Ques-1 $f_i(x) = \frac{1}{\pi b} \cdot \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2}$; $i=1,2$

equal priors. let $(a_1 < a_2)$

- $p_1(x) = p_2(x) = 1/2$

(a) for class conditional densities to be PDF

$$\int_{-\infty}^{\infty} f_x(x) \cdot dx = 1$$

let $\frac{x-a_i}{b} = t$

$$\Rightarrow \int_{-\infty}^{\infty} \frac{1}{\pi b} \cdot \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2} \cdot dx$$

$\Rightarrow \frac{dx}{b} = dt$

$$\Rightarrow \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{1+t^2} dt = \frac{1}{\pi}\left[\tan^{-1} t\right]_{-\infty}^{\infty}$$

$$\Rightarrow \frac{1}{\pi}\left[\frac{\pi}{2} - \left(\frac{-\pi}{2}\right)\right] = \frac{1}{\cancel{\pi}} \cdot \cancel{\pi} = 1$$

$\therefore$ probability density function.

(b) Baye's Boundary : $q_1(x) - q_2(x) = 0$

$$\Rightarrow \frac{p_1 f_1(x)}{\sum p_i f_i(x)} - \frac{p_2 f_2(x)}{\sum p_i f_i(x)} = 0$$

$$\Rightarrow \frac{1}{2} \frac{1}{\pi b}\left(\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} - \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}\right) = 0$$

$$\Rightarrow \left(\frac{x-a_1}{b}\right)^2 - \left(\frac{x-a_2}{b}\right)^2 = 0 \Rightarrow (x-a_1)^2 - (x-a_2)^2 = 0$$

$$\Rightarrow \boxed{x = \frac{a_1+a_2}{2}}$$

(c) $P_{err} = P(h(x) \neq Y)$

$= P(h(x)=1, Y=2) + P(h(x)=2, Y=1)$

$= P(h(x)=1 \mid Y=2) \cdot P(Y=2) + P(h(x)=2 \mid Y=1) \cdot p(Y=1)$

$= p \cdot \int_{-\infty}^{\frac{a_1+a_2}{2}} f_2(x) dx + p \cdot \int^{\infty} f_1(x) dx$

let $\frac{x-a_2}{b} = t_2$

$$= p.\int_{-\infty}^{\frac{a_1+a_2}{2}} f_2(x)\,dx + p.\int_{\frac{a_1+a_2}{2}}^{\infty} f_1(x)\,dx$$

let $\frac{x-a_2}{b} = t_2$

$\frac{dx}{b} = dt_2$

$$= \frac{1}{2}\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1}{\pi b}\cdot\frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}dx + \frac{1}{2}\int_{\frac{a_1+a_2}{2}}^{\infty}\frac{1}{\pi b}\cdot\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}dx$$

let $\frac{x-a_1}{b} = t_1$

$\frac{dx}{b} = dt_1$

$$= \frac{1}{2\pi}\int_{-\infty}^{\frac{a_1-a_2}{2b}} \frac{1}{1+t_2^2}dt_2 + \int_{\frac{a_2-a_1}{2b}}^{\infty}\frac{1}{1+t_1^2}dt_1$$

$$= \frac{1}{2\pi}\left\{\left[\tan^{-1}t_2\right]_{-\infty}^{\frac{a_2-a_1}{2b}} + \left[\tan^{-1}t_1\right]_{\frac{a_2-a_1}{2}}^{\infty}\right\}$$

$$= \frac{1}{2\pi}\left[\pi - 2\tan^{-1}\left(\frac{a_2-a_1}{2b}\right)\right]$$

Ques-2: $h(x) = P(h(x)=i\,|\,x)$ a-classes $\{1,2,3,\dots,a\}$

priors: $p_i = P(Y=i)$

$f_i(x) = P(x|y=i)$

Posteriors: $q_i(x) = P(Y=i|x) = \frac{p_i f_i}{\sum_i p_i f_i}$

For 0-1 loss: $L(h(x), c_i) = \begin{cases} 0 & ; h(x)=c_i \\ 1 & ; h(x)\neq c_i \end{cases}$

Risk: $R(h) = E[L(h(x),Y)]$

$R(h(x)|x) = E[L(h(x),Y)\,|\,x]$

$$= \sum_{i=1}^{a} q_i(x).\,L(h(x), c_i)$$

$$= \underset{h(x)=c_0}{0.q_0(x)} + \sum_{h(x)\neq c_i} 1.\,q_i(x)$$

$$= \sum_{h(x)\neq c_i} q_i(x)$$

we know

$\sum q_i(x) = 1$

$$= \sum_{h(x) \neq C_i} q_i(x)$$

we know $\sum q_i(x) = 1$

$$= 1 - \underset{h(x)=C_0}{q_0(x)}$$

(b) To improve the classifier, we can associate different loss values for different classes rather than just 0-1 loss.

Ques-3:
$$L(h(x)=i, Y=j) = \begin{cases} 0 & ; \; i=j \\ \lambda_r & ; \; i=k+1 \quad \text{(rejection)} \\ \lambda_m & ; \; \text{otherwise} \quad \text{(misclassification)} \end{cases}$$

$$R[h(x)=i|x] = \begin{cases} \lambda_r \\ E[L(h(x), C_i|X] \end{cases}$$

$$= \begin{cases} \lambda_r \\ \sum q_j(x).\lambda m \end{cases} = \begin{cases} \lambda r \\ \lambda m(1-q_i(x)) \end{cases}$$

To minimize the risk:

$$R[h(x)=i|x] \leq R[h(x)=j|x] \quad \forall j$$

$$\Rightarrow \lambda_m(1-q_i(x)) \leq \begin{cases} \lambda r \\ \lambda_m(1-q_j(x)) \end{cases}$$

$$\Rightarrow \quad q_i(x) \geq \begin{cases} 1 - \dfrac{\lambda r}{\lambda m} \\ q_j(x) \end{cases}$$

- if $\lambda_r = 0$; $\quad q_i(x) \geq 1 - \dfrac{0}{\lambda_m} \Rightarrow q_i(x) \geq 1$

  $\therefore$ always reject //

- if $\lambda_r > \lambda_m$; $\quad q_i(x) \geq 1 - \dfrac{\lambda r}{\lambda m} \Rightarrow q_i(x) > 0$

  $\therefore$ never reject //

$\therefore$ never reject //

Ques-4: Derive MLE for:

(a) Exponential function

$P(x,\lambda) = \lambda e^{-\lambda x}$ $\qquad\qquad \mathcal{D} = \{x_i\}_{i=1}^{n}$

likelihood:

$$L(\lambda,\mathcal{D}) = P(\mathcal{D}|\lambda) = \prod_{i=1}^{n} P(x_i,\lambda)$$

$$= \lambda^n e^{-\lambda\sum_i^n x_i}$$

log-likelihood:

$$\ell(\lambda) = n\log\lambda - \lambda\sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum x_i = 0 \quad \Rightarrow \quad \boxed{\lambda_{MLE} = \frac{n}{\sum_{i=1}^{n} x_i}}$$

(b) Multivariate Gaussian distribution:

$$P(x,\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^d|\Sigma|}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

likelihood: $L = \prod_{i=1}^{n} P(x_i,\mu,\Sigma)$

log-likelihood: $\ell = \sum_{i=1}^{n} \log P(x_i,\mu,\Sigma)$

$$= \sum_{i=1}^{n}\left[\log\left(\frac{1}{\sqrt{(2\pi)^d|\Sigma|}}\exp\left\{-\frac{1}{2}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu)\right\}\right)\right]$$

$$\ell = \sum_{i=1}^{n}\log\frac{1}{\sqrt{(2\pi)^d|\Sigma|}} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu)$$

- $$\frac{\partial \ell}{\partial \mu} = 0 - \frac{1}{2}\sum_{i=1}^{n}\Sigma^{-1}(x_i-\mu) = 0$$

$$\Rightarrow \quad \mu_{MLE} = \frac{\sum_{i=1}^{n} x_i}{\quad}$$

$$\Rightarrow \boxed{\mu_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

- $$\frac{\partial l}{\partial \Sigma} = \sum_{i=1}^{n}\left(-\frac{1}{2}\cdot\frac{1}{\Sigma} + \frac{1}{2}(x_i-\mu)^T(x_i-\mu)\frac{1}{\Sigma^2}\right) = 0$$

$$\Rightarrow \boxed{\Sigma_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^T(x_i-\mu)}$$

Ques 5. $f_x(x) = \mathcal{N}(\mu, \sigma^2)$

$p(\mu) = \mathcal{N}(\mu_0, \sigma_0^2)$

- $$f(D|\mu) = \prod_{i=1}^{n} \mathcal{N}(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2}\cdot\frac{(x_i-\mu)^2}{\sigma^2}\right\}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

- $f(\mu) = \mathcal{N}(\mu_0, \sigma_0^2)$

$$\Rightarrow f(\mu) = \frac{1}{\sigma_0\sqrt{2\pi}} \cdot \exp\left\{\frac{-1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

using baye's rule:

$$f(\mu|D) \propto f(\mu) \cdot f(D|\mu)$$

$$= \underbrace{k}_{\text{some constant}} \exp\left\{\frac{-1}{2\sigma_0^2}(\mu-\mu_0)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

log-likelihood: $$l(\mu|D) = \log k - \frac{1}{2\sigma_0^2}(\mu-\mu_0)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial l}{\partial \mu} = 0 - \frac{1}{\sigma_0^2}(\mu-\mu_0) + \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\Rightarrow \frac{n\mu}{\sigma^2} + \frac{\mu}{\sigma_0^2} = \frac{\mu_0}{\sigma_0^2} + \frac{\sum x_i}{\sigma^2}$$

$$\Rightarrow \frac{n\mu}{\sigma^2} + \frac{\mu}{\sigma_0^2} = \frac{\mu_0}{\sigma_0^2} + \frac{\Sigma x_i}{\sigma^2}$$

$$\Rightarrow \boxed{\mu = \frac{\sigma_0^2 \sum_{i=1}^{n} x_i + \sigma^2 \mu_0}{\sigma_0^2 n + \sigma^2}}$$

- since mode occurs at mean in $N(\mu, \sigma^2)$, both will have same value of μ.

Ques-6:

(a) Mixture of gaussians

$$f(x_i | \mu_j, \sigma_j^2) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{(x_i - \mu_j)^2}{\sigma_j^2}\right\}$$

log-likelihood:

$$\begin{aligned} \ell(\theta) &= \sum_{i=1}^{n} \log(f(x_i | \theta)) \\ &= \sum_{i=1}^{n} \sum_{j} \log(f(x_i, z_i = j | \theta)) \\ &= \sum_{i=1}^{n} \log \sum_{j} \pi_j f(x_i | \mu_j, \sigma_j^2) \end{aligned}$$

- $$\begin{aligned} \frac{\partial f(x_i | \mu_j, \sigma_j^2)}{\partial \mu_j} &= \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left\{\frac{-1}{2} \frac{(x_i - \mu_j)^2}{\sigma_j^2}\right\} \cdot \frac{-1}{2\sigma_j^2} \cdot (-2)(x_i - \mu_j) \\ &= f(x_i | \mu_j, \sigma_j^2) \cdot \frac{(x_i - \mu_j)}{\sigma_j^2} \quad \text{— (i)} \end{aligned}$$

- $$\begin{aligned} \frac{\partial f(x_i | \mu_j, \sigma_j^2)}{\partial \sigma_j} &= \frac{-1}{\sigma_j^2 \sqrt{2\pi}} \exp\left\{\frac{-1}{2} \frac{(x_i - \mu_j)^2}{\sigma_j^2}\right\} \\ &\quad + \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left\{\frac{-1}{2} \frac{(x_i - \mu_j)^2}{\sigma_j^2}\right\} \cdot \frac{-1}{2} (x_i - \mu_j)^2 \cdot \frac{(-2)}{\sigma_j^3} \\ &= f(x_i | \mu_j, \sigma_j^2) \cdot \left[\frac{(x_i - \mu_j)^2}{\sigma_j^3} - \frac{1}{\sigma_j}\right] \quad \text{— (ii)} \end{aligned}$$

Now,

$$\theta_{MLE} = \underset{\tilde{\theta}}{\operatorname{argmax}} \; \ell(\tilde{\theta}) \;\Big|\; \sum_j \pi_j = 1$$

Now,

$$\theta_{MLE} = \underset{\tilde{\theta}}{\operatorname{argmax}} \ \ell(\tilde{\theta}) \ \Big| \ \sum_j \pi_j = 1$$

Lagrangian: $\mathcal{L}(\theta,\lambda) = \underset{\theta}{\operatorname{argmax}} \ \ell(\theta) + \lambda\left(1 - \sum_j \pi_j\right)$

- $$\frac{\partial \mathcal{L}}{\partial \mu_j} = \frac{\partial}{\partial \mu_j}\left[\sum_{i=1}^{n} \log \sum_j \pi_j f(x_i|\mu_j,\sigma_j^2) + \lambda\left(1 - \sum_j \pi_j\right)\right]$$

$$= \sum_{i=1}^{n}\left(\frac{1}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)} \cdot \frac{\partial}{\partial \mu_j} \pi_j f(x_i|\mu_j,\sigma_j^2)\right)$$

$$= \sum_{i=1}^{n}\left(\frac{\pi_j \cdot f(x_i|\mu_j,\sigma_j^2)}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)} \cdot \frac{(x_i-\mu_j)}{\sigma_j^2}\right) \quad \text{(using eq (i))}$$

- $$\frac{\partial \mathcal{L}}{\partial \sigma_j} = \frac{\partial}{\partial \sigma_j}\left[\sum_{i=1}^{n} \log \sum_j \pi_j f(x_i|\mu_j,\sigma_j^2) + \lambda\left(1 - \sum_j \pi_j\right)\right]$$

$$= \sum_{i=1}^{n}\left(\frac{1}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)} \cdot \frac{\partial}{\partial \sigma_j} \pi_j f(x_i|\mu_j,\sigma_j^2)\right)$$

$$= \sum_{i=1}^{n}\left[\frac{\pi_j \cdot f(x_i|\mu_j,\sigma_j^2)}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)}\left\{\frac{(x_i-\mu_j)^2}{\sigma_j^3} - \frac{1}{\sigma_j}\right\}\right] \quad \text{(using eq (ii))}$$

- $$\frac{\partial \mathcal{L}}{\partial \pi_j} = \frac{\partial}{\partial \pi_j}\left[\sum_{i=1}^{n} \log \sum_j \pi_j f(x_i|\mu_j,\sigma_j^2) + \lambda\left(1 - \sum_j \pi_j\right)\right]$$

$$= \sum_{i=1}^{n}\left[\frac{1}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)} \frac{\partial}{\partial \pi_j} \pi_j f(x_i|\mu_j,\sigma_j^2)\right] - \lambda = 0$$

$$\sum_j\left(\sum_{i=1}^{n}\left[\frac{\pi_j f(x_i|\mu_j,\sigma_j^2)}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)}\right] - \lambda \pi_j\right) \quad \text{(multiplying both sides by } \textstyle\sum_j \pi_j\text{)}$$

put $$\frac{\pi_j f(x_i|\mu_j,\sigma_j^2)}{\sum_j \pi_j f(x_i|\mu_j,\sigma_j^2)} = \Gamma_{ij}$$

- $$\frac{\partial \mathcal{L}}{\partial \mu_j} = 0 \Rightarrow \sum_{i=1}^{n} \Gamma_{ij} \frac{(x_i-\mu_j)}{\sigma_j^2} = 0$$

- $\frac{\partial \ell}{\partial \mu_j} = 0 \Rightarrow \sum_{i=1}^{n} \Gamma_{ij} \frac{(x_i - \mu_j)}{\sigma_j^2} = 0$

$$\boxed{\mu_j = \sum_{i=1}^{n} \Gamma_{ij} x_i \Big/ \sum_{i=1}^{n} \Gamma_{ij}}$$

- $\frac{\partial \ell}{\partial \sigma_j} = 0 \Rightarrow \sum_{i=1}^{n} \Gamma_{ij} (x_i - \mu_j)^2 \left[\frac{1}{\sigma_j^3} - \frac{1}{\sigma_j}\right] = 0$

$$\Rightarrow \boxed{\sigma_j^2 = \sum_{i=1}^{n} \Gamma_{ij} (x_i - \mu_j)^2 \Big/ \sum_{i=1}^{n} \Gamma_{ij}}$$

- $\frac{\partial \ell}{\partial \pi_j} = 0 \Rightarrow \sum_j \left(\sum_{i=1}^{n} \Gamma_{ij} - \lambda \pi_j \right) = 0$

$$\Rightarrow \sum_j \sum_{i=1}^{n} \Gamma_{ij} - \lambda \cancel{\sum_j \pi_j}^{\,1} = 0$$

$$\Rightarrow \lambda = \sum_j \sum_i \Gamma_{ij} \Rightarrow \lambda = \sum_j 1 \Rightarrow \lambda = n$$

$$\Rightarrow \boxed{\pi_j = \frac{1}{n} \sum_{i=1}^{n} \Gamma_{ij}}$$

EM algorithm:

1). Choose $\theta^{(0)}$ at random.
2). Loop $t = 1, 2, \ldots$ till convergence

E-step: $$\Gamma_{ij}^{(t+1)} = \frac{\pi_j^{(t)} f(x_i \mid \mu_j^{(t)}, \sigma_j^{2(t)})}{\sum_j \pi_j^{(t)} f(x_i \mid \mu_j^{(t)}, \sigma_j^{2(t)})}$$

M-step: $$\mu_j^{(t+1)} = \frac{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)} \cdot x_i}{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}}$$

$$\sigma_j^{2(t+1)} = \frac{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)} (x_i - \mu_j^{(t+1)})^2}{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}}$$

$$\pi_j^{(t+1)} = \frac{1}{n} \sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}$$

$$\pi_j^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}$$

(b) Mixture of Bernoulli

$$f(x|p) = p^x (1-p)^{(1-x)}$$

- $$\frac{\partial f(x_i|p_j)}{\partial p_j} = x_i p_j^{x_i - 1}(1-p_j)^{(1-x_i)} - p_j^{x_i}(1-x_i)(1-p_j)^{-x_i}$$

$$= x_i p_j f(x_i|p_j) - (1-x_i)(1-p_j) f(x_i|p_j)$$

$$= f(x_i|p_j)\left[\cancel{x_i p_j} - 1 + p_j + x_i - \cancel{x_i p_j}\right]$$

$$= f(x_i|p_j)[p_j + x_i - 1]$$

log-likelihood:

$$\ell(\theta) = \sum_{i=1}^{n} \log f(x_i|\theta)$$

$$\ell(\theta,\pi) = \sum_{i=1}^{n} \log \sum_j f(x_i, z_i = j|\theta)$$

$$= \sum_{i=1}^{n} \log \sum_j \pi_j f(x_i|\theta_j)$$

$$\theta_{MLE} = \underset{\tilde{\theta}}{\operatorname{argmax}}\ \ell(\theta,\pi) \ \Big|\ \sum_j \pi_j = 1$$

Lagrangian: $\mathcal{L}(\theta,\pi,\lambda) = \underset{\tilde{\theta}}{\operatorname{argmax}}\ \ell(\tilde{\theta},\pi) + \lambda(1 - \sum_j \pi_j)$

$$\mathcal{L} = \sum_{i=1}^{n} \log \sum_j \pi_j f(x_i|p_j) + \lambda(1 - \sum_j \pi_j)$$

- $$\frac{\partial \mathcal{L}}{\partial p_j} = \sum_{i=1}^{n} \frac{\pi_j f(x_i|p_j)}{\sum_j \pi_j f(x_i|p_j)}[p_j + x_i - 1]$$

- $$\frac{\partial \mathcal{L}}{\partial \pi} = \sum_{i=1}^{n} \frac{f(x_i|p_j)}{\sum_j \pi_j f(x_i|p_j)} - \lambda = 0$$

(multiplying both sides by $\sum_j \pi_j = 1$)

$$= \sum_j \left(\sum_{i=1}^{n} \frac{\pi_j f(x_i|p_j)}{\sum_j \pi_j f(x_i|p_j)} - \lambda \pi_j\right)$$

$$= \sum_j \left[\sum_{i=1} \frac{\pi_j f(x_i|p_j)}{\sum_j \pi_j f(x_i|p_j)} - \lambda \pi_j \right]$$

put $\quad \Gamma_{ij} = \dfrac{\pi_j f(x_i|p_j)}{\sum_j \pi_j f(x_i|p_j)}$

- $\dfrac{\partial \mathcal{L}}{\partial p_j} = 0 \Rightarrow \sum_{i=1}^{n} \Gamma_{ij} [p_j + x_i - 1] = 0$

$$\Rightarrow \boxed{p_j = \sum_{i=1}^{n} \Gamma_{ij}[x_i - 1] \Big/ \sum_{i=1}^{n} \Gamma_{ij}}$$

- $\dfrac{\partial \mathcal{L}}{\partial \pi_j} = 0 \Rightarrow \sum_j \left(\sum_{i=1}^{n} \Gamma_{ij} - \lambda \pi_j \right) = 0$

$$\Rightarrow \sum_j \sum_{i=1}^{n} \Gamma_{ij} - \lambda \cancelto{1}{\sum_j \pi_j} = 0$$

$$\Rightarrow \lambda = \sum_j \sum_i \Gamma_{ij} \Rightarrow \lambda = \sum_j 1 \Rightarrow \lambda = n$$

$$\Rightarrow \boxed{\pi_j = \frac{1}{n} \sum_{i=1}^{n} \Gamma_{ij}}$$

EM algorithm:

1). Choose $\theta^{(0)}$ at random.

2). Loop $t = 1, 2, \dots$ till convergence

E-step: $\quad \Gamma_{ij}^{(t+1)} = \dfrac{\pi_j^{(t)} f(x_i|p_j)}{\sum_j \pi_j^{(t)} f(x_i|p_j)}$

M-step: $\quad p_j^{(t+1)} = \dfrac{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)} [x_i - 1]}{\sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}}$

$$\pi_j^{(t+1)} = \frac{1}{n} \sum_{i=1}^{n} \Gamma_{ij}^{(t+1)}$$

Ques-7: $\theta \sim$ Random variable
$f(\theta) \sim$ prior
$f(\mathcal{D}|\theta) \sim$ likelihood

$$f(\theta|\mathcal{D}) = \frac{f(\theta).f(\mathcal{D}|\theta)}{\int f(\theta).f(\mathcal{D}|\theta)} \Rightarrow f(\theta|\mathcal{D}) \propto f(\theta).f(\mathcal{D}|\theta)$$

let $\Gamma(\alpha) = \int_0^\infty e^{-x}.x^{\alpha-1}dx$

$$\Rightarrow \int_0^1 p^{\alpha-1}(1-p)^{\beta-1}dp = \frac{\Gamma(\alpha).\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$\Rightarrow f(p) = p^{\alpha-1}(1-p)^{\beta-1}.\frac{\Gamma(\alpha).\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$\Rightarrow f(\mathcal{D}|p) = \prod_{i=1}^{n} f(p) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$$

$$f(p|\mathcal{D}) = k\, f(p).f(\mathcal{D}|p)$$
$$= k\, p^{\alpha-1+\Sigma x_i}(1-p)^{\beta-1+n-\Sigma x_i}$$

$$p_{MAP} = \underset{\tilde{p}}{\text{argmax}}\, f(\tilde{p}|\mathcal{D}) = \underset{\tilde{p}}{\text{argmax}}\, \ell(\tilde{p})$$

log-likelihood:
$$\ell(p) = \log k + (\alpha-1+\Sigma x_i)\log p + (\beta-1+n-\Sigma x_i)\log(1-p)$$

$$\frac{\partial \ell}{\partial p} = 0 + \frac{\alpha-1+\Sigma x_i}{p} + \frac{\beta-1+n-\Sigma x_i}{1-p} = 0$$

$$\boxed{p_{MAP} = \frac{\tilde{\alpha}-1}{\tilde{\alpha}+\tilde{\beta}-2}}$$ where $\tilde{\alpha} = \alpha+\Sigma x_i$, $\tilde{\beta} = \beta+n-\Sigma x_i$

Ques-8: $f(x|p) = \frac{n!}{\quad} \prod^{k} p_i^{x(i)}$

Ques-8: $f(x|p) = \dfrac{n!}{x^{(1)}!\,x^{(2)}!\ldots x^{(k)}!} \prod_{j=1}^{k} p_j^{x^{(j)}}$

prior: $f(p|\alpha) = \text{Dirichlet}(p, \alpha) = \dfrac{1}{B(\alpha)} \prod_{j=1}^{k} p_j^{\alpha_j - 1}$

where $B(\alpha) = \dfrac{\prod_{j=1}^{k} \Gamma(\alpha_j)}{\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}$

likelihood: $f(\mathcal{D}|p) = \prod_{i=1}^{n} f(x_i|p)$

$$= \prod_{i=1}^{n} \frac{}{x_i^{(1)}!\,x_i^{(2)}!\ldots x_i^{(k)}!} \prod_{j=1}^{k} p_j^{x_i^{(j)}}$$

$$f(\mathcal{D}|p) \propto \prod_{j=1}^{k} p_j^{\sum_{i=1}^{n} x_i^{(j)}}$$

posterior: $f(p|\mathcal{D}) \propto f(p).\,f(\mathcal{D}|p)$

$$f(p|\mathcal{D}) \propto \left(\prod_{j=1}^{k} p_j^{\alpha_j - 1}\right)\left(\prod_{j=1}^{k} p_j^{\sum_{i=1}^{n} x_i^{(j)}}\right)$$

$$f(p|\mathcal{D}) \propto \prod_{j=1}^{k} p_j^{\alpha_j + \sum_{i=1}^{n} x_i^{(j)} - 1}$$

$\therefore$ $f(p|\mathcal{D}) = \text{Dirichlet}(\tilde{\alpha})$ where $\tilde{\alpha} = \alpha + \sum_{i=1}^{n} x_i^{(j)}$

and $\sum_j p_j = 1$

taking log-likelihood:

$$\ell(p) \propto \sum_{j=1}^{k} \left(\alpha + \sum_{i=1}^{n} x_i^{(j)} - 1\right) \log(p_j)$$

Lagrangian: $\mathcal{L}(p, \lambda) \propto \sum_{j=1}^{k} \left(\alpha + \sum_{i=1}^{n} x_i^{(j)}\right) \log p_j + \lambda\left(1 - \sum_j p_j\right)$

$\dfrac{\partial \mathcal{L}}{\partial p_j} = 0$ and $\dfrac{\partial \mathcal{L}}{\partial \lambda} = 0$

we finally get: $p_{\ldots} = \alpha + N - 1$

∂p_j $\partial \lambda$

we finally get : $$\boxed{p_{j(MAP)} = \frac{\alpha + N - 1}{\sum_j \alpha_j + n - k}}$$

where $N = \{N_i\}_{i=1}^{k}$

Ques-9 :

E-step: posterior: $q(z) = P(z|X, \theta^{(t)})$

M-step: $$\theta^{(t+1)} = \underset{\theta^{(t)}}{\operatorname{argmax}} \sum_{i=1}^{n} \sum_{z} q(z_i) \log P(x_i, z_i | \theta^{(t)}) + \log P(\theta^{(t)})$$

The above could be considered as generalized EM for Maximizing Posterior. (MAP).