Eric Wan - ezw23@drexel.edu

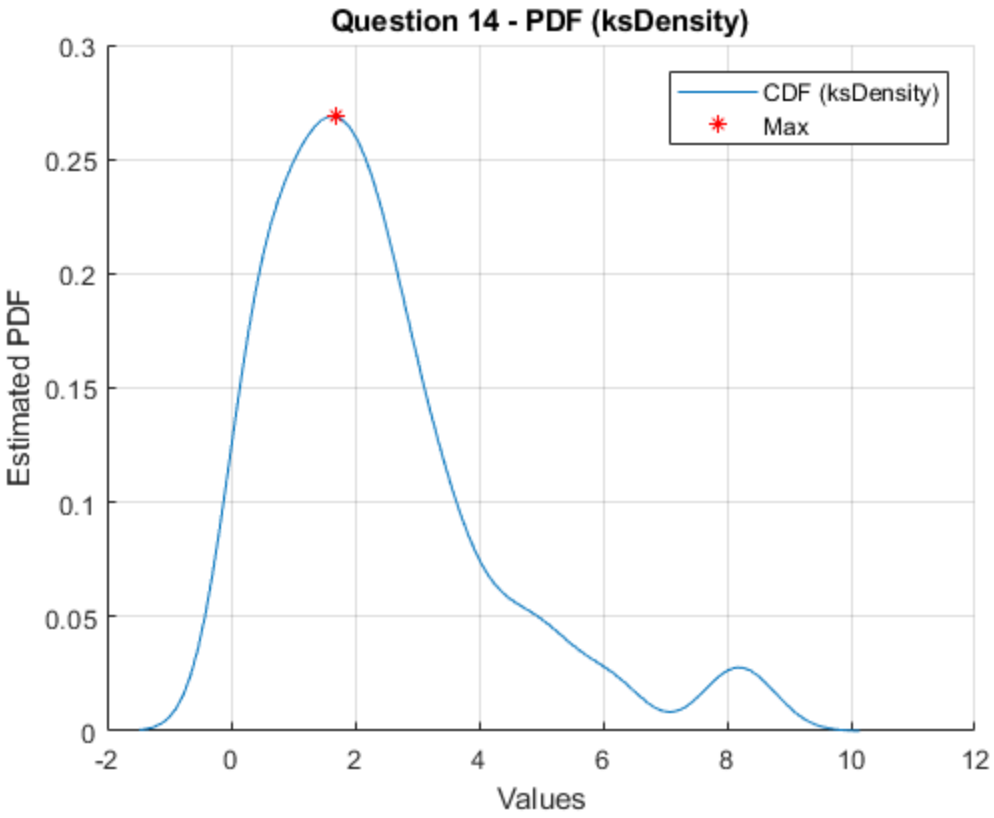
```
clear, clc, warning('on'), close all
```

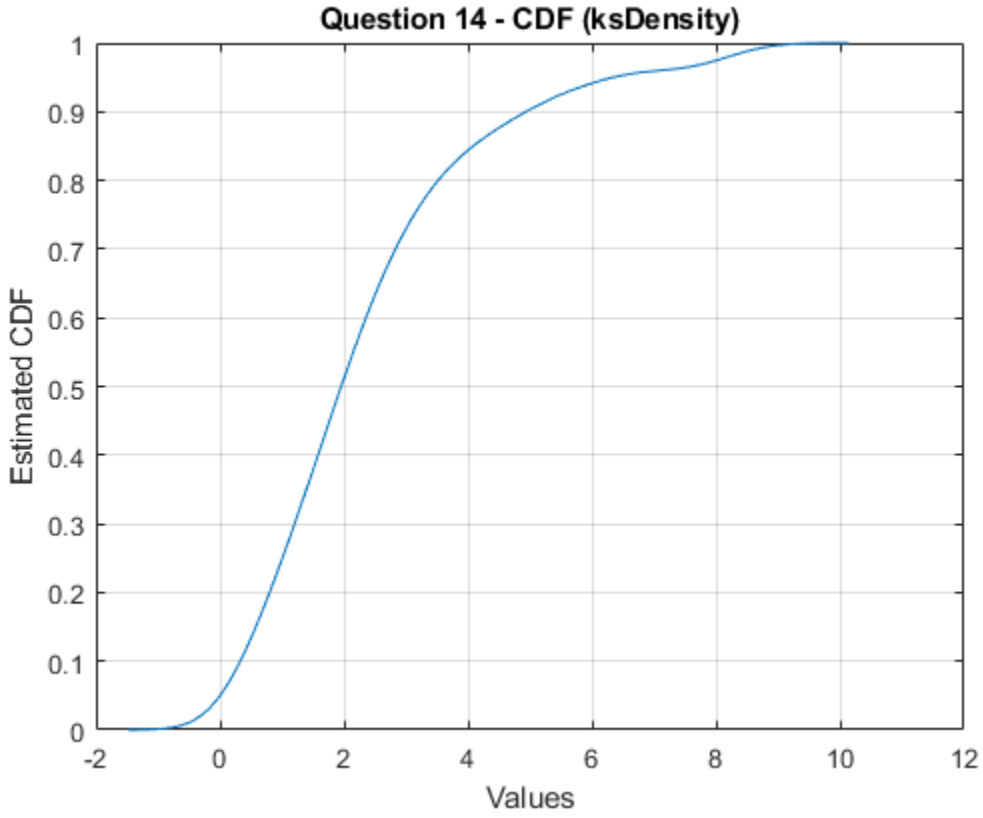
Question 14

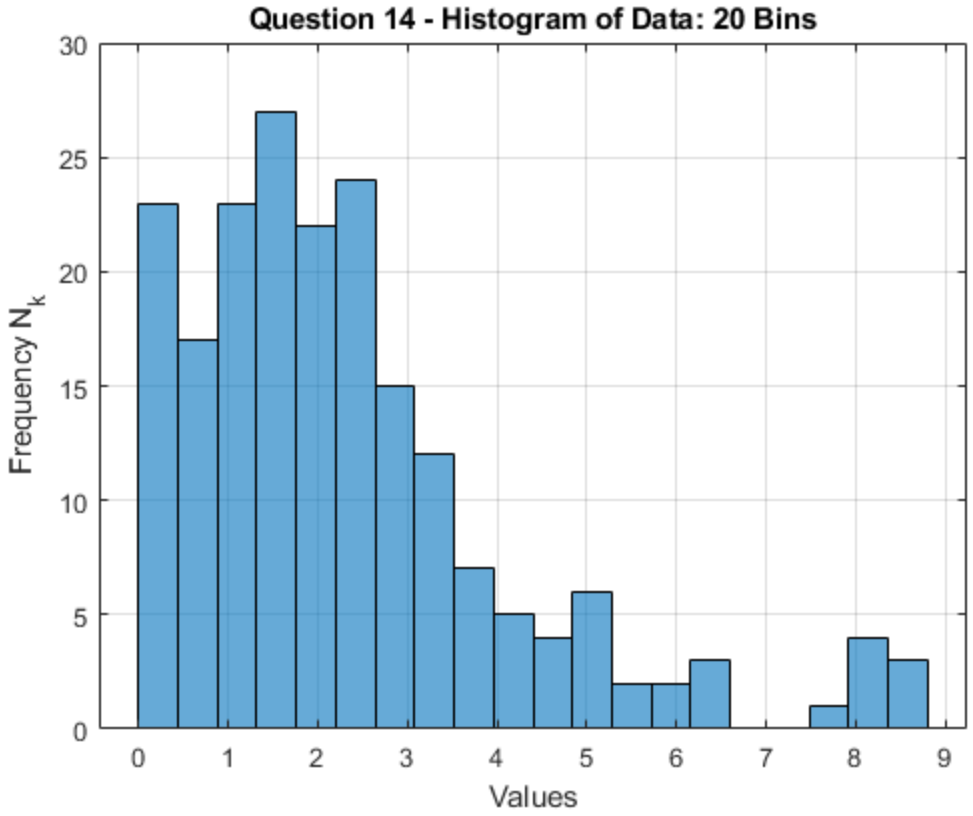
```
%{
14. You are given a data set collected with 200 entries. Estimate the
 pdf,
CDF, mean and variance from the data.  Also determine the location of
 the
peak of the density. Verify the mean from the calculated density using
 the
definition of the mean. Verify your CDF estimate by obtaining the
 histogram
of the data.

There is a file named HW3_data_shankar_Spring. The column headings
 have the
last names of the students. Choose the set meant for you. Sample
 results
are shown. You must produce all the displays containing the
 information
shown. [Hint: Use ksdensity(.) in Matlab. It can generate the
 estimated pdf
and the CDF for any given data. Most of what is needed with this HW
 problem
can be undertaken using the Matlab published document posted in week #
 1
under the Supplementary Materials Tab.]
%}
data = xlsread("HW3_data_shankar_Spring(1).xlsx", 1, "CF:CF");
shank = xlsread("HW3_data_shankar_Spring(1).xlsx", 1, "CL:CL");
datasort = sort(data);

% PDF function using ksDensity
[p, i] = ksdensity(data); % pdf = probability density, i = increment
 over range of data
figure;
grid on;
hold on;
plot(i, p);
plot(i(p == max(p)), p(p == max(p)), 'r*'); % peak point
title("Question 14 - PDF (ksDensity)");
xlabel("Values");
ylabel("Estimated PDF");
legend("CDF (ksDensity)", "Max")

% CDF function using ksDensity
[c, i] = ksdensity(data, 'Function', 'cdf');
figure;
```

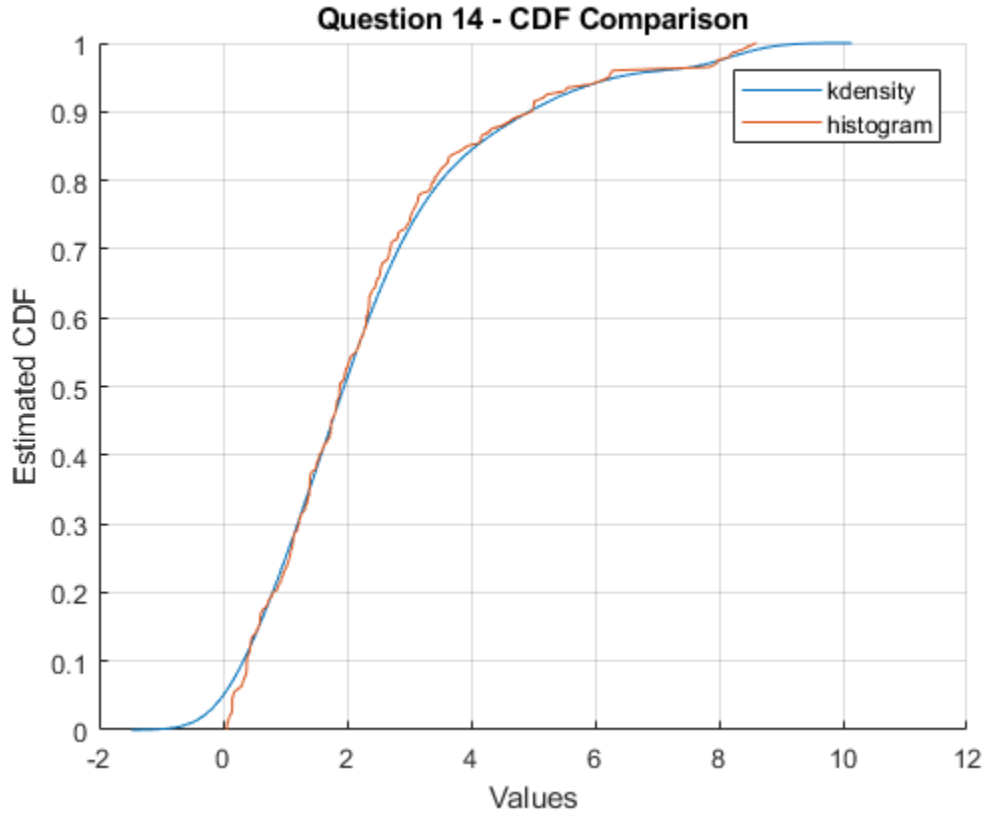
```
plot(i, c);
grid on;
title("Question 14 - CDF (ksDensity)");
xlabel("Values");
ylabel("Estimated CDF");

% Histogram of Data
figure;
histogram(data, 20);
grid on;
title("Question 14 - Histogram of Data: 20 Bins");
xlabel("Values");
ylabel("Frequency N_k");

% CDF function using Histogram
figure;
grid on;
hold on;
plot(i, c);
[c, i] = ecdf(data); % cdf = cumulative distrib, i = increment over
 range of data
plot(i, c);
title("Question 14 - CDF Comparison");
xlabel("Values");
ylabel("Estimated CDF");
legend("kdensity", "histogram");
fprintf("Mean of data is: %.3f, Variance of data is: %.3f,
 Mean(ksDensity) of data is: %.3f\n", mean(data), var(data), mean(i));
```

Mean of data is: 2.340, Variance of data is: 3.478, Mean(ksDensity) of data is: 2.328


Question 14 - PDF (ksDensity)
CDF (ksDensity)
Max
Estimated PDF
Values
0.3
0.25
0.2
0.15
0.1
0.05
0
-2
0
2
4
6
8
10
12


Question 14 - CDF (ksDensity)
Estimated CDF
Values
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-2
0
2
4
6
8
10
12


Question 14 - Histogram of Data: 20 Bins
Frequency N_k
Values
0
5
10
15
20
25
30
0
1
2
3
4
5
6
7
8
9


Question 14 - CDF Comparison
Estimated CDF
Values
kdensity
histogram
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-2
0
2
4
6
8
10
12

Published with MATLAB® R2017b