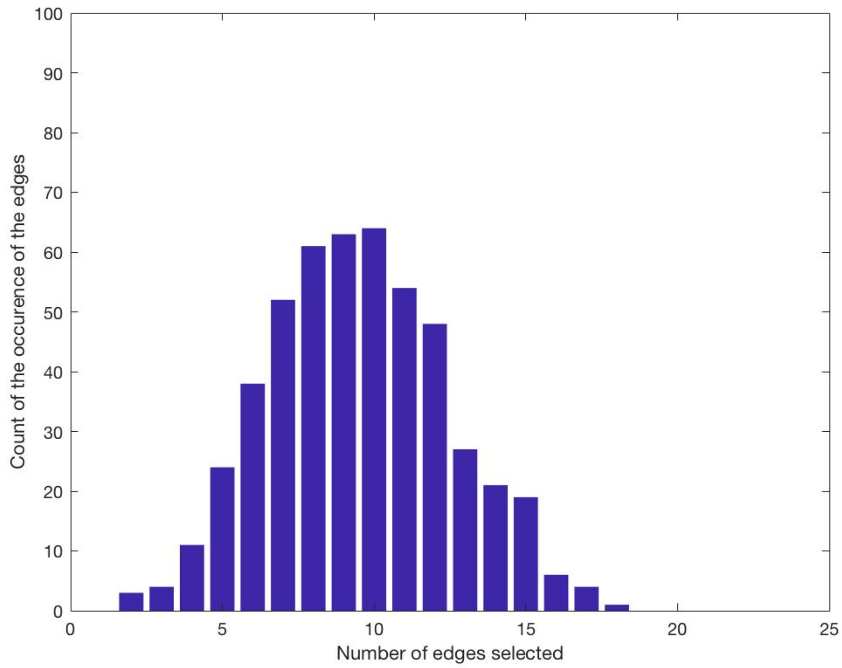
```
clc;
close all;

%Initialization of variables
n=20;
size = nchoosek(20,2);   %To choose 190 combinations
prob = 0.05;
samp = 1000;  %100,500,1000 samples
temp1=zeros(samp);
count_final=zeros(size+1,1);

%Generate Random Numbers and count successes for all trials
for j = 1:samp
   count=0;
   for i = 1:size
      S(i) = rand(1);
      if S(i)<=prob
      count = count + 1;
      end
   end
   temp1(j) = count;
end

%count successes for all samples
for j = 1:size+1
   for i = 1:samp
     if temp1(i) == j-1
        count_final(j,1) = count_final(j,1)+1;
     end
   end
end

%Potting Histogram
figure(1);
bar(0:size,count_final);
xlim([0,25]);
ylim([0,150]);
xlabel('Number of edges selected');
ylabel('Count of the occurence of the edges');
```

Output :-

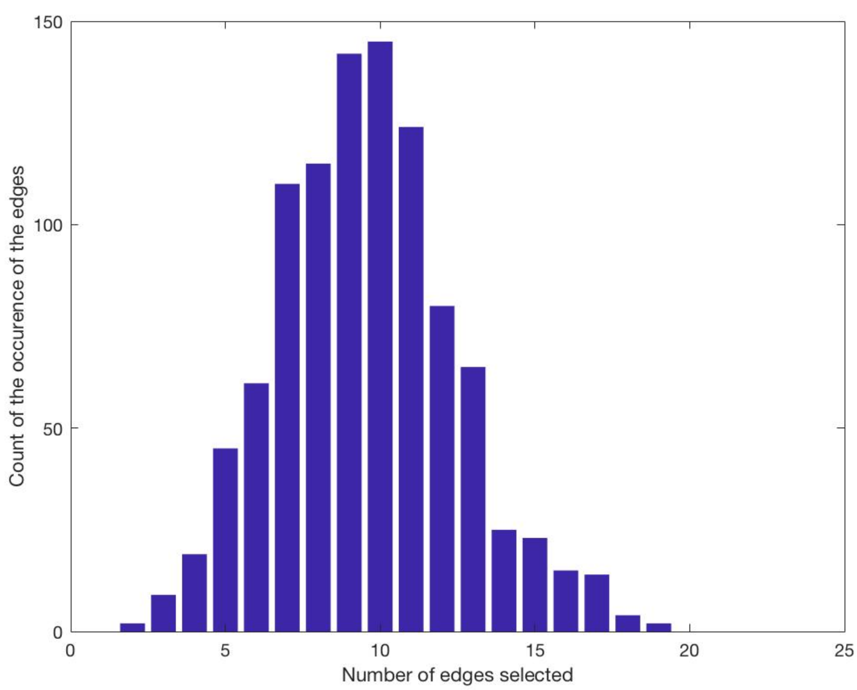
Histogram for 100 samples



Histogram for 500 samples



Histogram for 1000 samples



A routine to count the number of edges selected with a probability of 0.05 was implemented and the code was written to run 100,500,1000 such samples. The histogram of the number of edges selected with the count of each edge was plotted which resembles a normal curve.

1) For a group of n=20 people, when 2 people are selected at random in an unordered fashion, the total combinations are 20choose2 i.e. 20*19/2=190. Hence, the total number of edges in a group of n=20 people is 190.
2) If the random uniform number generated has a probability less than 0.05, it is considered as a success. This is how the successes are counted.
3) As seen from the histograms, as the number of simulations increases from 100 to 1000, the curve approximates more to a Gaussian curve with the data centered around mean and the variance is also low. When the random variable is made to undergo 100,500,1000 simulations, it resembles a normal curve, hence has a Gaussian distribution.