# Finite Gaussian Mixture with unknown parameters and unknown membership probability

**0. Reference**

- 연세대학교 임종호 교수님의 missing data analysis 강의

**1. Introduction**

확률변수 $Y_1, \cdots, Y_n$ 가 주어졌을 때, 이들의 모집단에 대해서 생각해보자. 가장 쉽게 생각할 수 있는 분포는 정규분포이다.

$$Y_1, \cdots, Y_n \sim Normal$$

만약 $Y_i$ 가 정규분포를 따른다고 가정한다면, 데이터의 형태를 정규분포의 특성에 따라 unimodal로 가정하는 것이고 이를 characterize하는 두 개의 모수 $\mu, \sigma^2$ 를 찾을 것이다. 하지만 만약 $Y_i$ 가 대한민국 남여 13세에서 50세 국민의 신장이라고 가정하자. 이러한 데이터에 대해서 하나의 모집단으로 데이터를 잘 설명할 수 있을까?
확률변수 $Y_i$ 의 실현 값(realized value)을 $y_i$ 라고 하자. $y_1$ 은 25세 남성의 신장, $y_2$ 는 45세 여성의 신장이라고 한다면, $y_1, y_2$ 는 전혀 다른 특성을 가지고 있다. 이와 같이 어떤 데이터가 하나의 집단이라기 보다, 여러 집단에서 비롯되었다고 보는 것이 합당할 때가 있다. 이는 EDA를 통해 파악할 수도 있고 이질적인 집단을 파악하는 대표적인 ML 기법인 클러스터링을 통해서 확인할 수도 있다. 그렇다면 다른 ML 기법에 비해서 Gaussian Mixture을 함으로써 얻는 이득은 무엇일까? 아래와 같이 정리해보았다.

- 집단의 특징을 분포의 추정된 모수를 통해 확인할 수 있다. 클러스터링은 각 클러스터의 변수 평균, 중앙값, 최빈값 등 summary statistics을 통해 이질성을 확인한다. 이에 대해서 표준 오차(Standard Error)을 구할 수 없기 때문에 통계적 추론을 할 수 없다. 하지만 GM을 통해 얻은 모수 추정치는 표준 오차를 구할 수 있기 때문에 이를 기반으로 도출한 모수 추정치에 대해 신뢰구간이나 가설 검정을 해서 uncertainty을 control 할 수 있다.
- 클러스터링은 각 집단에 대한 확률의 개념이 없다. 즉, 어떤 새로운 데이터가 어느 집단에서 가장 잘 나타날 것인가에 대해서 답변을 해주지 못한다. 하지만 GM은 membership probability 라는 개념을 통해 어떤 집단이 dominant한지 확률의 개념을 통해서 알려준다.

이제 $Y_i$ 를 GM의 시각에서 바라보자.

**2. Model**

먼저 아래와 같이 notation을 정의한다.

- Assume following Gaussian mixture model

$$Y_1, \cdots, Y_n \overset{iid}{\sim} \pi_1 f_1(y;\, \mu_1, \sigma^2) + \pi_2 f_2(y;\, \mu_2, \sigma^2) + \cdots + \pi_K f_K(y;\, \mu_K, \sigma^2) = \sum_{k=1}^{K} \pi_k f_k(y;\, \mu_k, \sigma^2)$$

- Let

$$\eta = (\theta, \gamma),\ \theta = (\pi_1, \cdots, \pi_{K-1}),\ \gamma = (\gamma_1, \cdots, \gamma_K),\ \gamma_k = (\mu_k, \sigma^2)$$

  여기서 $\pi_K = 1 - \sum_{k=1}^{K-1} \pi_k$ 이고 common variance $\sigma^2$ 을 가정한다.

모수는 $\eta = (\theta, \gamma)$ 이므로 이에 대해서 likelihood function을 세운다.

$$\begin{aligned} log\, L_{obs}(\eta) &= \sum_{i=1}^{n} log\, f(y_i; \eta) \\ &= \sum_{i=1}^{n} log \left\{ \sum_{k=1}^{K} \pi_k f_k(y_i; \gamma_k) \right\} \end{aligned} \tag{1}$$

membership probability인 $\pi_k$ 와 관련된 확률 분포는 multinomial distribution이다. 즉, 아래와 같이 indicator function을 정의하자.

$$z_{ik} = \begin{cases} 1 & if\ unit\ i\ belongs\ to\ kth\ group\ with\ probability\ \pi_k \\ 0 & o.w \end{cases} \tag{2}$$

(1) 에서 정의한 likelihood function은 observed likelihood이다. 즉, 우리가 관측한 데이터는 $y_i$ 뿐인데, $z_{ik}$ 전부를 missing data라고 보는 것이다. 따라서 (1) 은 $z_{ik}$ 는 포함되지 않고 우리가 관측한 데이터에만 의존하는 observed likelihood이다. 모수 $\eta$ 에 대한 MLE를 구하기 위해서 complete likelihood을 세워보자.

$$\begin{aligned} logL_{com}(\theta) &= log \prod_{i=1}^{n} \prod_{k=1}^{K} \pi_k^{z_{ik}} f_k(y_i)^{z_{ik}} \\ &= \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} log\pi_k + \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} log f_k(y_i) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} log\pi_k + \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} \left\{ -\frac{1}{2} log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_i - \mu_k)^2 \right\} \end{aligned} \tag{3}$$

그런데 EM 알고리즘의 E-step에서 정의하는 조건부 기대값은 $Q\left(\eta \mid \eta^{(t)}\right) = E\left\{ logf(\boldsymbol{y}, \boldsymbol{\delta}; \eta) \mid \boldsymbol{y}_{obs}, \boldsymbol{\delta}, \eta^{(t)} \right\}$ 이므로

$$\begin{aligned} Q\left(\eta \mid \eta^{(t)}\right) &= E\left\{ logf(\boldsymbol{y}, \boldsymbol{\delta}; \eta) \mid \boldsymbol{y}_{obs}, \boldsymbol{\delta}, \eta^{(t)} \right\} \\ &= E\left\{ logL_{com}(\theta) \mid \boldsymbol{y}_{obs}, \boldsymbol{\delta}, \eta^{(t)} \right\} \\ &= E\left\{ \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} log\pi_k + \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} \left\{ -\frac{1}{2} log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_i - \mu_k)^2 \right\} \mid \boldsymbol{y}_{obs}, \boldsymbol{\delta}, \eta^{(t)} \right\} \\ &= \sum_{i=1}^{n} \sum_{k=1}^{K} E\left\{ z_{ik} \mid y_i, \eta^{(t)} \right\} log\pi_k^{(t)} + \sum_{i=1}^{n} \sum_{k=1}^{K} E\left\{ z_{ik} \mid y_i, \eta^{(t)} \right\} \left\{ -\frac{1}{2} \left( log\sigma^2 + \frac{1}{\sigma^2}(y_i - \mu_k)^2 \right) \right\} + C \end{aligned} \tag{4}$$

그런데 여기서

$$
\begin{aligned}
E\left\{z_{ik} \mid y_i, \eta^{(t)}\right\} &= P\left(z_{ik}=1 \mid y_i, \eta^{(t)}\right) \\
&= \frac{P\left(z_{ik}=1, y_i, \eta^{(t)}\right)}{f(y_i, \eta^{(t)})} \\
&= \frac{P\left(z_{ik}=1\right) f\left(y_i \mid \eta^{(t)}, z_{ik}=1\right)}{\sum_{z_{ik}} f\left(y_i, z_{ik}, \eta^{(t)}\right)} \\
&\overset{let}{\equiv} z_{ik}^{(t)}
\end{aligned}
$$

현재 모수는 $\pi_k, \mu_k, \sigma^2$ 이므로 (4) 의 $Q\left(\eta \mid \eta^{(t)}\right)$ 를 maximize하는 값들을 해당 모수에 대해서 찾는 것이 곧 MLE를 찾는 것이므로 $t$번째 모수 추정치가 주어졌을 때, $Q\left(\eta \mid \eta^{(t)}\right)$ 을 maximize하는 $t+1$ 번째 모수 추정치를 반복적으로 찾는다.

- $\pi_k^{(t+1)}$

  $Q\left(\eta \mid \eta^{(t)}\right)$ 를 maximize하는 $\pi_k$ 를 찾되, $\sum_{k=1}^{K} \pi_k = 1$ 이라는 제약 조건을 Lagrange multiplier 을 통해 두면

$$
L(\pi_k, \lambda) = \sum_{k=1}^{K} \sum_{i=1}^{n} z_{ik}^{(t)} log\pi_k + \lambda\left(1 - \sum_{k=1}^{K} \pi_k\right)
$$

$$
\therefore \frac{\partial}{\partial \pi_k} L(\pi_k, \lambda) = \frac{\partial}{\partial \pi_k} \sum_{k=1}^{K} \sum_{i=1}^{n} z_{ik}^{(t)} log\pi_k + \frac{\partial}{\partial \pi_k} \lambda\left(1 - \sum_{k=1}^{K} \pi_k\right) = 0
$$

$$
\sum_{i=1}^{n} \frac{z_{ik}^{(t)}}{\pi_k} - \lambda = 0
$$

$$
\therefore \pi_k^{(t+1)} = \frac{\sum_{i=1}^{n} z_{ik}^{(t)}}{\lambda}
$$

  그런데 여기서

$$
\begin{aligned}
&\sum_{i=1}^{n} \frac{z_{ik}^{(t)}}{\pi_k} - \lambda = 0 \\
\Longleftrightarrow &\sum_{k=1}^{K} \sum_{i=1}^{n} z_{ik}^{(t)} = \lambda \sum_{k=1}^{K} \pi_k \\
\Longleftrightarrow &\sum_{i=1}^{n} 1 = \lambda \\
&\therefore \lambda = n \\
&\therefore \pi_k^{(t+1)} = \frac{\sum_{i=1}^{n} z_{ik}^{(t)}}{n}
\end{aligned}
$$

- $\mu_k^{(t+1)}$

$$L(\mu_k) = \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}\left\{-\frac{1}{2}\left(log\sigma^2 + \frac{1}{\sigma^2}(y_i - \mu_k)^2\right)\right\}$$

$$\frac{\partial}{\partial \mu_k}L(\mu_k) = \frac{1}{\sigma^2}\sum_{i=1}^{n} z_{ik}^{(t)}(y_i - \mu_k) = 0$$

$$\sum_{i=1}^{n} z_{ik}^{(t)} y_i = \sum_{i=1}^{n} z_{ik}^{(t)} \mu_k$$

$$\therefore \mu_k^{(t+1)} = \frac{\sum_{i=1}^{n} z_{ik}^{(t)} y_i}{\sum_{i=1}^{n} z_{ik}^{(t)}}$$

- $\sigma^{2(t+1)}$

$$L(\sigma^2) = \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}\left\{-\frac{1}{2}\left(log\sigma^2 + \frac{1}{\sigma^2}(y_i - \mu_k)^2\right)\right\}$$

$$\frac{\partial}{\partial \sigma^2}L(\sigma^2) = -\frac{1}{2}\left\{\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}/\sigma^2 - \frac{1}{\sigma^4}\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}(y_i - \mu_k)^2\right\} = 0$$

$$\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}/\sigma^2 = \sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}(y_i - \mu_k)^2/\sigma^4$$

$$\therefore \sigma^{2(t+1)} = \frac{\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}(y_i - \mu_k^{(t+1)})^2}{\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}} = \frac{\sum_{i=1}^{n}\sum_{k=1}^{K} z_{ik}^{(t)}(y_i - \mu_k^{(t+1)})^2}{n}$$

이상으로 GM에서 latent variable $z_{ik}$ 을 가정함으로써 $\eta$에 대한 MLE을 구했다.
$Q\left(\eta \mid \eta^{(t)}\right) = E\left\{logf(\boldsymbol{y}, \boldsymbol{\delta}; \eta) \mid \boldsymbol{y}_{obs}, \boldsymbol{\delta}, \eta^{(t)}\right\}$ 을 maximize하는 $\eta^{(t+1)}$ 을 찾는 것은 $\bar{S}(\eta) = E\left\{S_{com}(\eta) \mid \mathbf{y}_{obs}, \boldsymbol{\delta}; \eta^{(t)}\right\} = 0$ 의 해를 찾는 것과 동일하다. 따라서 위에서 구한 final estimates인 $\hat{\eta} = \left(\hat{\pi}_k, \hat{\mu}_k, \hat{\sigma}^2\right)$ 을 $\bar{S}(\eta)$ 에 대입하면 근사적으로 0에 가까워야 제대로 추정치를 구했다는 의미이다. 이를 아래와 같이 확인해볼 것이다. 즉, $\eta$에 대한 Mean Score Function인 $\bar{S}(\eta) = E\left\{S_{com}(\eta) \mid \mathbf{y}_{obs}, \boldsymbol{\delta}; \eta^{(t)}\right\} = E\left\{\frac{\partial}{\partial \eta} log\, L_{obs}(\eta) \mid \mathbf{y}_{obs}, \boldsymbol{\delta}; \eta^{(t)}\right\}$ 에 각 모수의 final estimates을 넣어봄으로써 Mean Score Function의 estimates가 0에 가까운지 확인해본다.

**3. Experiments**

간단하게 데이터를 R에서 모수를 설정하여 generate 한 후에, 이 데이터에 GM을 적용하여 실제 모수와 유사한 추정치를 내는지 확인해본다.

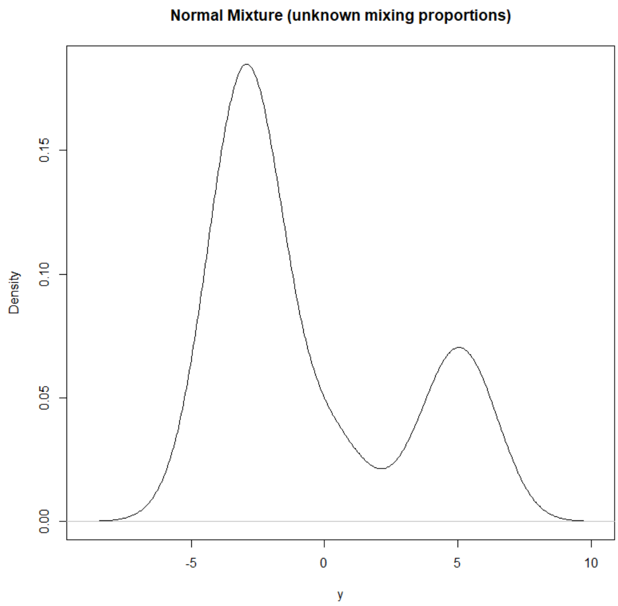
**Step 1: Iterate EM Algorithm to derive final Estimates**

모수를 아래와 같이 설정하였다.

$$\pi = (0.1, 0.2, 0.7)$$

$$f_1(y;\, 0, 1),\ f_2(y;\, 5, 1),\ f_3(y;\, -3, 1)$$

의도한바는 세 개의 봉우리를 가지도록 하였으나 막상 만들어보니 아래와 같이 bimodal의 형태를 띄었다.




세 개의 분포에서 생성하였으므로 우선 $K = 3$ 으로 GM을 할 것이지만 당장 보이는 분포가 bimodal 이므로 결과가 좋을지 확신이 서지는 않는다.
membership probability의 초깃값, $\pi^{(0)} = (\pi_1^{(0)}, \pi_2^{(0)}, \pi_3^{(0)})$ 에 따른 최종 수렴 값을 아래와 같이 정리하였다.

| | $\pi^{(0)} = (0.33, 0.33, 0.34)$ | $\pi^{(0)} = (0.2, 0.2, 0.6)$ |
|---|---|---|
| $\pi_1^{(t)}$ | 0.357 | 0.132 |
| $\pi_2^{(t)}$ | 0.263 | 0.247 |
| $\pi_3^{(t)}$ | 0.380 | 0.621 |

$\pi^{(0)} = (0.33, 0.33, 0.34)$ 로 초깃값을 설정하였을 때는, 모수로 설정한 $(0.1, 0.2, 0.7)$ 과 다소 거리가 있는 결과를 보였다. 사실 EDA로 대략 두 개의 bimodal, 그 중에서도 음수쪽에 봉우리가 큰 형태를 인지한다면 초깃값을 $\pi^{(0)} = (0.33, 0.33, 0.34)$ 와 같이 동등하게 부여하는 것이 아니라 $\pi^{(0)} = (0.2, 0.2, 0.6)$ 와 같이 어느 정도 치우치게 설정하는 것이 합리적이다. 그에 따른 결과도 모수 $(0.1, 0.2, 0.7)$ 와 나름 가깝게 구해졌다.
다음으로 membership probability의 초깃값은 $(0.2, 0.2, 0.6)$ 으로 고정하고 $\mu$의 초깃값을 아래와 같이 변화하며 결과를 살펴보았다.

| | $\mu^{(0)} = (0, 3, -2),\ \sigma^{2(0)} = 2$ | $\mu^{(0)} = (0, 1, 2),\ \sigma^{2(0)} = 2$ |
|---|---|---|
| $\mu_1^{(t)}$ | 0.005 | -2.676 |
| $\mu_2^{(t)}$ | 5.024 | -2.671 |
| $\mu_3^{(t)}$ | -3.049 | 4.617 |
| $\sigma^{2(t)}$ | 0.88 | 1.817 |

결과는 꽤 흥미롭다. 사실, 데이터를 생성하는 모수를 모르는 현실에서는 초깃값을 $\mu^{(0)} = (0, 3, -2)$ 와 같이 설정하지 않을 것이다. 애초에 bimodal 모양이기 때문이다. 하지만 여기서는 모수를 알고 있기 때문에 그 모수를 잘 찾아가는지 알아보고 싶어서 초깃값으로 $\mu^{(0)} = (0, 3, -2)$ 을 설정해보았다. 실제 모수는 $(0, 5, -3)$ 인데 이와 매우 비슷한 최종 추정치가 나왔다.
다음은 좀 더 현실적인 초깃값 $\mu^{(0)} = (0, 1, 2)$ 을 살펴보자. 사실 이것도 EDA를 안 한 사람의 초깃값일 것이다. EDA를 한다면 분명 bimodal이 음수, 양수 양쪽에 있기 때문이다. 어찌됐든, 아주 niave한 추정치를 시도해보면 결과는 두 모평균에 대한 추정치는 -2.6과 유사하고 나머지는 4.6이다. 이는 위 그림에서도 알 수 있듯이 이 데이터가 -3, 5 부근에서 bimodal 형태를 보이기 때문에 이 값과 비슷하게 수렴이 된 것으로 보인다.

**Step 2: Derive estimates of Standard Error of Final Estimates**

Step2에서는 MSF의 추정치가 0에 가까움을 확인하여 최종 추정치가 제대로 구해졌는지 확인하고 $Q$의 second order을 구함으로써 추정치의 standard error을 유도해본다.
$E\left\{\frac{\partial}{\partial\eta} log(\boldsymbol{y}, \boldsymbol{\delta}; \eta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\}$ 의 estimate을 구하기 위해서 mean score fuction부터 유도한다. 우선 complete log likelihood는 (3) 에서 유도하였으므로, 여기서 세 개의 모수인 $\pi_k, \mu_k, \sigma^2$ 에 대해서 미분을 수행하여 기댓값을 취한다. 여기서 $\hat{\eta}$는 Step1에서 데이터를 동일하게 생성하고 초깃값은 $\pi^{(0)} = (0.2, 0.2, 0.6)$, $\mu^{(0)} = (0, 3, -2)$, $\sigma^{2(0)} = 2$ 으로 설정한 하여 EM의 결과로 나온 모수들의 추정치이다.

- $\pi_k$ 에 대한 MSF

$$\frac{\partial}{\partial\pi_k} logL_c(\theta) = \sum_{i=1}^{n} z_{ik}/\pi_k - \lambda$$

$$\begin{aligned}\therefore E\left\{\frac{\partial}{\partial\pi_k} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\} &= \sum_{i=1}^{n} E\left\{z_{ik} \mid y_i, \hat{\eta}\right\}/\hat{\pi}_k - \lambda \\ &= \sum_{i=1}^{n} \frac{\hat{\pi}_k f_k(y_i; \hat{\eta})}{f(y_i; \hat{\eta})}/\hat{\pi}_k - \lambda\end{aligned}$$

  R로 확인해본 결과, $\pi_1, \pi_2, \pi_3$ 에 대한 MSF의 estimates 값은 각각 -0.0.206, -0.00042, 0.0045로 거의 0에 가까움을 확인하였다.

- $\mu_k$ 에 대한 MSF

$$\frac{\partial}{\partial\mu_k} logL_c(\theta) = \sum_{i=1}^{n} z_{ik}\frac{1}{\sigma^2}(y_i - \mu_k)$$

$$\therefore E\left\{\frac{\partial}{\partial \mu_k} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\} = \sum_{i=1}^{n} E\left\{z_{ik} \mid y_i, \hat{\eta}\right\}\left(y_i - \hat{\mu}_k\right)/\hat{\sigma}^2$$
$$= \sum_{i=1}^{n} \frac{\hat{\pi}_k f_k(y_i;\hat{\eta})}{f(y_i;\hat{\eta})}\left(y_i - \hat{\mu}_k\right)/\hat{\sigma}^2$$

R로 확인해본 결과, $\mu_1, \mu_2, \mu_3$에 대한 MSF의 estimates 값은 각각 0.0058, 0.0002, 0.0041로 거의 0에 가까움을 확인하였다.

- $\sigma^2$에 대한 MSF

$$\frac{\partial}{\partial \sigma^2} logL_c(\theta) = \sum\sum z_{ik}\left\{-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}\left(y_i - \mu_k\right)^2\right\}$$

$$E\left\{\frac{\partial}{\partial \sigma^2} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\} = \sum\sum E\left\{z_{ik} \mid y_i, \hat{\eta}\right\}\left\{-\frac{1}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\left(y_i - \hat{\mu}_k\right)^2\right\}$$
$$= \sum\sum \frac{\hat{\pi}_k f_k(y_i;\hat{\eta})}{f(y_i;\hat{\eta})}\left\{-\frac{1}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\left(y_i - \hat{\mu}_k\right)^2\right\}$$

R로 확인해본 결과, $\sigma^2$에 대한 MSF의 estimates 값은 -0.00237로 거의 0에 가까움을 확인하였다.

- $E\left\{-\frac{\partial^2}{\partial \pi_k^2} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\}$의 추정치

$$-\frac{\partial}{\partial \pi_k^2} logL_c(\theta) = \sum_{i=1}^{n} z_{ik}/\pi_k^2$$

$$E\left\{-\frac{\partial^2}{\partial \pi_k^2} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\} = \sum_{i=1}^{n} z_{ik}^{(t)}/\hat{\pi}_k^2$$

R을 통해, $k = 1, 2, 3$일 때, 각각 2269, 1215, 483로 확인하였다.

- $E\left\{-\frac{\partial^2}{\partial \pi_k \partial \mu_k} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial \pi_k \partial \mu_k} logL_c(\theta) = 0$$

따라서 추정치 또한 0이다.

- $E\left\{-\frac{\partial^2}{\partial \pi_k \partial \sigma^2} logL_c(\theta) \mid \boldsymbol{y}_{obs}, \delta, \hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial \pi_k \partial \sigma^2} logL_c(\theta) = 0$$

따라서 추정치 또한 0이다.

- $E\left\{-\frac{\partial^2}{\partial\mu_k^2}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\}$의 추정치

$$-\frac{\partial}{\partial\mu_k^2}logL_c(\theta) = \frac{1}{\sigma^2}\sum_{i=1}^{n} z_{ik}$$

$$E\left\{-\frac{\partial}{\partial\mu_k^2}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\} = \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n} z_{ik}^{(t)}$$

R을 통해, $k = 1, 2, 3$일 때, 각각 44, 83, 209로 확인하였다.

- $E\left\{-\frac{\partial^2}{\partial\mu_k\partial\pi_k}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial\mu_k\partial\pi_k}logL_c(\theta) = 0$$

따라서 추정치 또한 0이다.

- $E\left\{-\frac{\partial^2}{\partial\mu_k\partial\sigma^2}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial\mu_k\partial\sigma^2}logL_c(\theta) = \frac{1}{\hat{\sigma}^4}\sum_{i=1}^{n} z_{ik}^{(t)}(y_i - \hat{\mu}_k)$$

R을 통해, $k = 1, 2, 3$일 때, 각각 0.0065, 0.00032, 0.00469로 확인하였다.

- $E\left\{-\frac{\partial^2}{\partial(\sigma^2)^2}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial(\sigma^2)^2}logL_c(\theta) = -\sum\sum z_{ik}\left\{\frac{1}{2\sigma^4} - 2\frac{1}{2\sigma^6}\left(y_i - \mu_k\right)^2\right\}$$

R을 통해, 0.190으로 확인하였다.

- $E\left\{-\frac{\partial^2}{\partial(\sigma^2)\partial\mu_k}logL_c(\theta) \mid \boldsymbol{y}_{obs},\delta,\hat{\eta}\right\}$의 추정치

$$-\frac{\partial^2}{\partial(\sigma^2)\partial\mu_k}logL_c(\theta) = -\sum\sum z_{ik}\left\{-\frac{1}{\sigma^4}\left(y_i - \mu_k\right)\right\}$$

R을 통해, 0.0116으로 확인하였다.