Q1. (1) $p(X|\theta) = \sum_{k=1}^{K} \pi_k \, p(x|\mu_k) = \sum_{k=1}^{K} \pi_k \prod_{j=1}^{m} \mu_k[j]^{x[j]} (1-\mu_k[j])^{1-x[j]}$

(2) $L(\theta) = \log p(X|\theta) = \log \prod_{i=1}^{n} p(x^{(i)}|\theta) = \sum_{i=1}^{n} \log p(x^{(i)}|\theta)$

$$= \sum_{i=1}^{n} \log\left(\sum_{k=1}^{K} \pi_k \left(\prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]} \right)\right)$$

(3) $p(x_1, x_2|\theta) = \sum_{k=1}^{K} \pi_k \prod_{j=1}^{m-d} \mu_k[j]^{x_1[j]} (1-\mu_k[j])^{1-x_1[j]} \cdot \prod_{j=m-d+1}^{m} \mu_k[j]^{x_2[j]} (1-\mu_k[j])^{1-x_2[j]}$

$$p(x_1|x_2;\theta) = \frac{p(x_1, x_2|\theta)}{p(x_2|\theta)} = \frac{\sum_{k=1}^{K} \pi_k \, p(x_1|\mu_k)\, p(x_2|\mu_k)}{\sum_{k=1}^{K} \pi_k \, p(x_2|\mu_k)}$$

$$= \sum_{k=1}^{K} \left(\frac{\pi_k \, p(x_2|\mu_k)}{\sum_{i=1}^{K} \pi_i \, p(x_2|\mu_i)} \right) p(x_1|\mu_k)$$

$$= \sum_{k=1}^{K} \pi_k' \cdot p(x_1|\mu_k), \text{ where } \pi_k' = \frac{\pi_k \, p(x_2|\mu_k)}{\sum_{i=1}^{K} \pi_i \, p(x_2|\mu_i)}, \ \forall k.$$

Therefore, new mixing coefficients are

$$\pi_k' = \frac{\pi_k \prod_{j=m-d+1}^{m} \mu_k[j]^{x_2[j]} (1-\mu_k[j])^{1-x_2[j]}}{\sum_{i=1}^{K} \pi_i \prod_{j=m-d+1}^{m} \mu_i[j]^{x_2[j]} (1-\mu_i[j])^{1-x_2[j]}}$$

component density is $p(x_1|\mu_k) = \prod_{j=1}^{m-d} \mu_k[j]^{x_1[j]} (1-\mu_k[j])^{1-x_1[j]}$

(4) $p(z^{(i)}|\pi) = \prod_{k=1}^{K} \pi_k^{z_k^{(i)}}$

$z_k^{(i)} = 1 \Rightarrow p(x^{(i)}|z^{(i)}, \theta) = \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}$.

Generally,

$$\Rightarrow p(x^{(i)}|z^{(i)}, \theta) = \prod_{k=1}^{K} \left(\prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]} \right)^{z_k^{(i)}}$$

$$P(Z,X|\theta) = \prod_{i=1}^{n} P(Z^{(i)}, X^{(i)}|\theta) = \prod_{i=1}^{n} P(Z^{(i)}|\pi) P(X^{(i)}|Z^{(i)},\theta)$$

$$= \prod_{i=1}^{n} \prod_{k=1}^{K} \left(\pi_k \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}\right)^{z_k^{(i)}}$$

(5) $$E(Z_k^{(i)}|X^{(i)},\theta) = \sum_{k=1}^{K} z_k^{(i)} P(Z_k^{(i)}|x^{(i)},\theta)$$

$$= P(Z_k^{(i)}=1|x^{(i)},\theta)$$

$$= \frac{P(X^{(i)}|Z^{(i)},\theta) P(Z_k^{(i)}=1|\theta)}{P(X^{(i)}|\theta)}$$

$$= \frac{\prod_{k=1}^{K}\left(\prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}\right)^{z_k^{(i)}} \cdot \prod_{k=1}^{K} \pi_k^{z_k^{(i)}}}{\sum_{k=1}^{K} \pi_k \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}}$$

$$= \frac{\pi_k \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}}{\sum_{k=1}^{K} \pi_k \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}}$$

(6) $$P(Z,X|\theta) = \prod_{i=1}^{n} \prod_{k=1}^{K} \left(\pi_k \prod_{j=1}^{m} \mu_k[j]^{x^{(i)}[j]} (1-\mu_k[j])^{1-x^{(i)}[j]}\right)^{z_k^{(i)}}$$

$$\ln P(Z,X|\theta) = \sum_{i=1}^{n} \sum_{k=1}^{K} z_k^{(i)} \left(\ln \pi_k + \sum_{j=1}^{m} (x^{(i)}[j] \ln \mu_k[j] + (1-x^{(i)}[j]) \ln(1-\mu_k[j]))\right)$$

$$E(\ln P(Z,X|\theta)|X,\theta) = \sum_{i=1}^{n} \sum_{k=1}^{K} E\left[z_k^{(i)} \left(\ln \pi_k + \sum_{j=1}^{m} (x^{(i)}[j] \ln \mu_k[j] + (1-x^{(i)}[j]) \ln(1-\mu_k[j]))\right)\right]$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} E(z_k^{(i)}|X,\theta) \left(\ln \pi_k + \sum_{j=1}^{m} (x^{(i)}[j] \ln \mu_k[j] + (1-x^{(i)}[j]) \ln(1-\mu_k[j]))\right)$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} \phi(z_k^{(i)}) \cdot \left(\ln \pi_k + \sum_{j=1}^{m} (x^{(i)}[j] \ln \mu_k[j] + (1-x^{(i)}[j]) \ln(1-\mu_k[j]))\right)$$

(7) $E(\ln P(Z, X|\theta) | X, \theta)$

$$= \sum_{i=1}^{n}\sum_{k=1}^{K} \phi(Z_k^{(i)}) \cdot \left(\ln \pi_k + \sum_{j=1}^{m} (X^{(i)}[j] \ln \mu_k[j] + (1-X^{(i)}[j]) \ln (1-\mu_k[j]))\right) \quad ①$$

$$L = ① + \lambda \left(\sum_{k=1}^{K} \pi_k - 1\right)$$

$$\frac{\partial L}{\partial \pi_k} = \sum_{i=1}^{n} \frac{\phi(Z_k^{(i)})}{\pi_k} + \lambda = 0 \Rightarrow \sum_{i=1}^{n} \frac{\phi(Z_k^{(i)})}{\pi_k} = -\lambda, \quad \sum_{i=1}^{n} \phi(Z_k^{(i)}) = -\lambda \pi_k$$

$$\sum_{k=1}^{K} \sum_{i=1}^{n} \phi(Z_k^{(i)}) = -\lambda \sum_{k=1}^{K} \pi_k = -\lambda \Rightarrow \sum_{k=1}^{K} N_k = -\lambda.$$

$$\tilde{\pi}_k^* = \frac{\sum_{i=1}^{n} \phi(Z_k^{(i)})}{\sum_{k=1}^{K} N_k} = \frac{N_k}{\sum_k N_k}$$

$$\frac{\partial L}{\partial \mu_k[j]} = \sum_{i=1}^{n} \phi(Z_k^{(i)}) \cdot \left(\frac{X^{(i)}[j]}{\mu_k[j]} - \frac{1-X^{(i)}[j]}{1-\mu_k[j]}\right) = 0$$

$$\Rightarrow \sum_{i=1}^{n} \phi(Z_k^{(i)}) \left(X^{(i)}[j] - \mu_k[j]\right) = 0.$$

$$\Rightarrow \sum_{i=1}^{n} \phi(Z_k^{(i)}) X^{(i)}[j] = \tilde{\mu}_k[j] \sum_{i=1}^{n} \phi(Z_k^{(i)}) = \tilde{\mu}_k[j] N_k$$

$$\Rightarrow \tilde{\mu}_k[j] = \frac{\sum_{i=1}^{n} \phi(Z_k^{(i)}) X^{(i)}[j]}{N_k}$$

$$\tilde{\mu}_k = \frac{\sum_{i=1}^{n} \phi(Z_k^{(i)}) X^{(i)}}{N_k}$$

Q2. (1) $\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{20+15}{20+15+10+5} = \frac{35}{50} = \frac{7}{10}$

$\text{Precision} = \frac{TP}{TP+FP} = \frac{20}{20+10} = \frac{2}{3}$

$\text{Recall} = \frac{TP}{TP+FN} = \frac{20}{20+5} = \frac{4}{5}$

$F_1\ \text{Score} = \frac{2}{\text{Recall}^{-1}+\text{Precision}^{-1}} = \frac{2}{\frac{5}{4}+\frac{3}{2}} = \frac{8}{5+6} = \frac{8}{11}$

(2) $TP = 540 \quad FP = 23+96+2 = 121 \quad FN = 10+21+14 = 45$

$TN = 1550 - 540 - 121 - 45 = 844$

(3). $K=4$

a: $TP_a = 5 \quad FP_a = 177 \quad FN_a = 57 \quad TN_a = 1311$

b: $TP_b = 540 \quad FP_b = 121 \quad FN_b = 45 \quad TN_b = 844$

c: $TP_c = 436 \quad FP_c = 43 \quad FN_c = 372 \quad TN_c = 699$

d: $TP_d = 87 \quad FP_d = 141 \quad FN_d = 8 \quad TN_d = 1314$.

$\text{Macro Precision} = \frac{1}{4}\left(\frac{5}{5+177} + \frac{540}{540+121} + \frac{436}{436+43} + \frac{87}{87+141}\right)$

$= 0.534$

$\text{Macro Recall} = \frac{1}{4}\left(\frac{5}{5+57} + \frac{540}{540+45} + \frac{436}{436+372} + \frac{87}{87+8}\right)$

$= 0.6148 \approx 0.615$

$\text{Balanced Accuracy} = \frac{1}{4}\left(\frac{5}{62} + \frac{540}{585} + \frac{436}{808} + \frac{87}{95}\right) = 0.6148 \approx 0.615$

$\text{Accuracy} = \frac{5+540+436+87}{1550} = \frac{534}{775} = 0.6890 \neq \text{Balance Accuracy}$

Because they use different formulas, and Accuracy emphasizes on the majority class, Balanced Accuracy treats 4 classes equally.