$$"w N(\mu_1, \sigma_1^2) + (1-w) N(\mu_2, \sigma_2^2)"$$

$$(w, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$$

Ex 45 $X_1, \dots, X_n$ i.i.d $f(x; \underset{\sim}{\theta}) = w \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + (1-w) \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$

IDENTIFIABILITY: $\Theta = \{ w \in (0, \frac{1}{2}], \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 > 0, (\mu_1, \sigma_1^2) \neq (\mu_2, \sigma_2^2) \}$

OR $\Theta = \{ w \in [\frac{1}{2}, 1), \mu_1, \mu_2 \in \mathbb{R} \dots \dots \}$

STANDARD LOG-LIKELIHOOD:

$$l_n(\underset{\sim}{\theta}) = \sum_{i=1}^{n} \log \left(\frac{w}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(X_i-\mu_1)^2}{2\sigma_1^2}} + \frac{(1-w)}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(X_i-\mu_2)^2}{2\sigma_2^2}} \right)$$

$$\frac{\partial l_n(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} \overset{!}{=} 0$$

EM-algorithm approach:

- introduce $\underset{\sim}{Z_i} = \begin{pmatrix} Z_{i1} \\ Z_{i2} \end{pmatrix}$, $Z_{ij} = \begin{cases} 1, & X_i \text{ generated from } N(\mu_j, \sigma_j^2) \\ 0, & \text{otherwise} \end{cases}$

$$\underset{\sim}{Z_i} \sim \text{Mult}_2(1, (w, 1-w)) \qquad j \in \{0, 1\}$$

$$\mathcal{L}\left(X_i \mid \underset{\sim}{Z_i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \sim N(\mu_1, \sigma_1^2)$$

$$\mathcal{L}\left(X_i \mid \underset{\sim}{Z_i} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \sim N(\mu_2, \sigma_2^2)$$

→ COMPLETE LOG-LIKELIHOOD

$$l_n^c(\underset{\sim}{\theta}) = \sum_{i=1}^{n} \log f_{X, \underset{\sim}{Z}}(X_i, \underset{\sim}{Z_i}), \text{ where}$$

$$f_{X,\underset{\sim}{Z}}(x, \underset{\sim}{z}) = f_{X|\underset{\sim}{Z}}(x|\underset{\sim}{z}) f_{\underset{\sim}{Z}}(\underset{\sim}{z}) =$$

$$= \Big(z_1 \underbrace{f_1(x; \mu_1, \sigma_1^2)}_{\frac{1}{\sigma_1} \varphi\left(\frac{x-\mu_1}{\sigma_1}\right)} + z_2 \underbrace{f_2(x; \mu_2, \sigma_2^2)}_{\frac{1}{\sigma_2} \varphi\left(\frac{x-\mu_2}{\sigma_2}\right)} \Big) w^{z_1} (1-w)^{z_2}$$

$$\ell_n^c(\underset{\sim}{\theta}) = \sum_{i=1}^{n} \log\left(Z_{i1} f_1(X_i; \mu_1, \sigma_1^2) + Z_{i2} f_2(X_i; \mu_2, \sigma_2^2)\right) + \log\left(w^{Z_{i1}} (1-w)^{Z_{i2}}\right)$$

$$= \sum_{i=1}^{n} Z_{i1} \log f_1(X_i; \mu_1, \sigma_1^2) + Z_{i2} \log f_2(X_i; \mu_2, \sigma_2^2) + \sum_{i=1}^{n} Z_{i1} \log w + Z_{i2} \log(1-w)$$

$\boxed{\text{E - STEP}}$

$$Q(\underset{\sim}{\theta}, \underset{\sim}{\hat\theta}^{(\ell)}) = \mathbb{E}_{\hat\theta^{(\ell)}}\left[\ell_n^c(\underset{\sim}{\theta}) \mid \underset{\sim}{X}\right]$$

$$= \mathbb{E}_{\hat\theta^{(\ell)}}\left[Z_{i1} \log f(X_i; \mu_1, \sigma_1^2) + Z_{i2} \log f(X_i; \mu_2, \sigma_2^2) \mid \overset{X_i}{X_1, \dots, X_n}\right]$$

$$+ \mathbb{E}_{\hat\theta^{(\ell)}}\left[Z_{i1} \log w + Z_{i2} \log(1-w) \mid \underset{X_i}{X_1, \dots, X_n}\right]$$

WE NEED

$$\mathbb{E}_{\hat\theta^{(\ell)}}\left[Z_{i1} \mid X_i\right] = P_{\hat\theta^{(\ell)}}(Z_{i1} = 1 \mid X_i) = f_{Z_1 \mid X}(1 \mid X_i; \underset{\sim}{\hat\theta}^{(\ell)})$$

$$= \frac{f_{X \mid Z_1}(X_i \mid 1; \hat\theta^{(\ell)})\, f_{Z_1}(1; \underset{\sim}{\hat\theta}^{(\ell)})}{f_X(X_i; \underset{\sim}{\hat\theta}^{(\ell)})}$$

$\mathcal{L}(X \mid Z_1 = 1) \sim f_1(X_i; \mu_1, \sigma_1^2)$

$Z_1 \sim Be(w)$

$$= \frac{f_1(X_i; \hat\mu_1^{(\ell)}, \hat\sigma_1^{2(\ell)}) \cdot \hat w^{(\ell)}}{f_1(X_i; \hat\mu_1^{(\ell)}, \hat\sigma_1^{2(\ell)})\, \hat w^{(\ell)} + f_2(X_i; \hat\mu_2^{(\ell)}, \hat\sigma_2^{2(\ell)})(1 - \hat w^{(\ell)})}$$

DENOM = $f_1(X_i; \hat\mu_1^{(\ell)}, \hat\sigma_1^{2(\ell)})\, \hat w^{(\ell)} + f_2(X_i; \hat\mu_2^{(\ell)}, \hat\sigma_2^{2(\ell)})(1 - \hat w^{(\ell)})$

$$=: r_{i1}^{(\ell)}$$

ANALOGOUSLY : $\mathbb{E}_{\hat\theta^{(\ell)}}\left[Z_{i2} \mid X_i\right] = \dfrac{f_2(X_i; \hat\mu_2^{(\ell)}, \hat\sigma_2^{2(\ell)})(1 - \hat w^{(\ell)})}{\text{DENOM}}$

$$=: r_{i2}^{(\ell)}$$

NOTE THAT $r_{i2}^{(\ell)} = 1 - r_{i1}^{(\ell)}$

$$\boxed{\text{M-STEP}}$$

$$
\begin{aligned}
Q(\underset{\sim}{\theta}\,|\,\hat{\underset{\sim}{\theta}}^{(k)}) = & \sum_{i=1}^{n} \mathbb{E}_{\hat{\underset{\sim}{\theta}}^{(k)}}[Z_{i1}|X_i] \log f_1(X_{i1};\mu_1,\sigma_1^2) \\
& + \mathbb{E}_{\hat{\underset{\sim}{\theta}}^{(k)}}[Z_{i2}|X_i] \log f_2(X_{i2};\mu_2,\sigma_2^2) \\
& + \sum_{i=1}^{n} \mathbb{E}_{\hat{\underset{\sim}{\theta}}^{(k)}}[Z_{i1}|X_i] \log w + \mathbb{E}_{\hat{\underset{\sim}{\theta}}^{(k)}}[Z_{i2}|X_i] \log(1-w)
\end{aligned}
$$

$$
\begin{aligned}
= & \overbrace{\sum_{i=1}^{n} r_{i1}^{(k)} \log f_1(X_i;\mu_1,\sigma_1^2)}^{\text{"}\mu_1,\sigma_1^2\text{"}} + \sum_{i=1}^{n} r_{i2}^{(k)} \log f_2(X_i;\mu_2,\sigma_2^2) \\
& + \underbrace{\left(\sum_{i=1}^{n} r_{i1}^{(k)} \log w + (1-r_{i1}^{(k)}) \log(1-w)\right)}_{\text{"}w\text{"}}
\end{aligned}
$$

$$\hat{\underset{\sim}{\theta}}^{(k+1)} = \underset{\underset{\sim}{\theta}\in\Theta}{\operatorname{argmax}}\ Q(\underset{\sim}{\theta}\,|\,\hat{\underset{\sim}{\theta}}^{(k)})$$

$$
\begin{aligned}
\rightsquigarrow \hat{w}^{(k+1)} &= \underset{w}{\operatorname{argmax}} \sum_{i=1}^{n} r_{i1}^{(k)} \log w + (1-r_{i1}^{(k)}) \log(1-w) \\
&= \frac{1}{n} \sum_{i=1}^{n} r_{i1}^{(k)}
\end{aligned}
$$

$$\rightarrow (\hat{\mu}_1^{(k+1)}, \hat{\sigma}_1^{2(k+1)}) = \underset{\mu_1,\sigma_1^2}{\operatorname{argmax}} \underbrace{\sum_{i=1}^{n} r_{i1}^{(k)} \log\left(\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{1}{2\sigma_1^2}(X_i-\mu_1)^2\right\}\right)}_{\ell_1(\mu_1,\sigma_1^2)}$$

$$\ell_1(\mu_1,\sigma_1^2) = \sum_{i=1}^{n} \frac{-r_{i1}^{(k)}}{2} \log \sigma_1^2 - \frac{r_{i1}^{(k)}}{\sigma_1^2}(X_i-\mu_1)^2$$

$$\frac{\partial \ell_1(\mu_1,\sigma_1^2)}{\partial \mu_1} = -\sum_{i=1}^{n} \frac{r_{i1}^{(k)}}{\sigma_1^2}(X_i-\mu_1)(-2) \overset{!}{=} 0$$

$$\rightsquigarrow \hat{\mu}_1^{(k+1)} = \frac{\sum_{i=1}^{n} r_{i1}^{(k)} X_i}{\sum_{i=1}^{n} r_{i1}^{(k)}}$$

$$\frac{\partial \ell_1(\hat{\mu}_1^{(k+1)},\sigma_1^2)}{\partial \sigma_1^2} = -\sum_{i=1}^{n} \frac{r_{i1}^{(k)}}{2\sigma_1^2} + \frac{r_{i1}^{(k)}}{\sigma_1^4}(X_i-\hat{\mu}_1^{(k+1)})^2 = 0$$

$$\rightsquigarrow \hat{\sigma}_1^{2(k+1)} = \frac{\sum_{i=1}^{n} r_{i1}^{(k)} (X_i-\hat{\mu}_1^{(k+1)})^2}{\sum_{i=1}^{n} r_{i1}^{(k)}}$$

ANALOGOUSLY:

$$\hat{\mu}_2^{(l+1)} = \frac{\sum_{i=1}^{n} r_{i2}^{(l)} X_i}{\sum_{i=1}^{n} r_{i2}^{(l)}}$$

$$\hat{\sigma}_2^{(l+1)} = \frac{\sum_{i=1}^{n} r_{i2}^{(l)} \left(X_i - \hat{\mu}_2^{(l+1)}\right)^2}{\sum_{i=1}^{n} r_{i2}^{(l)}}$$

Ex. 44 $Y_1, \ldots, Y_n \sim$ i.i.d. $\mathrm{EXP}(\lambda)$, $f(y;\lambda) = \lambda e^{-\lambda y} \mathbb{1}\{y > 0\}$

$0 = d_0 < d_1 < \ldots < d_M = \infty$

WE ONLY KNOW THAT $Y_i \in (d_{q_i - 1}, d_{q_i}]$,

FOR SOME $q_i \in \{1, \ldots, M\}$