Homework 2 (Due: 10/3)

For problem 1 and 2, scan your answer and submit them on i-campus. For problem 3, submit your R script file on i-campus (please only include R function).

1. In the augmented category example, Compute approximated $I(\theta)$ from original data. Also compute approximated $I_o(\theta)$ by computing $I_c(\theta)$ and $I_m(\theta)$ from augmented data.

2. Derive the EM algorihtm to compute the MLE of $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p)$ for the mixture model $X \sim pN(\mu_1, \sigma_1^2) + (1-p)N(\mu_2, \sigma_2^2)$.

3. Using the result of Question 2, make **R** function to find the MLE. For your function, the input argument should contain one-dimensional data vector and the vector of the initial value of $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p)$. For the stopping criterion for your EM algorithm, if the difference of the consecutive log-likelihood values is less than 10^{-10}, stop your algorithm. I expect your function works as follows:

```
> # Data generation
> n=500;p=0.3;mu1=-2;mu2=2;sigma1=3;sigma2=1
> x1=rnorm(n,mu1,sqrt(sigma1))
> x2=rnorm(n,mu2,sqrt(sigma2))
> z=rbinom(n,1,p)
> x=z*x1+(1-z)*x2
>
> result=Kim8(x,c(-0.5,0.5,1,1,0.1))
EM converges in 144 steps
> result$param
[1] -1.7536370  1.9713058  3.2629705  0.7765643  0.3044439
```