STATISTICAL PATTERN RECOGNITION & LEARNING
FALL 2016
EXERCISES FOR NON-PARAMETRIC DENSITY ESTIMATION

No.	x_1	x_2	Label
1	0	0	-1
2	1	0	+1
3	5	1	+1
4	5	4	+1
5	1	3	-1
6	2	5	-1
7	4	1	+1
8	5	2	+1
9	3	1	+1
10	1	2	+1
11	0	0	-1
12	0	1	-1
13	2	5	-1
14	4	5	+1
15	1	3	+1
16	0	2	+1
17	4	1	+1
18	1	5	+1
19	3	5	+1
20	1	5	+1

QUESTION 1
From the table above compute the classification of the point (0,0) using the kernel density approach and naive histogram approach. You have to apply MAP first and then solve the same with ML. Set the interval h to 1.

QUESTION 2
From the table above compute the classification of the point (0,0) using the nearest neighbor density approach. You have to apply MAP first and then solve the same with ML. Set the value of K to 3.

QUESTION 3
Repeat the above with nearest neighbor kernel density approach.

QUESTION 4
Apply MAP and ML assuming kernel density approach. Set h = 2. The kernel to be used is given by:

$w(\boldsymbol{u}) = 1$ if $|u_i| <= 3$, for all i
$w(\boldsymbol{u}) = 0$ otherwise

QUESTION 5
Apply MAP and ML assuming kernel density approach. Set h=1. The kernel to be used is given by:

$w(\boldsymbol{u}) = 1-|u|^3$ if $|u_i| <= 3$, for all i
$w(\boldsymbol{u}) = 0$ otherwise

QUESTION 6
Apply MAP and ML assuming kernel density approach. Set h = 2. The kernel to be used is given by:

$w(\boldsymbol{u}) = 1-|u|^2$ if $|u_i| <= 1/2$, for all i
$w(\boldsymbol{u}) = 0$ otherwise

QUESTION 7
Given the following data:

No.	x_1	x_2	x_3	Label
1	0	0	1	-1
2	1	0	3	+1
3	5	1	4	+1
4	5	4	1	+1
5	1	3	2	-1
6	2	5	5	-1
7	4	1	1	+1
8	5	2	4	+1
9	3	1	1	+1
10	1	2	2	+1
11	0	0	3	-1
12	0	1	1	-1
13	2	5	2	-1
14	4	5	3	+1
15	1	3	4	+1
16	0	2	1	+1
17	4	1	5	+1
18	1	5	1	+1
19	3	5	1	+1
20	1	5	2	+1

Apply MAP and ML assuming kernel density approach. The kernel to be used is given by:

$w(\boldsymbol{u}) = 1-|u|^3$ if $|u_i| <= 1$, for all i

$w(\boldsymbol{u}) = 0$ otherwise

The interval vector $\mathbf{h} = [2,4,3]^T$

QUESTION 8
The table below gives the density calculated from training set using the following weighting function:

$w(u) = 1$ if $|u| <= 5$

$w(u) = 0$ otherwise

and the following density function:

$$p(x) = \frac{1}{Nh}\sum_{t=1}^{N} w(\frac{x - x^t}{h})$$

N= number of points, h = 1.

x	p(x)
40	0
36	0
35	5/25
32	5/25
31	5/25
30	15/25
27	15/25
26	15/25
20	10/25
15	0

10	10/25
7	10/25

Suppose we know that we have 25 sample points in the training set. There are only integers in the training set (no floating point values).

a. What can you say about the actual training data? Is it possible to determine a range where the actual points may lie?

b. What would be p(x=5), p(x = 0), p(x= -5)?

SOLUTION (change the font color to see)