CS 540 Homework 1

Name: Halit Erdogan
Login: halit
Section: CS540-001
Date: 9/21/2011

Question 1a

The sample space is:

$$\Omega = \{<1,2>, <1,3>, <1,4>, <1,5>, <2,1>, <2,3>, <2,4>, <2,5>, <3,1>, <3,2>, <3,4>, <3,5>, <4,1>, <4,2>, <4,3>, <4,5>, <5,1>, <5,2>, <5,3>, <5,4>\}$$

Each of the basic outcomes in Ω is equally likely, and thus has probability 1/20. The event of our interest is:

$$E = \{<1,5>, <2,4>, <2,5>, <3,4>, <3,5> <4,2>, <4,3>, <4,5> <5,1>, <5,2>, <5,3>, <5,4>\}.$$

Then,

$$P(E) = |A|/|\Omega| = 12/20 = 3/5$$

Question 1b

I use the notation $f(x)'$ to indicate $\frac{df(x)}{dx}$ for any function f(x).

So the question asks for

$((lnx + 3x)^2 + xe^x)^{'}$
$= (ln^2x + 9x^2 + 6xlnx)^{'} + (xe^x)^{'}$
$= (ln^2x)^{'} + (9x^2)^{'} + (6xlnx)^{'} + (xe^x)^{'}$
$= \frac{2lnx}{x} + 18x + 6(lnx + 1) + e^x + xe^x$
$= \frac{2lnx}{x} + 18x + 6lnx + 6 + e^x + xe^x$

Question 1c

$lim_{x\to0^+} 1/x = \infty$ and $lim_{x\to0^+} lnx = -\infty$; therefore we can apply L'Hopital's rule on $lim_{x\to0^+} \frac{lnx}{1/x}$ to compute $lim_{x\to0^+} xlnx$.
According to the L'Hopital's rule $lim_{x\to0^+} xlnx = lim_{x\to0^+} \frac{lnx'}{(1/x)'}$. Let's compute the derivatives:

$lnx' = 1/x$ and $(1/x)' = -1/x^2$.

Therefore $lim_{x\to0^+} xlnx = lim_{x\to0^+} \frac{1/x}{-1/x^2}$ which is $lim_{x\to0^+} - x = 0$

Question 2A

Iteration	c_1	c_2	c_1 members	c_2 members	c_1 distortion	c_2 distortion	updated c_1	updated c_2
1	(4,2)	(6,2)	(0,0), (3,3), (5,2)	(6,3), (7,0)	6.88	3.41	(2.66,1.66)	(6.5,1.5)
2	(2.66,1.66)	(6.5,1.5)	(0,0), (3,3)	(5,2), (6,3), (7,0)	4.41	4.74	(1.5,1.5)	(6,1.66)
3	(1.5,1.5)	(6,1.66)	(0,0), (3,3)	(5,2), (6,3), (7,0)	4.24	4.33	(1.5,1.5)	(6,1.66)

Question 2B

Iteration	c_1	c_2	c_1 members	c_2 members	c_1 distortion	c_2 distortion	updated c_1	updated c_2
1	(-1,1)	(0,5)	(0,0), (7,0)	(3,3), (5,2), (6,3)	9.47	15.76	(3.5,0)	(4.66,2.66)
2	(3.5,0)	(4.66,2.66)	(0,0), (7,0)	(3,3), (5,2), (6,3)	7	3.81	(3.5,0)	(4.66,2.66)

Question 2C

Different initial center points result in different clustering in k-means algorithm. It effects number of iterations, final center points, and the distortion. Finally we can observe that different initial center points result in different clustering of the points.

Question 3a

$f(x_1) = +,\ f(x_2) = +,\ f(x_3) = +,\ f(x_4) = +,\ f(x_5) = +$. Therefore the 0-1 loss will be 2/5.

Question 3b

$f(x_1) = +,\ f(x_2) = +,\ f(x_3) = +,\ f(x_4) = +,\ f(x_5) = +$. Therefore the 0-1 loss will be 2/5.

Question 3c

$f(x_1) = +,\ f(x_2) = +,\ f(x_3) = +,\ f(x_4) = -,\ f(x_5) = +$. Therefore the 0-1 loss will be 3/5.

Question 3d

$f(x_1) = -,\ f(x_2) = +,\ f(x_3) = +,\ f(x_4) = +,\ f(x_5) = +$. Therefore the 0-1 loss will be 2/5.

Question 3e

$f(x_1) = -,\ f(x_2) = +,\ f(x_3) = -,\ f(x_4) = +,\ f(x_5) = +$. Therefore the 0-1 loss will be 4/5.

Question 3f

No! If we try to minimize the error rate in the training data it may lead the risk of overfitting. In such a case, our model is biased and specific too much for the training data. It causes to make more mistakes for the unseen data. In order to prevent this, test set error rate should be preferred (i.e., one should at least divide the data as training and test. Then train the model without using the test data, and test the resulting model on the test data).