Xin Zhao (ID:2002778508)

EE 511 (Brandon Franzke)

Project 2 (Due 1/26/2017)

1. **Simulate sampling uniformly on the interval [-3,2].**
2. **Generated a histogram of the outcomes**

I use MATLB to achieve it. Using “rand()” to generate an array[-3,5]

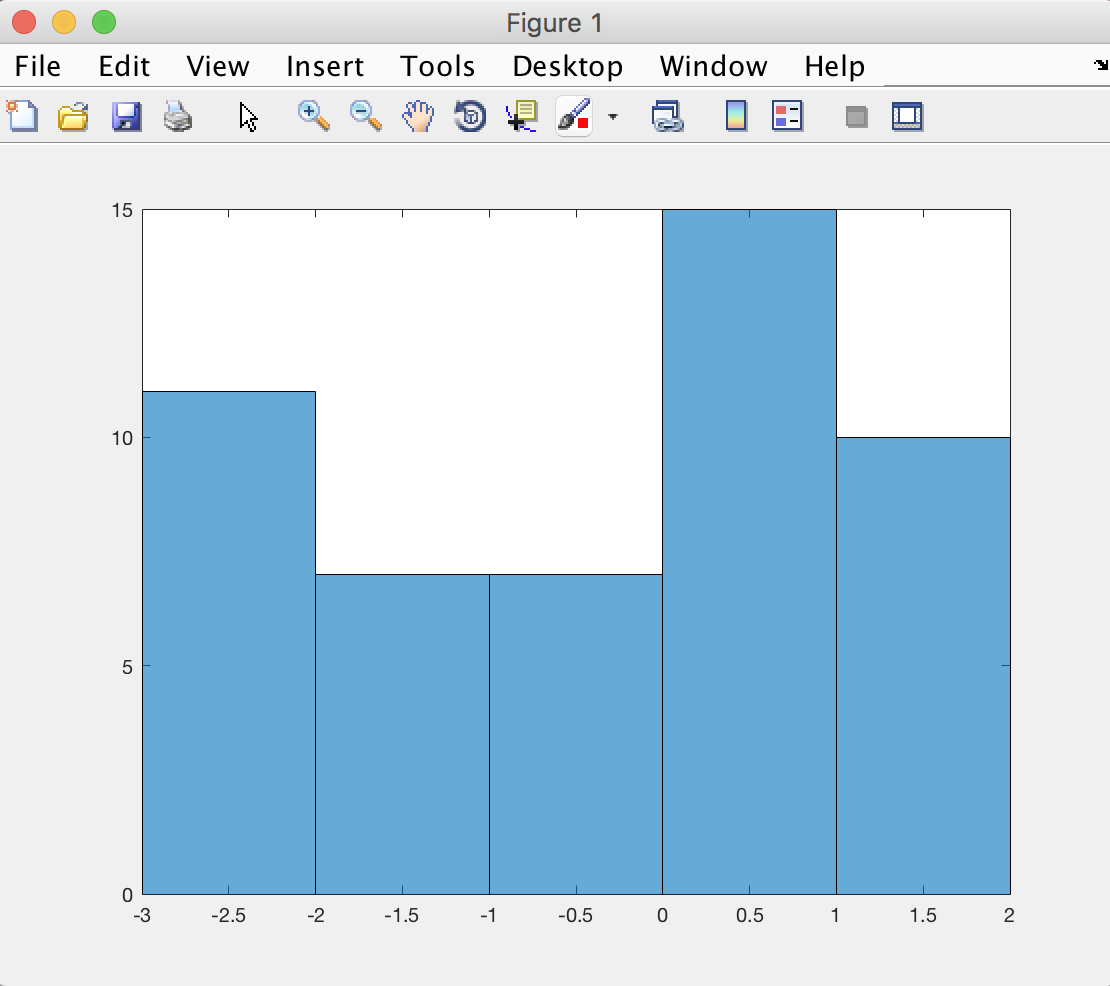
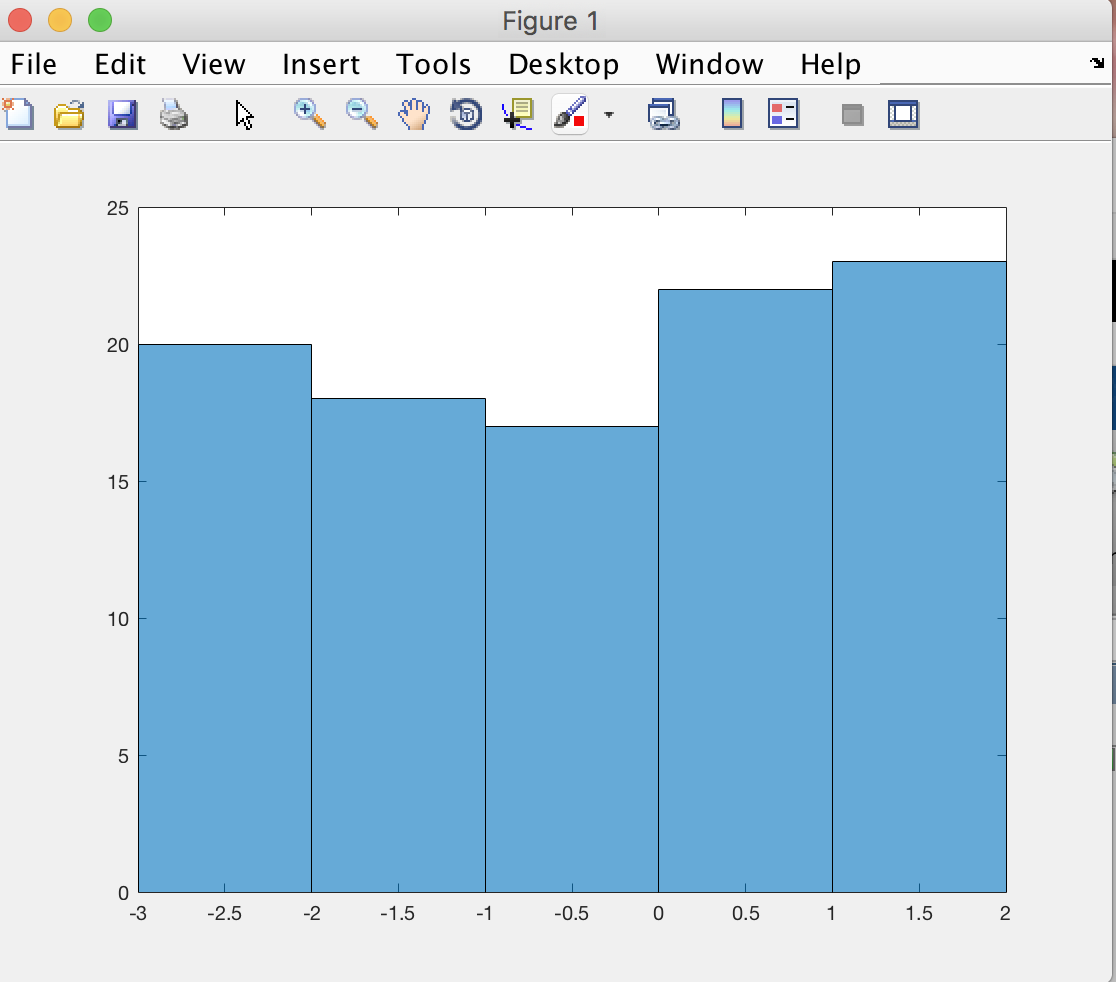
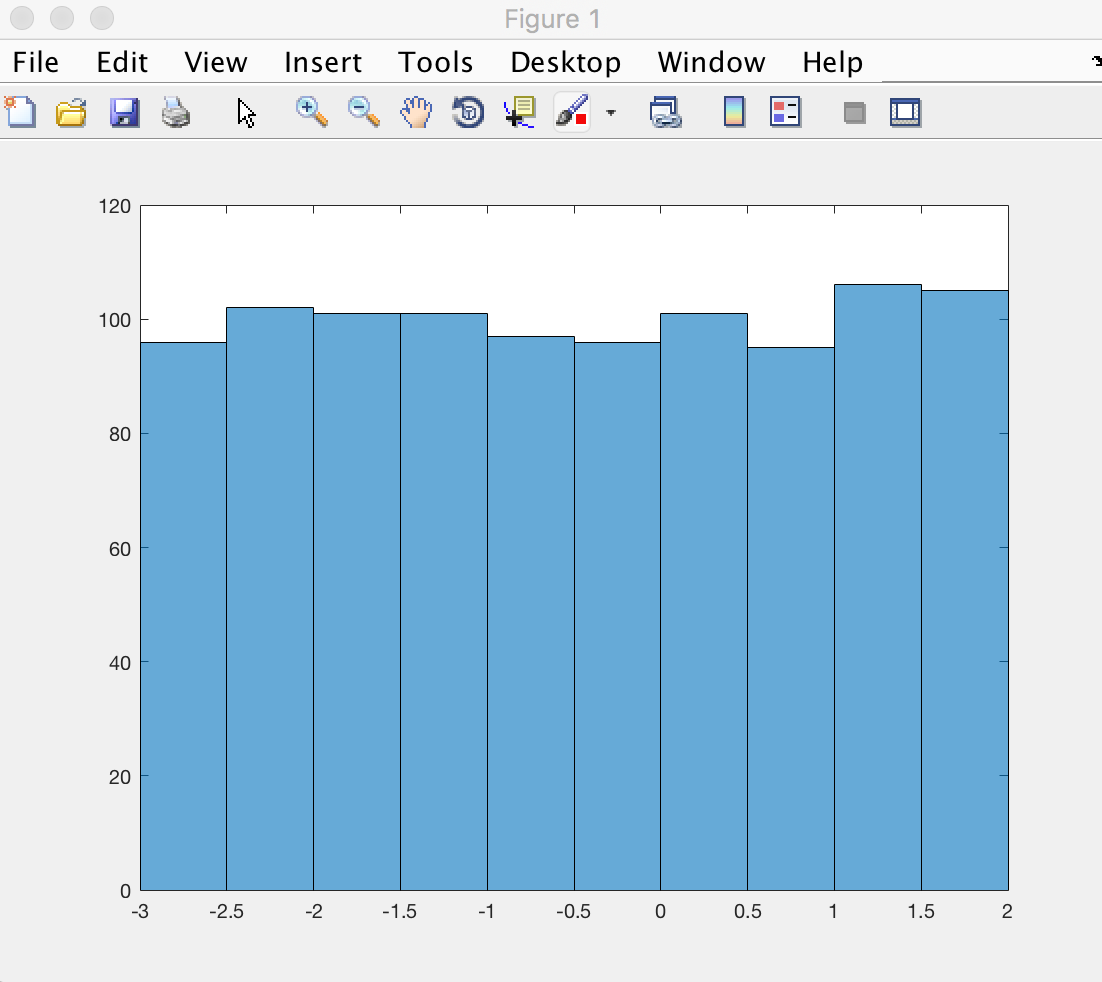
Here is the coding and histogram:

n=input('number\n');

b=-3+5\*(rand(1,n));

h=histogram(b);

The sample number is 50 The sample number is 100 The sample number is 1000

1. **Compute the sample mean and sample variance for your samples. How do these values compare to the theoretical values? If you repeat the experiment will you compute a different sample mean or sample variance?**

I use the build-in function “mean” and “var” to calculate directly.

The theory value is :

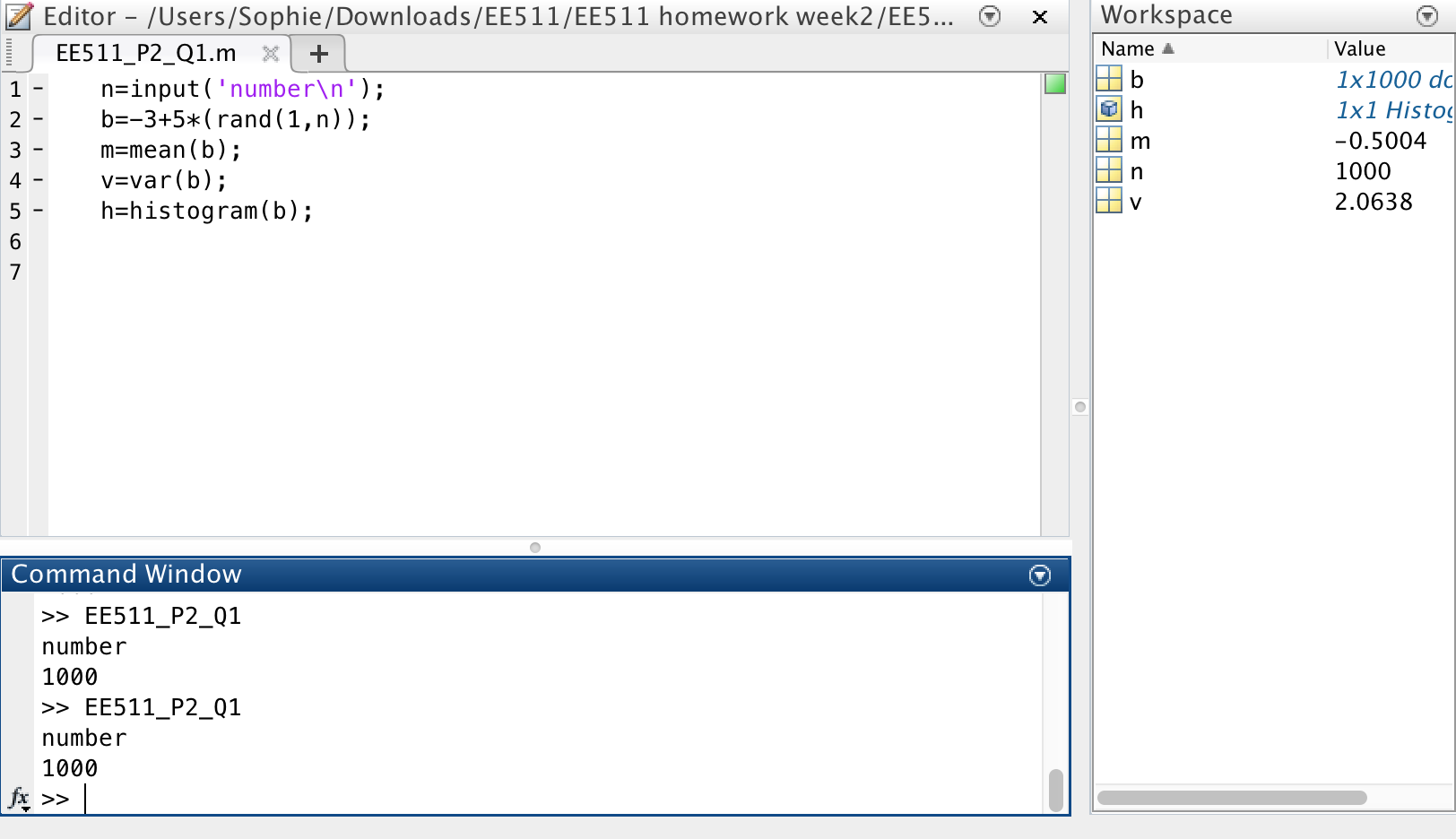
Here is the coding:

n=input('number\n');

b=-3+5\*(rand(1,n));

m=mean(b);

v=var(b);



1. **Compute the bootstrap confidence interval for the sample mean and sample standard deviation.**

According to the method of calculating bootstrap confidence interval, I need to draw too many times resample and calculate each mean. Then order the mean from small number to large number to observe whether the position of 2.5% and position of 97.5% is very close. In MATLAB, I use “randsample()” to repeat drawing resample, and use “sort()” to order the array of mean value.

Here is the coding:

b=-3+5\*(rand(1,100));

a=rand(1,1000);

for i=1:1000

sample=randsample(b,100,1);

m=mean(sample);

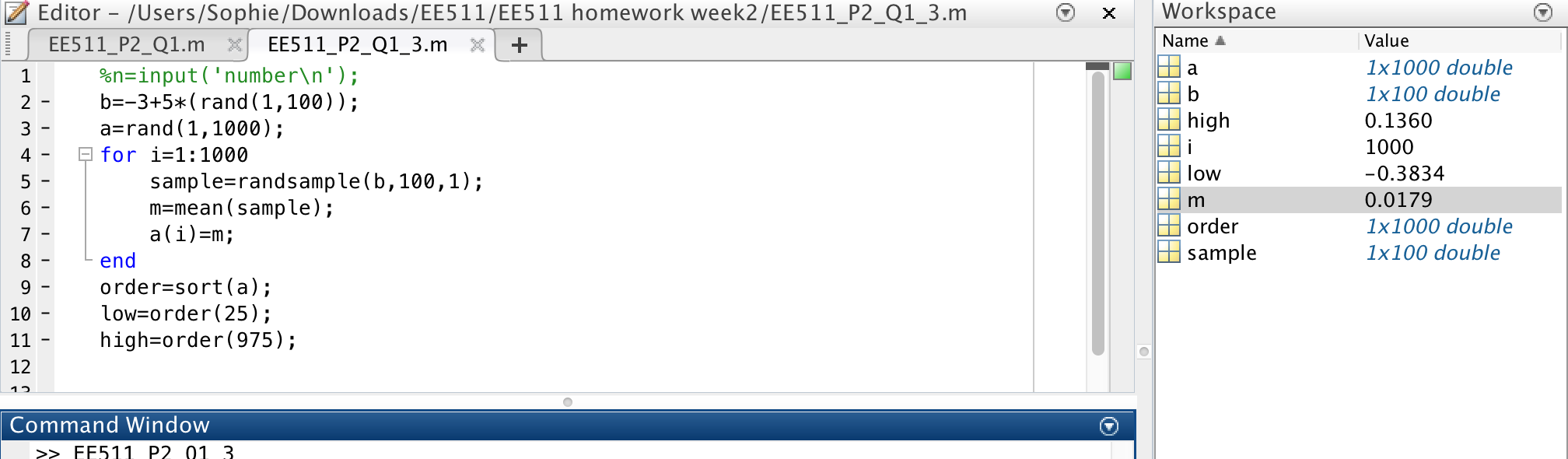
a(i)=m;

end

order=sort(a);

low=order(25);

high=order(975);



1. **Produce a sequence by drawing samples from a standard uniform random variable**.
2. **Compute . Are and uncorrelated? What can you conclude about the independence of and?**

I draw the resample form X, and named it , then I shift right one bit from and add zero in the first bit to generate. Then I calculate the covariance of and . From the outcome, I found the number of very close to 0, in this time, I think they are uncorrelated. But it cannot judge they are independent, because only the and independent, we can conclude they are uncorrelated. Conversely, it is not true.

Here is the coding:

n=input('sample\n');

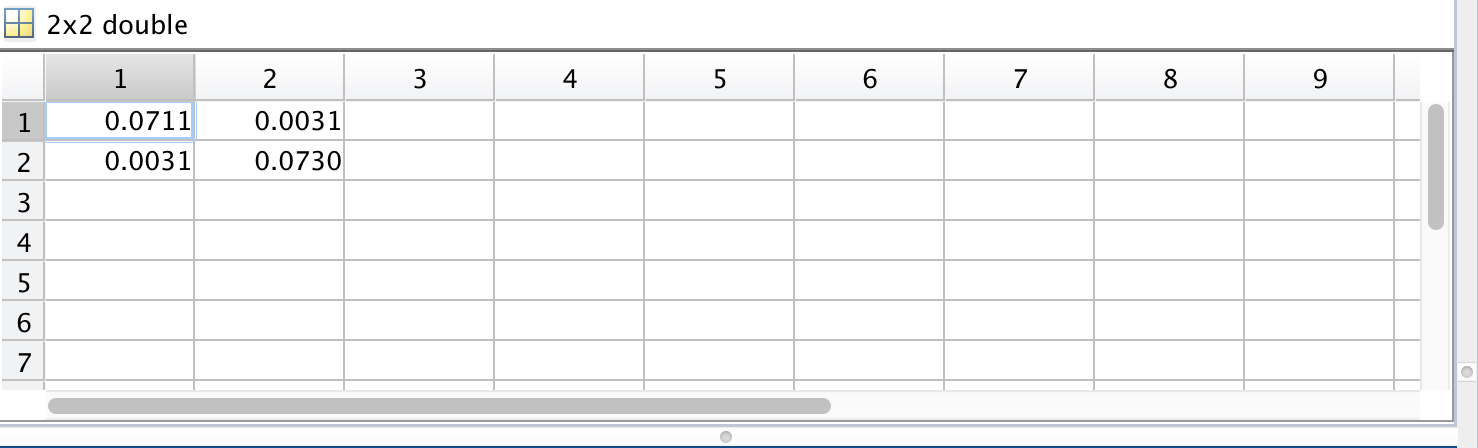
x=rand(1,1000);

x\_1=randsample(x,n,1);

x\_2=[0,x\_1(1:n-1)];

c=cov(x\_1,x\_2);

covariance value:



1. **Compute a new sequence Y where:**

**. Assume X[k]=0 for**

**. Compute . Are and uncorrelated?**

X[k-1],X[k-2],X[k-3] means make a one, two, three bits of right shit of x[k]. And put them in the formula Y to calculate the covariance of X and Y. Even though the number is little bigger than the question a, I still think they are uncorrelated**.**

Here is the coding:

k=input('sample\n');

x=rand(1,10000);

x=randsample(x,k,1);

y\_1=[0,x(1:k-1)];

y\_2=[0,0,x(1:k-2)];

y\_3=[0,0,0,x(1:k-3)];

if k<=0

x(k)=0;

y\_1=x(k-1);

y\_2=x(k-2);

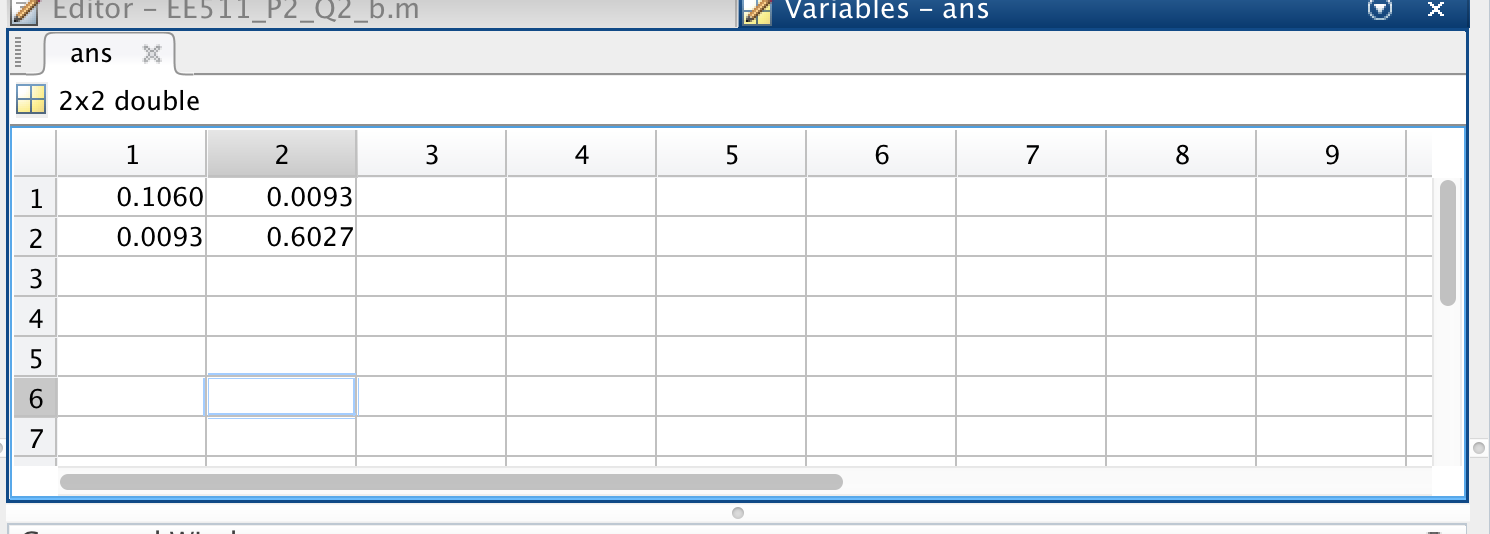
y\_3=x(k-3);

end

y=x(k)-2\*y\_1+0.5\*y\_2-y\_3;

cov(x,y);

covariance value:



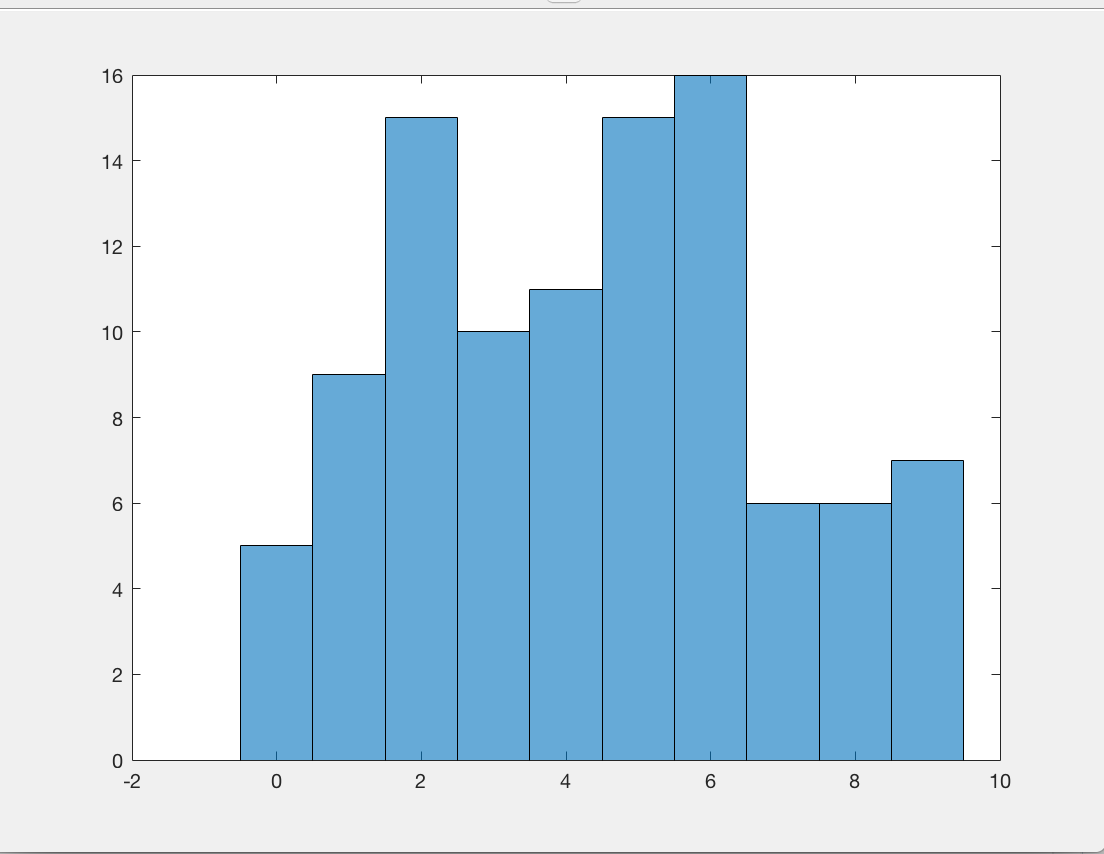
1. **Let M = 10. Simulate (uniform) sampling with replacement from the outcomes 0, 1, 2, 3, ..., M‐1.**
2. **Generate a histogram of the outcomes**.

There are two way to express the integer number the first one is using “randi()”, another is using “round(rand())”.

Here is coding and histogram:

r=round(9\*rand(1,100));

h=histogram(r);



1. **Perform a statistical goodness‐of‐fit test to conclude at the 95% confidence level if your data fits samples from a discrete uniform distribution 0, 1, 2, ..., 9.**

In the MATLAB, I need to record the total number of each sample, and build an array to put the expectation values. Then according to the formula to calculate the threshold and test number. Generally, the test number should be smaller than the threshold number, because the majority area is in the field which is from zero to threshold number.

Here is the coding:

r=round(9\*rand(1,100));

r\_1=rand(1:10);

h=histogram(r\_1);

e=rand(1,10);

for i=1:10

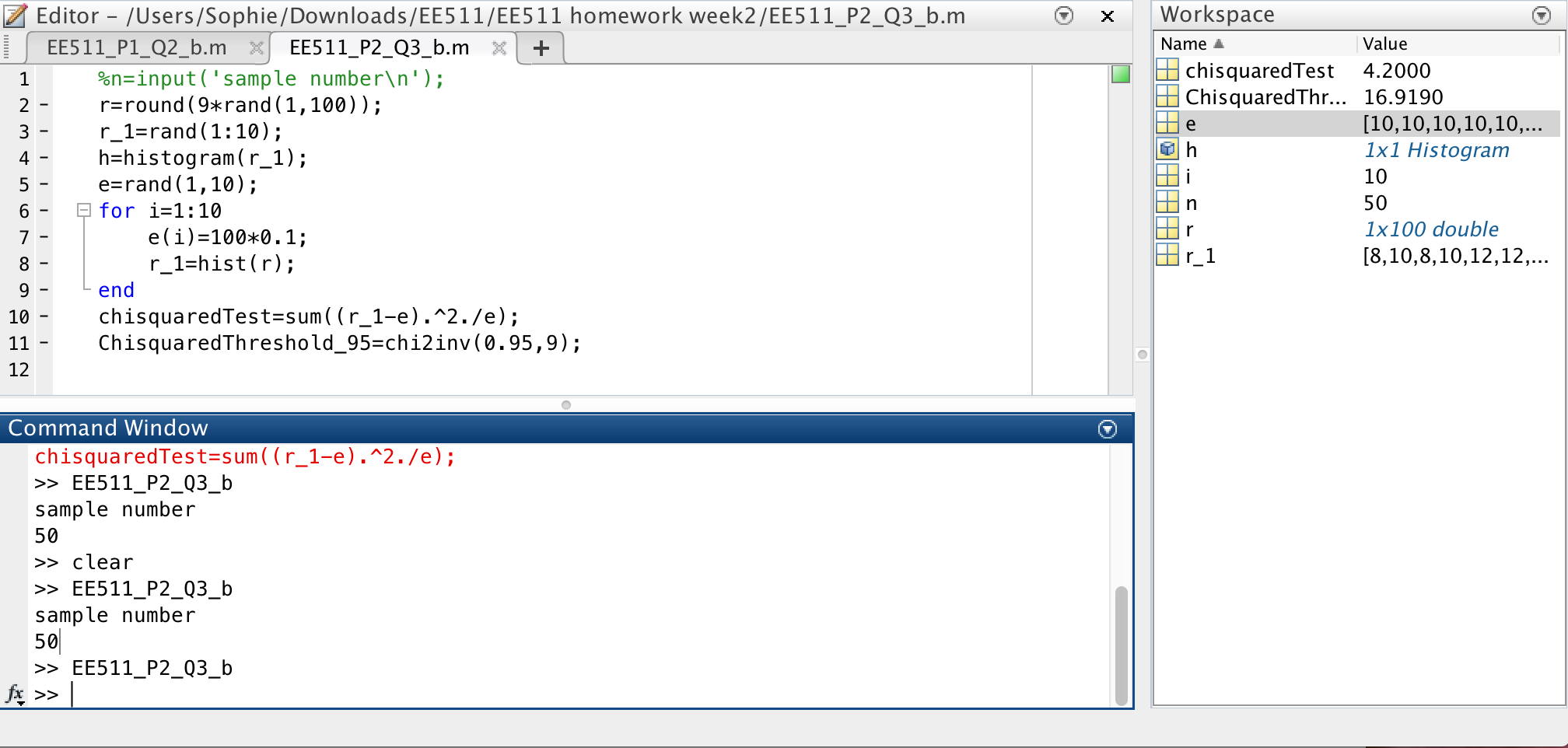
e(i)=100\*0.1;

r\_1=hist(r);

end

chisquaredTest=sum((r\_1-e).^2./e);

ChisquaredThreshold\_95=chi2inv(0.95,9);



1. **Repeat (b) to see if your data (the same data from b) instead fit an alternate uniform distribution 1, 2, 3, ..., 10.**

The different between b and c is we change the discrete uniform from 0 1 2 3 4 5 6 7 8 9 to 1 2 3 4 5 6 7 8 9 10, which means the total number of 0 is 0, and there is no the total number of 10, cause the array is 1-by-10. Then I just make a bit right shift of new array and calculate the “test”, and then I find that the “test” is bigger than the threshold. It means chi-square test is affected by data.

Here is coding:

r=round(9\*rand(1,100));

r\_1=rand(1:10)+1;

h=histogram(r\_1);

e=rand(1,10);

for i=1:10

e(i)=100\*0.1;

r\_1=hist(r);

end

r\_2=[0,r\_1(1:9)]

chisquaredTest=sum((r\_2-e).^2./e);

ChisquaredThreshold\_95=chi2inv(0.95,9);

