(Part I of the final exam. The paper work must be submitted between 10:00 am - 11:59 am of July 3rd, and the problem of the oral test will be given during the mentioned shedule.)

(1) Assume thay $(Y_1, Y_2, Y_3, Y_4) \sim$ multinomial with $(0.5 - 0.5\theta, 0.25\theta, 0.25\theta, 0.5)$. When the incomplete data $(y_1, y_2, y_3 + y_4) = (38, 34, 125)$ are observed, please use EM algorithm to estimate θ with the initial value $=0.5$.

Assume that E-Step is replaced by MC step (Monte Carlo integration). Please compare the convergence property with the previous result, and whether the monotonical property for log-likelihood function still holds.

(2) Assume that $X_1, X_2, ..., X_n \sim^{i.i.d} \pi(x|\theta)$, where $\pi(x|\theta) = \sum_{j=1}^{3} p_j \phi_{\sigma_j}(x - \mu_j)$, ϕ is the p.d.f of the standard normal distribution and $\sum_{j=1}^{3} p_j = 1$. Please estimate the unknown parameter $(\mu_i, \sigma_j, p_j)_{j=1}^{3}$ by EM algorithm. (Hint: introduce the latent variable Z_i to present the label information. For example, $Z_i = j$ if the ith observation (X_i) comes from the jth component.)