Assignment 1

CS331 Machine Learning

Name :- Srajan Chourasia
Roll No. :- 2003135
Branch : C.S.E
Date :- 28 Feb 2023

Q1)

Sol. Given : Class Conditional Density

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

2 Class with equal Priors :

$$P_1 = P_2 = 1/2$$

a) we have $f_i(x) > 0 \ \forall \ x \in \mathbb{R}, \ i = 1, 2$

Now,

for class conditional Density to be PDF,

$$\int_{-\infty}^{\infty} f_i(x)\, dx = 1$$

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

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

$$\frac{1}{b}\, dx = dt$$

$$= \int_{-\infty}^{\infty} \left[\frac{1}{\pi \cancel{b}} \times \frac{1}{1 + t^2} \right] \times \cancel{b}\, dt$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dt}{1 + t^2}$$

$$= \frac{1}{\pi} \left[\tan^{-1} t \right]_{-\infty}^{\infty}$$

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

$$= 1/\pi \times \pi$$

$$= 1$$

(b) We will get Bayes Decision Boundary with the usual 0-1 loss function when,

$$q_1(x) = q_2(x)$$

$$C P_1(x) f_1(x) = C P_2(x) f_2(x)$$ where, C is a Constant.

Since,

$$P_1(x) = P_2(x)$$

$\Rightarrow$

$$f_1(x) = f_2(x)$$

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

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

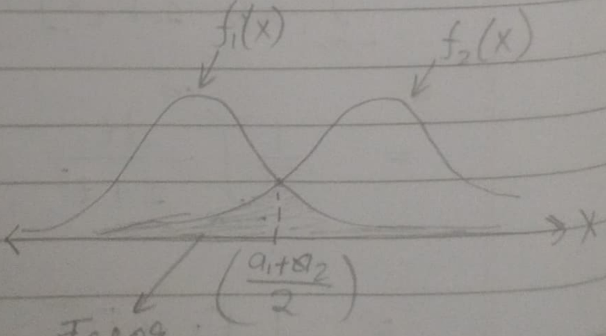
$$(x-a_1)^2 - (x-a_2)^2 = 0$$

$$(x-a_1+x-a_2)(x-a_1-x+a_2) = 0$$

$$\boxed{x = \frac{a_1+a_2}{2}}$$ Bayes Decision Boundary

c) Probability of Error :-

$$P_{error} = \frac{1}{2}\int_{-\infty}^{\frac{a_1+a_2}{2}} f_2(x)\,dx + \frac{1}{2}\int_{\frac{a_1+a_2}{2}}^{\infty} f_1(x)\,dx$$



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

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

Let, $\frac{x-a_2}{b} = m$ and $\frac{x-a_1}{b} = n$

$\frac{dx}{b} = dm$ and $\frac{dx}{b} = dn$

at. $x = \frac{a_1+a_2}{2}$

$m = \frac{a_1-a_2}{2b}$ $\qquad n = \frac{a_2-a_1}{2b}$

$$= \frac{1}{2\pi \cancel{b}}\left\{ \int_{-\infty}^{\frac{a_1-a_2}{2b}} \frac{\cancel{b}\, dm}{1+m^2} + \int_{\frac{a_2-a_1}{2b}}^{\infty} \frac{\cancel{b}\, dn}{1+n^2} \right\}$$

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

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

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

Ans

Q2>

Sol. Given:
$P(h(x) = i \mid x)$

Assume: Consider there are total 'k' classes $\{C_1, C_2, \ldots, C_k\}$

a) Resulting Risk:

$Risk[h(x) = i \mid x]$

risk of choosing i^{th} class, when actual class is Y.

i.e.

$$E[L(h(x), Y) \mid x] \quad \forall \text{ Classes } (k)$$

$$\Rightarrow \sum_{j=1}^{k} p_j(x) \, L(h(x), C_j)$$

where; p_j is the probability of getting j^{th} class.

Now:

$$L(h(x), C_j) = \begin{cases} 0 & \text{if } h(x) = C_j \\ 1 & \text{if } h(x) \neq C_j \end{cases}$$

$$= \underset{C_j = h(x)}{0 \cdot p_j(x)} + \sum_{C_j \neq h(x)} 1 \cdot p_j(x)$$

Since, $\cancel{\Sigma q_j}$ $\sum p_j(x) = 1$

Thus:

$$Risk[h(x) = i \mid x] = 1 - p_j(x) \quad \text{where, } C_j = h(x)$$

b) Improvement in the Classifier:-

- Improve the loss functions as different classes should have different values of loss function.

Ans

Q3>

Sol. Loss Function

$$L(h(x) = i, Y = j) = \begin{cases} 0, & i = j \\ \lambda_r, & i = k+1 \\ \lambda_m, & \text{otherwise} \end{cases}$$

where; $\lambda_r \to$ loss incurred for choosing the $(k+1)^{th}$ class, rejection
$\lambda_m \to$ loss incurred for misclassification.

Now:

$$\text{Risk}\left[h(x) = i \mid x\right] = \begin{cases} E\left[L(h(x), y) \mid x\right] \\ \lambda_r \end{cases}$$

$$= \begin{cases} \sum q_i(x) \cdot \lambda_m \\ \lambda_r \end{cases}$$

if $c_i = h(x)$ {match}

$$= \begin{cases} \lambda_m(1 - q_i(x)) & \text{if misclassification} \\ \lambda_r & \text{if rejection} \end{cases}$$

Now: Minimizing Risk

$$\text{Risk}\left[h(x) = j \mid x\right] \geq \text{Risk}\left[h(x) = i \mid x\right]$$

For Case I :-

$$\lambda_m(1 - q_j(x)) \geq \lambda_m(1 - q_i(x))$$

$$\boxed{q_j(x) \leq q_i(x)}$$

For Case II :-

$$\lambda_r \geq \lambda_m(1 - q_i(x))$$

$$\boxed{q_i(x) \geq 1 - \lambda_r/\lambda_m}$$

Now :- if: $\lambda_r = 0 \Rightarrow q_i(x) = 1$

if $\lambda_r > \lambda_m$ i.e. $\lambda_r/\lambda_m > 1$

$$\Rightarrow \lambda_r/\lambda_m - 1 > 0 \Rightarrow 1 - \frac{\lambda_r}{\lambda_m} < 0$$

$$\Rightarrow q_i(x) \geq 0$$

Ans

Q4)

Sol.

a) Exponential Distribution

Let, $f(x|\lambda) = \begin{cases} \lambda . e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

Assume, data $\mathcal{D}$ contain total n datapoints.

Likelihood Function $L(\theta, \mathcal{D}) = \prod_{i=1}^{n} f(x_i|\theta)$

where, $\theta = \lambda$

$$L(\theta, \mathcal{D}) = \lambda^n e^{-\lambda(\sum_i^n x_i)}$$

Taking Log Both Side:

$$l(\theta, \mathcal{D}) = \ln(\lambda^n) + (-\lambda)\sum_i^n x_i \qquad \text{where, } \theta = \lambda$$

Differentiate both Side w.r.t λ

$$\frac{\partial l(\lambda, \mathcal{D})}{\partial \lambda} = \frac{n}{\lambda} - \sum_i^n x_i$$

$$0 = \frac{n}{\lambda} - \sum_i^n x_i$$

$$\boxed{\lambda_{MLE} = \frac{n}{\sum_i^n x_i}}$$

MLE for exponential Distribution.

Ans

b) Multivariate Gaussian Distribution

Assume: X be d dimensional Vector of features,

Σ = Covariance of matrix

$$f(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|_{det}} \cdot \exp\left\{\frac{-1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right\}$$

where, $\vec{\mu}, \Sigma$ are parameter.

Now: ~~taking log~~ likelihood of the above function.

$$L(\theta, (\vec{\mu}, \Sigma)) = \prod_{i=1}^{n} f(x_i | \mu, \Sigma)$$

taking log of above function:-

$$l(\theta, (\vec{\mu}, \Sigma)) = \sum_{i=1}^{n} -\frac{d}{2}\log(2\pi) - \log(|\Sigma|) - \frac{1}{2}(\vec{x}_i - \vec{\mu})^T \Sigma^{-1} (\vec{x}_i - \vec{\mu})$$

Now: $\frac{dl}{d\mu} = 0$ differentiate and equate it to 0.

$$\Rightarrow \sum_{i=1}^{n} \Sigma^{-1} (\vec{\mu} - x_i) = 0$$

$$\Rightarrow \boxed{\vec{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i}$$

$\therefore \frac{dl}{d\Sigma} = 0$

$$\Rightarrow n\Sigma - \sum_{i=1}^{n} (\vec{x}_i - \vec{\mu})^T (\vec{x}_i - \vec{\mu}) = 0$$

$$\Rightarrow \boxed{\Sigma = \frac{1}{n} \sum_{i=1}^{n} (\vec{x}_i - \vec{\mu})^T (\vec{x}_i - \vec{\mu})}$$

Ans

Q5)

Sol. Given:

$f_X(x) = N(x ; \mu, \sigma^2)$, where μ is the unknown and σ^2 is known.

Assum: $p(\mu) \cong N(\mu_0, \sigma_0^2)$

Likelihood:

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

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

$\therefore$ $f(\mu|D) = C \cdot f(\mu) f(D|\mu)$ where, μ is a Constant

$$\Rightarrow f(\mu|D) \propto \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2\right\} \cdot \exp\left\{\frac{-1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

$$\propto \exp\left\{\frac{-1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2 + \frac{1}{\sigma_0^2}(\mu-\mu_0)^2\right]\right\}$$

$$\propto \exp\left\{\frac{-1}{2}\left(\mu^2\left(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}\right) - 2\mu\left(\frac{\sum x_i}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right) + \left(\frac{\sum x_i^2}{\sigma^2}+\frac{\mu_0^2}{\sigma_0^2}\right)\right)\right\}$$

Taking log liklihood.

$$l(\mu|D) \propto \frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 + \left(\frac{-1}{\sigma_0^2}(\mu-\mu_0)^2\right)$$

Now: differentiating w.r.t "μ" and equat to 0.

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

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

Now: ~~k exp{-1~~

$$f_X(x) = K \cdot \exp\{a\mu^2 + b\mu + c\}$$

where, $a = \frac{-n}{2\sigma^2} - \frac{1}{2\sigma_0^2}$

$$b = \frac{1}{2\sigma^2} \times 2\sum_{i=1}^{n} x_i + \frac{2\mu_0}{2\sigma_0^2}$$

$$c = \frac{-1}{2\sigma^2}\sum x_i^2 - \frac{1}{2\sigma_0^2}\mu_0^2$$

Such that $\int K\exp(a\mu^2 + b\mu + c) = 1$

As posterior and prior have the same form and likelihood, the distribution will be normal and mean and mode will be same.

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

Q6)

Sol.

a) Mixture of Gaussian :

$$f(x,\theta) = \sum_i \lambda_i f_i(x|\theta)$$

$$f_j(x|\theta) \simeq N(\mu_j, \sigma_j)$$

Likelihood $L(\theta) = \sum_{i=1}^{N}\sum_{j=1}^{k} z_{ij} \ln(\lambda_j f_j(x_i|\theta_j))$

Now : E-Step :

$$\gamma_{ij} = \cancel{E[z_{ij}]} \frac{\lambda_j f_j(x_i|\theta_j)}{\sum_k \lambda_k f_k(x_i|\theta_k)}$$

M- Step :

$$\mu_k = \frac{1}{\sum_{i=1}^{N} \gamma_{ij}} \times \sum_{i=1}^{N} \gamma_{ij} x_i$$

$$\Sigma_k = \frac{1}{\sum_{i=1}^{N} \gamma_{ij}} \times \sum_{i=1}^{N} \gamma_{ij} (x_i - \mu_j)^T (x_i - \mu_j)$$

b) Mixture of Bernoulli

$$f(x|\theta) = \lambda_1 f(x|\theta_1) + \lambda_2 f(x|\theta_2) \ldots\ldots\ldots + \lambda_k f(x|\theta_k)$$

Let, $P(z_{ij} = 1) = \lambda_i \quad \forall i$

$$f(x_i = z_{ij}) = \phi(x_i|\theta_j) = P_j^{x_i} (1 - P_j)^{1-x_i}$$

$$f(x_i|z_i, \theta) = \prod_{j=1}^{k} (\lambda_j \phi(x_i|\theta_j))^{z_{ij}}$$

Likelihood: $f(x, \theta) = \prod_{i=1}^{N} \left(\prod_{j=1}^{k} (\lambda_j \phi(x_i|\theta_j))^{z_{ij}} \right)$

Take log:-

$$l(x, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{k} z_{ij} \ln(\lambda_j \phi(x_i|\theta_j))$$

Now: E-Step:

$$\gamma_{ij}^{(t+1)} = \frac{\lambda_j^{(t)} (P_j)^{x_i} (1-P_j)^{1-x_i}}{\sum_{j=1}^{k} \lambda_j^{(t)} (P_j)^{x_i} (1-P_j)^{1-x_i}}$$

M-Step:

$$\therefore \quad p_j^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{ij} x_i}{\sum_{i=1}^{N} \gamma_{ij}}$$

$$\therefore \quad \lambda_j^{(t+1)} = \frac{1}{n} \sum_{i=1}^{N} \gamma_{ij}^{(t)}$$

Q7>

Sol. Derive the MAP estimate of a Bernoulli distribution based on n iid samples. (Conjugate prior : Beta distribution)

$$f(P) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} P^{a-1} (1-P)^{b-1} \quad \text{where, } p \in [0,1], \; a,b \geq 1$$

$$f(P|D) = k\, f(D|P)\, f(P) \quad \text{where, } k \text{ is a Constant}$$

$$= \tilde{k}\, P^{\sum x_i} (1-P)^{n-\sum x_i}\, P^{a-1} (1-P)^{b-1}$$

$$= \tilde{k}\, P^{a+\sum x_i - 1} (1-P)^{b+n-\sum x_i - 1}$$

$$f(P|D) = \text{Beta}(a+\sum x_i,\; b+n-\sum x_i)$$

Take log :

$$\log f(P|D) = \log \tilde{k} + (a+\sum x_i - 1)\log P + (b+n-\sum x_i - 1)\log(1-P)$$

Differentiate w.r.t P and equate it to 0.

$$0 = \frac{(a+\sum x_i - 1)}{P} - \frac{(b+n-\sum x_i - 1)}{1-P}$$

$$\Rightarrow \quad P_{MAP} = \frac{\sum x_i + a - 1}{a+b+n-2}$$

$$P_{MAP} = \frac{\tilde{a} - 1}{\tilde{a} + \tilde{b} - 2} \quad \text{where, } \tilde{a} = a + \sum x_i, \; \tilde{b} = b + n - \sum x_i$$

Ans

Q8)

Sol. Derive ~~generalized EM for~~ MAP the MAP estimate of a multinomial distribution based on n iid Samples.

(Conjugate Prior : Dirichlet Distribution)

Dirichlet Distribution :-

$$f(\theta) = f(x_1, x_2, \ldots, x_K, \alpha_1, \alpha_2, \ldots, \alpha_K) = \frac{1}{B(\alpha)} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

Such that $\sum_{i=1}^{K} x_i = 1$ and $x_i \in [0, 1] \;\forall\, i \in \{1, \ldots, K\}$

$$B(\alpha) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)}, \quad \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$$

Likelihood

$$f(x|\theta) = \frac{n!}{x_1!\, x_2! \ldots x_K!} \prod_{i=1}^{K} \theta_i^{x_i}$$

Now :

$$f(\theta|x, \alpha) = C \cdot f(x|\theta) f(\theta|\alpha) \quad \text{where, } C \text{ is a Constant}$$

$$f(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{K} \theta_i^{\alpha_i - 1}, \quad \text{where}$$

$$B(\alpha) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma(\alpha_0)}, \quad \text{where } \alpha_0 = \sum_{i=1}^{K} \alpha_i$$

$$f(\theta|x, \alpha) = \frac{C \cdot n!}{\prod_{i=1}^{K} x_i!} \prod_{i=1}^{K} \theta_i^{x_i} \frac{1}{B(\alpha)} \prod_{i=1}^{K} \theta_i^{\alpha_i - 1}$$

$$f(\theta|x, \alpha) = C \cdot \frac{n!}{\left(\prod_{i=1}^{K} x_i!\right) B(\alpha)} \prod_{i=1}^{K} \theta_i^{x_i + \alpha_i - 1}$$

Now : ~~Set~~

$$\theta_{MAP} = \underset{\theta}{\operatorname{argmax}}\; f(\theta|x, \alpha)$$

$$\text{s.t} \quad \sum_{i=1}^{K} \theta_i = 1$$

$$f(\theta|x,\alpha) \propto \prod_{i=1}^{k} \theta_i^{x_i+\alpha_i-1} \Rightarrow \log(f(\theta|x,\alpha)) \propto \sum_{i=1}^{k}(x_i+\alpha_i-1)\log\theta_i$$

Now:
$$L(\theta,\lambda) = \tilde{c}\cdot\prod_{i=1}^{k}\theta_i^{x_i+\alpha_i-1} - \lambda\left(\sum_{i=1}^{k}\theta_i - 1\right)$$

$$\cancel{\frac{\partial l}{\partial \theta_i}} \Rightarrow l(\theta,\lambda) = \log\tilde{c} + \sum_{i=1}^{k}(x_i+\alpha_i-1)\log\theta_i - \lambda\left(\sum_{i=1}^{k}\theta_i - 1\right)$$

$$\frac{\partial l}{\partial \theta_i} = (x_i+\alpha_i-1)\times\frac{1}{\theta_i} - \lambda = 0$$

$$\Rightarrow \theta_i = \frac{x_i+\alpha_i-1}{\lambda}$$

$$\frac{\partial l}{\partial \lambda} = -1\left(\sum_{i=1}^{k}\theta_i - 1\right) = 0$$

$$\Rightarrow \sum_{i=1}^{k}\frac{x_i+\alpha_i-1}{\lambda} - 1 = 0$$

$$\Rightarrow \boxed{\lambda = \sum_{i=1}^{k}(x_i+\alpha_i-1)}$$

$$\boxed{\theta_i = \frac{(x_i+\alpha_i-1)}{\sum_{j=1}^{k}(x_j+\alpha_j-1)}}$$

Thus,
$$\theta_{MLE} = \left\{\frac{x_1+\alpha_1-1}{\sum_{j=1}^{k}(x_j+\alpha_j-1)}, \frac{x_2+\alpha_2-1}{\sum_{j=1}^{k}(x_j+\alpha_j-1)}, \ldots\ldots, \frac{x_k+\alpha_k-1}{\sum_{j=1}^{k}(x_j+\alpha_j-1)}\right\}$$

Ans

Q9>

Sol. Derive generalized EM for MAP estimation.

Consider a probabilistic model in which we collectively denote all of the objectserved variables by X and all of the hidden Variables by Z.

The Joint distribution $p(x, z|\theta)$ is governed by a set of parameters denoted by θ.

Goal : Maximize $p(x|\theta) = \sum_z p(x, z|\theta)$

Here, we are assuming z is discrete

Now : $\ln(p(x|\theta)) = L(q, \theta) + KL(q \| qp)$

where, $q(z)$ is a distribution defined over the latent variables.

$$\mathcal{L}(q, \theta) = \sum_z q(z) \ln\left\{\frac{p(x, z|\theta)}{q(z)}\right\}$$

$$KL(q \| p) = -\sum_z q(z) \ln\left\{\frac{p(z|x, \theta)}{q(z)}\right\}$$

$\therefore$ $\ln(p(x, z|\theta)) = \ln(p(z|x, \theta)) + \ln(ap(x|\theta))$

Now :- $$L(q, \theta) = \sum_z p(z|x, \theta^{old}) \ln(p(x, z|\theta)) - \sum_z p(z|x, \theta^{old}) \ln(p(z|x, \theta^{old}))$$

$$= Q(\theta, \theta^{old}) + \text{Constant}$$

Now : $$p(z|x, \theta) = \frac{p(x, z|\theta)}{\sum_z p(x, z|\theta)}$$

$$= \frac{\prod_{i=1}^{N} p(x_i, z_i|\theta)}{\sum_z \prod_{i=1}^{N} p(x_i, z_i|\theta)}$$

$$= \prod_{i=1}^{N} p(z_i|x_i, \theta)$$

Now :

- $\mu_k^{new} = \mu_k^{old} + \left(\frac{\sqrt{}^{new}(z_{mk}) - \sqrt{}^{old}(z_{mk})}{N_k^{new}} \right) (X_m - \mu_k^{old})$

- $N_k^{new} = N_k^{old} + \sqrt{}^{new}(z_{mk}) - \sqrt{}^{old}(z_{mk})$

Ans