hw 2Xiyu Xie

1 Problem 1

$l(p_1, p_2, ..., p_k) = log(\prod_{t=1}^{N} \prod_{i=1}^{K} p_i^x i) = \sum \sum x_i log p_i$

use Lagrange mutiplier
maximize $l(p_1, p_2, ..., p_k)$
subject to $\sum p_i = 1$
$L(p_1, p_2, ..., p_k, \lambda) = l(p_1, p_2, ..., p_k) + \lambda(\sum p_i - 1)$
By posing all the derivatives to be 0, we get
$\hat(p_i) = \frac{\sum x_i}{N}$

2 Problem 2

$f(\theta|x_1, x_2, ..., x_n) = \frac{1}{(2\pi\sigma^2)^(n/2)} exp(\frac{-1}{2\sigma^2} \sum (x_i - \theta)^2 - \frac{1}{2\sigma_0^2}(\theta - \mu_0)^2)$

$modifyandget \theta(x_1, x_2, ..., x_n) = \frac{\sum x_i \sigma_0^2 + \mu_0 \sigma^2}{\sigma^2 + n * \sigma_0^2}$

3 Problem 3

$P(M) = 1\%$
$P(B) = 30\%$
$P(O) = 69\%$

$P(+|M) = 95\%$
$P(-|M) = 5\%$

$P(-|\ M) = 95\%$
$P(\ M|-) = 97.5\%$

$P(M|-) = 2.5\%$

$P(-) = P(-|M) * P(M) + P(-|\ M) * P(\ M) = 0.05 * 0.01 + 0.95 * 0.99 = 0.941$ $P(+) = 1 - P(-) = 0.059$ or $P(-) = P(-|\ M) * P(\ M)/P(\ M|-) = 0.95 * 0.99/0.975 = 0.965$
$P(+) = 1 - P(-) = 0.035$

$P(M|+) = P(+|M) * P(M)/P(+) = 0.95 * 0.01/0.059 = 0.16 or 0.27$

4 Problem 4

(c)0.9225352 (d)0.9225352 (e)0.7605634 decreased by overfitting

5 Problem 5

(b)Euclidean is better than manhattan (c)not accuracy with k¿=5