#1.

- 用 MC-step (Monte Carlo)取代 E-step.

由於

$$Q(\theta|\theta^{(n)}) = E_{\theta^{(n)}}(\log f(\boldsymbol{Y}|\theta)|\boldsymbol{x}) \quad ,where\ \ \boldsymbol{x} = (y_1, y_2, y_3 + y_4) = (38,34,125)$$

$$= E_{\theta^{(n)}}(\log h(\boldsymbol{Y})|\boldsymbol{x}) + 38\log\left(\frac{1}{2} - \frac{\theta}{2}\right) + 34\log\left(\frac{\theta}{4}\right) + \log(\frac{\theta}{4})E_{\theta^{(n)}}(\boldsymbol{Y_3}|\boldsymbol{x})$$

因此我們要用 Monte Carlo 估計$E_{\theta^{(n)}}(\boldsymbol{Y_3}|\boldsymbol{x}) = \int_0^{125} y f_{Y_3|\boldsymbol{x}}(y)dy,$

$\Rightarrow$ 1. 生成$U_1, U_2, \dots, U_k \sim U(0,1)$

2. 令$T_i = 125U_i \ \sim U(0,125)\ \ , i = 1, \dots, k$

3. 估計量 $\hat{I} = \frac{1}{k}\sum_{i=1}^{k}\frac{T_i f_{Y_3|\boldsymbol{x}}(T_i)}{f_T(T_i)}$

- M-step

$\max_\theta Q(\theta|\theta^{(n)}) \Leftrightarrow \max_\theta W(\theta|\theta^{(n)})$,

$$where\ \ W(\theta|\theta^{(n)}) = 38\log\left(\frac{1}{2} - \frac{\theta}{2}\right) + 34\log\left(\frac{\theta}{4}\right) + \log\left(\frac{\theta}{4}\right)\hat{I}$$

這裡用 Newton Method 來求$W(\theta|\theta^{(n)})$最大值

$\Rightarrow\ \ \theta_{m+1} = \theta_m - \frac{w'(\theta|\theta^{(n)})}{w''(\theta|\theta^{(n)})}$,直到滿足$|\theta_{m+1} - \theta_m| < 10^{-5}$.

$\Rightarrow\ \ \theta^{(n+1)} = \arg\max_\theta W(\theta|\theta^{(n)})$

- ***Simulation:***

Initial value $\theta^{(0)} = 0.5$,stopping rule: $|\theta^{(n+1)} - \theta^{(n)}| < 10^{-7}$.

Case1:

在 MC-step 時，取 k=1000，發現很難達成停止條件 $|\theta^{(n+1)} - \theta^{(n)}| < 10^{-7}$.

因為 Monte Carlo 估計$E_{\theta^{(n)}}(\boldsymbol{Y_3}|\boldsymbol{x})$本身就存在誤差，所以考慮透過增加樣本數方式減少誤差。

Case2:

在 MC-step 時，第 n 步取 $k^{(n+1)} = k^{(n)} + 1000$,$k^{(0)} = 0$，

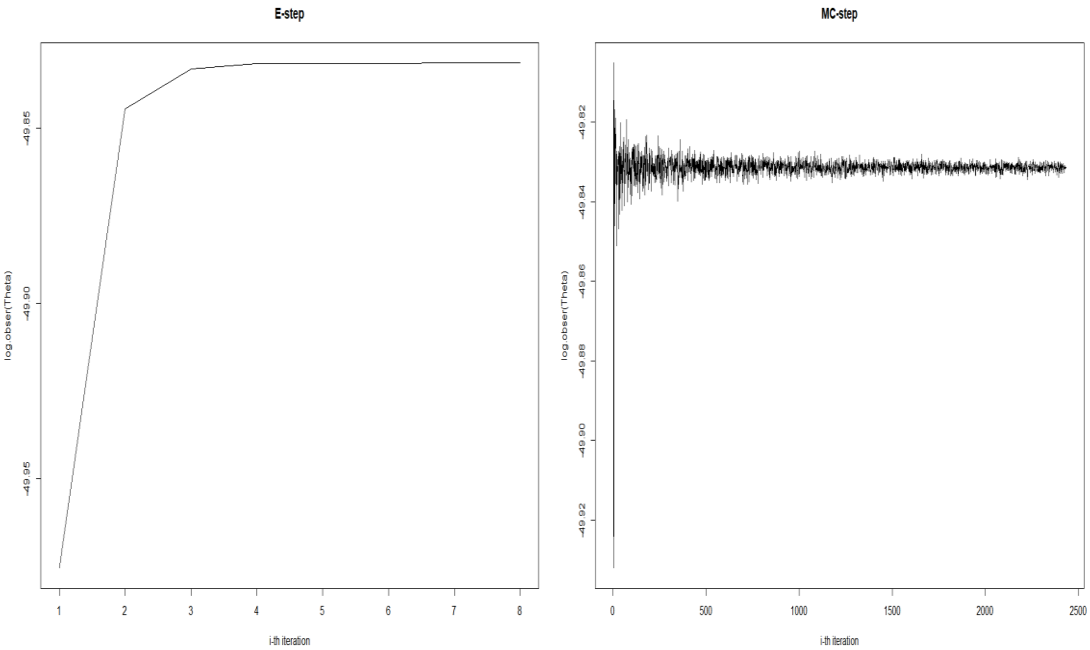
則當 k=2434000，滿足停止條件 $|\theta^{(n+1)} - \theta^{(n)}| < 10^{-7}$, $\hat{\theta}_{MC} = 0.6267990521$.

Note:

Observed $\mathbf{x} = (y_1, y_2, y_3 + y_4) \ \Leftrightarrow\ \mathbf{x} = (y_1, y_2)$,$n = y_1 + y_2 + y_3 + y_4 = 197$

$\Rightarrow\ \mathbf{X} = (Y_1, Y_2) \sim \text{multinomial}(72, \frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2})$

$\Rightarrow$ loglikelihood function: $\log f_{\mathbf{X}}(\mathbf{x}|\theta^{(n)}) = \log f_{Y_1,Y_2}(y_1, y_2|\theta^{(n)})$



左圖為 E-step, y 軸為$\log f_{\boldsymbol{X}}(\mathbf{x}|\theta^{(n)})$, x 軸為第 n 次疊代.

右圖為 MC-step, y 軸為$\log f_{\boldsymbol{X}}(\mathbf{x}|\theta^{(n)})$, x 軸為第 n 次疊代.

觀察到我們用 MC-step 取代 E-step 時，它的值由於 monte carlo method 關係上下跳動，故無法保持理論上的性質: monotonical property for log-likelihood function.

然而，隨著 monte carlo method 中 k 的增加(用愈多的隨機變數去估計)，可以發現它是會收斂的，

$\hat{\theta}_{MC} = 0.6267990521$ (c.f. $\hat{\theta}_{E} = 0.6268214841$)