ECE 5100 – Probability

Assignment 1

Name: Negib Sherif

ID: 201904422

Date: Feb 2, 2022

Problem 1

1. For this section, we were tasked with importing the data given to us as a CSV file into MATLAB. This was originally done using the csvread function in MATLAB but after reading the official MATLAB documentation where the developers at MATLAB suggest using the readmatrix function, I decided to use the readmatrix function instead. The following is the code used to import the data given to us in MATLAB.

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

-----------------------------------------------------------------------------------------------

Code Sample 1. Importing CSV file

1. For this section, we were tasked with calculating the sample average, variance and standard deviation for each of Q5, Q6, and Q7. The code used to calculate this along with the results can be see below.

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

numOfColm = 3;

colIndex = 1;

while colIndex <= 3

columnData = data(:, colIndex); % gets values of each column

avg = mean(columnData);

variance = var (columnData);

standardD = std (columnData);

disp("average of column " + colIndex + " is: " + avg);

disp("variance of column" + colIndex + " is: " + variance);

disp("standard deviation of " + colIndex + " is: " + standardD);

colIndex = colIndex +1;

end

-----------------------------------------------------------------------------------------------

Code Sample 2. Calculating average, variance and standard deviation

Text

Description automatically generated

Figure 1. Results of average, variance and standard

1. I) For this section, we were asked to make a histogram-type or stem plot that plots the counts for each integer against the heads across the full range of the possible #Heads for a given N. The code used along with the plots obtained can be found below.

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

numOfColm = 3;

colIndex = 1;

increment = 0; %variable to keep track of the increment of values from q5 to q6, and to q7.

while colIndex <= numOfColm

columnData = data(:, colIndex); % gets values of each column

values = 0:1:(10+increment);

counts = sum (columnData == values);

subplot(3, 1, colIndex); % plot for each column

stem (values, counts);

xlabel('Number of Heads'), ylabel('Count'), title('Q' + string(4+colIndex) + ' Results by Counts');

colIndex = colIndex +1;

increment = increment + 5;

end

-----------------------------------------------------------------------------------------------

Code Sample 3. Getting counts vs Number of Heads Plot

Chart

Description automatically generatedFigure 2. Counts vs Number of Heads Plot

II) For this section, we were asked to make a stem plot that plots the relative frequency for each possible head.

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

numOfColm = 3;

Tossers = 52;

colIndex = 1;

increment = 0; %variable to keep track of the increment of values from q5 to q6, and to q7.

while colIndex <= numOfColm

columnData = data(:, colIndex); % gets values of each column

values = 0:1:(10+increment);

counts = sum (columnData == values);

subplot(3, 1, colIndex); % plots each column

stem (values, counts/Tossers);

xlabel('Number of Heads'), ylabel('Rel Frequency'), title('Q' + string(4+colIndex) + ' Results by Relative Frequency for Number of Heads');

colIndex = colIndex +1;

increment = increment + 5;

end

-----------------------------------------------------------------------------------------------

Code Sample 3. Getting Relative Frequency vs Number of Heads Plot

Chart, timeline

Description automatically generatedFigure 3. Relative frequency vs Number of Heads Plot

II) For this section, we were asked to make a stem plot that plots the relative frequency as obtained by individuals in the class.

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

numOfColm = 3;

Tossers = 52;

colIndex = 1;

increment = 0; %variable to keep track of the increment of values from q5 to q6, and to q7.

while colIndex <= numOfColm

columnData = data(:, colIndex); % gets values of each column

values = 0:1:(10+increment);

counts = sum (columnData == values);

subplot(3, 1, colIndex); % plots each column

stem (values/(10+increment), counts/Tossers);

xlabel('Estimated Pr[Head] from ' + string(4+colIndex) + ' Toss Data'), ylabel('Rel Frequency'), title('Q' + string(4+colIndex) + ' Results by Relative Frequency for Number of Heads');

colIndex = colIndex +1;

increment = increment + 5;

end

-----------------------------------------------------------------------------------------------

Code Sample 3. Getting Relative Frequency vs Number of Heads Plot

Chart

Description automatically generated with low confidence

Figure 4. Relative frequency vs Estimated Probability plot

1. For this section, we were tasked with reporting the probability that the estimated Pr[Head] for each N is between 0.4 and 0.6. To figure this out, what I did is look at the graphs in figure 4 and add the relative frequencies for each of the Estimated Pr[head] data points between 0.4 and 0.6. For instance, for Q5, there are three data points between 0.4 and 0.6. Similarly, for Q5, there are 4 data points from 0.4 to 0.6. Lastly, for Q7, there are 5 data points from 0.4 to 0.6. We would expect that the probability, that the estimated Pr[head] is between 0.4 and 0.6, is quite high. And as we would imagine, the results show the same as well. The results can be seen below.

Timeline

Description automatically generatedFigure 5. Plot showing Probability that the estimate Pr[head] is between 0.4 and 0.6

For Q5, the probability that the estimate Pr[head] is between 0.4 and 0.6 is equal to (0.211+0.288+0.192 = 0.691)

For Q6, the probability that the estimate Pr[head] is between 0.4 and 0.6 is equal to (0.135+0.135+0.385+0.135= 0.79)

For Q7, the probability that the estimate Pr[head] is between 0.4 and 0.6 is equal to (0.115 + 0.211 + 0.115 + 0.192 + 0.153 = 0.786)

As predicted these probabilities are pretty high.

1. For this section we were tasked with determining and reporting the sample estimate of Pr[head] using the data from the class for each value of N. The code used for this section along with the results obtained can be found below. As expected the probability is approximately 50% (or 0.5 or half).

-----------------------------------------------------------------------------------------------

data = readmatrix("CoinTossQ5to7RawData.csv");

numOfColm = 3;

Tossers = 52;

colIndex = 1;

increment = 0; %variable to keep track of the increment of values from q5 to q6, and to q7.

while colIndex <= numOfColm

columnData = data(:, colIndex); % gets values of each column

values = 0:1:(10+increment);

counts = sum (columnData == values);

Tossers = 52;

SampleProbHead = sum (columnData) ./ (Tossers \* (10+increment));

disp("Sample probability for Q" + string(4+colIndex) + " is: " + SampleProbHead);

colIndex = colIndex +1;

increment = increment + 5;

end

-----------------------------------------------------------------------------------------------

Text, letter

Description automatically generatedCode Sample 4. Getting Sample Probability of head

Figure 5. Results for sample probability of head