Mixture Models

mixture distn: $[\underline{\theta}] = \sum_{j=1}^{J} p_j [\underline{\theta}]_j$

for $\sum_j p_j = 1$, $\forall p_j > 0$

and valid $[\underline{\theta}]_j$ s.t. $\int [\underline{\theta}]_j \, d\underline{\theta} = 1$

note! This concept is useful for data, process, and parameters.

Example: 2-component mixture model

integrated likelihood specification

1.) $y_i \sim p[y_i | \underline{\theta}_1]_1 + (1-p)[y_i | \underline{\theta}_2]_2$

2.) $y_i \sim \begin{cases} [y_i | \underline{\theta}_1]_1 & , \text{ w.p. } p \\ [y_i | \underline{\theta}_2]_2 & , \text{ w.p. } 1-p \end{cases}$

latent variable specification

3.) $y_i \sim \begin{cases} [y_i | \underline{\theta}_1]_1 & , \ z_i = 1 \\ [y_i | \underline{\theta}_2]_2 & , \ z_i = 0 \end{cases}$

$z_i \sim \text{Bern}(p)$

Likelihood options w/ latent variables:

1.) $\prod_{i=1}^{n} z_i [y_i | \underline{\theta}_1]_1 + (1-z_i)[y_i | \underline{\theta}_2]_2$

2.) $\prod_{i=1}^{n} [y_i | \underline{\theta}_1]_1^{z_i} [y_i | \underline{\theta}_2]_2^{1-z_i}$

Priors: $p \sim \text{Beta}(\alpha, \beta)$, $\underline{\theta}_1 \sim [\underline{\theta}_1]$

$\underline{\theta}_2 \sim [\underline{\theta}_2]$

Generic Posterior:

$$[\underline{\theta}_1, \underline{\theta}_2, \underline{z}, p \mid \underline{y}] \propto$$

$$\propto \left(\prod_{i=1}^{n} [y_i|\underline{\theta}_1]_1^{z_i} [y_i|\underline{\theta}_2]_2^{1-z_i} [z_i|p]\right)[\underline{\theta}_1][\underline{\theta}_2][p]$$

Full-Conditionals:

$$[z_i|\cdot] \propto [y_i|\underline{\theta}_1]_1^{z_i} [y_i|\underline{\theta}_2]_2^{1-z_i} \underbrace{[z_i|p]}_{p^{z_i}(1-p)^{1-z_i}}$$

$$\propto \left(p[y_i|\underline{\theta}_1]_1\right)^{z_i} \left((1-p)[y_i|\underline{\theta}_2]_2\right)^{1-z_i}$$

$$= \text{Bern}(\tilde{p}_i), \quad \tilde{p}_i = \frac{p[y_i|\underline{\theta}_1]_1}{p[y_i|\underline{\theta}_1]_1 + (1-p)[y_i|\underline{\theta}_2]_2}$$

$$[p|\cdot] \propto \prod_{i=1}^{n} [z_i|p][p]$$

(from earlier models)

$$= \text{Beta}\left(\sum_{i=1}^{n} z_i + \alpha, \sum_{i=1}^{n}(1-z_i) + \beta\right)$$

$$[\underline{\theta}_1|\cdot] \propto \prod_{i=1}^{n} [y_i|\underline{\theta}_1]_1^{z_i} [\underline{\theta}_1]$$

$$\propto \prod_{\forall z_i = 1} [y_i|\underline{\theta}_2]_2^{z_i} [\underline{\theta}_1]$$

and similar for $\underline{\theta}_2$.

Gaussian mixture model:

$$y_i \sim \begin{cases} N(\mu_1, \sigma_1^2), & z_i = 1 \\ N(\mu_2, \sigma_2^2), & z_i = 0 \end{cases}, \quad i = 1, \dots, n$$

$$z_i \sim \text{Bern}(p)$$

$$\mu_1 \sim N(\mu_{10}, \sigma_{10}^2)$$

$$\mu_2 \sim N(\mu_{20}, \sigma_{20}^2)$$

$$\sigma_1^2 \sim IG(q_1, r_1)$$

$$\sigma_2^2 \sim IG(q_2, r_2)$$

Posterior:

$$[\underline{z}, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid y] \propto \left(\prod_{i=1}^{n} [y_i \mid \mu_1, \sigma_1^2]^{z_i} [y_i \mid \mu_2, \sigma_2^2]^{1-z_i} [z_i \mid p] \right) [\mu_1][\mu_2][\sigma_1^2][\sigma_2^2][p]$$

Full-conditional distns:

$$[z_i \mid \cdot] = \text{Bern}(\hat{p}_i), \quad \hat{p}_i = \frac{p[y_i \mid \mu_1, \sigma_1^2]}{p[y_i \mid \mu_1, \sigma_1^2] + (1-p)[y_i \mid \mu_2, \sigma_2^2]}$$

$$[p \mid \cdot] = \text{Beta}\left(\sum_{i=1}^{n} z_i + \alpha, \sum_{i=1}^{n} (1 - z_i) + \beta\right)$$

$$[\mu_1 \mid \cdot] \propto \prod_{i=1}^{n} [y_i \mid \mu_1, \sigma_1^2]^{z_i} [\mu_1]$$

$$\propto \exp\left\{-\frac{1}{2} \frac{\sum_{z_i=1} (y_i - \mu_1)^2}{\sigma_1^2}\right\} \exp\left\{-\frac{1}{2} \frac{(\mu_1 - \mu_{10})^2}{\sigma_{10}^2}\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(-2 \underbrace{\left(\frac{\sum_{z_i=1} y_i}{\sigma_1^2} + \frac{\mu_{10}}{\sigma_{10}^2}\right)}_{b} \mu_1 + \mu_1^2 \underbrace{\left(\frac{\sum_{i=1}^{n} z_i}{\sigma_1^2} + \frac{1}{\sigma_{10}^2}\right)}_{a}\right)\right\}$$

$$= N(a^{-1} b, a^{-1})$$

full-conditionals cont.

$$[\mu_2 \mid \cdot] = N(a^{-1}b, a^{-1})$$

$$a = \frac{\sum_{i=1}^{n} 1-z_i}{\sigma_2^2} + \frac{1}{\sigma_{12}^2}, \quad b = \frac{\sum_{\forall z_i = 0} y_i}{\sigma_2^2} + \frac{\mu_{20}}{\sigma_{20}^2}$$

$$[\sigma_1^2 \mid \cdot] \propto \prod_{i=1}^{n} [y_i \mid \mu_1, \sigma_1^2]^{z_i} [\sigma_1^2]$$

$$\propto (\sigma^2)^{-\frac{\sum_{i=1}^{n} z_i}{2}} \exp\left\{-\frac{1}{\sigma_1^2} \frac{\sum_{\forall z_i = 1}(y_i - \mu_1)^2}{2}\right\} (\sigma^2)^{-(q_1+1)} \exp\left\{-\frac{1}{\sigma^2 r_1}\right\}$$

$$\propto (\sigma^2)^{-\underbrace{\left(\frac{\sum_{i=1}^{n} z_i}{2} + q_1 + 1\right)}_{\tilde{q}}} \exp\left\{-\frac{1}{\sigma^2}\underbrace{\left(\frac{\sum_{\forall z_i = 1}(y_i - \mu_1)^2}{2} + \frac{1}{r_1}\right)}_{\tilde{r}^{-1}}\right\}$$

$$= IG(\tilde{q}, \tilde{r})$$

$$[\sigma_2^2 \mid \cdot] = IG(\tilde{q}, \tilde{r})$$

$$\tilde{q} = \frac{\sum_{i=1}^{n} 1 - z_i}{2} + q_2, \quad \tilde{r} = \left(\frac{\sum_{\forall z_i = 0}(y_i - \mu_2)^2}{2} + \frac{1}{r_2}\right)^{-1}$$

Note: all Gibbs updates!