1. $Z_i \sim \text{categorical}(\pi_1, \pi_2, 1-\pi_1-\pi_2)$

$$(X_i \mid Z_i = 1) \sim \text{Poisson}(\lambda_1)$$
$$(X_i \mid Z_i = 2) \sim \text{Poisson}(\lambda_2)$$
$$(X_i \mid Z_i = 3) \sim \text{Poisson}(\lambda_3)$$

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

a) Let $X = (X_1, \dots, X_{300})$ and $Z = (Z_1, \dots, Z_{300})$

Derive $Q(\theta, \theta^\circ) = E_{Z|X,\theta^\circ}[\log(P(X, Z|\theta))]$

Let $n = 300$

$$P(X, Z \mid \theta) = \prod_{i=1}^{n} P(X_i \mid Z_i, \theta) P(Z_i \mid \theta)$$
$$= \prod_{i=1}^{n} \prod_{k=1}^{3} [P(X_i \mid Z_i = k, \theta) P(Z_i = k \mid \theta)]^{I(Z_i = k)}$$

$$\log P(X, Z \mid \theta) = \sum_{i=1}^{n} \sum_{k=1}^{3} I(Z_i = k)[\log P(X_i \mid Z_i = k, \theta) + \log P(Z_i = k \mid \theta)]$$

$$Q(\theta, \theta^\circ) = \sum_{i=1}^{n} \sum_{k=1}^{3} P(Z_i = k \mid X_i, \theta^\circ)[\log P(X_i \mid Z_i = k, \theta) + \log P(Z_i = k \mid \theta)]$$
$$= \sum_{i=1}^{n} \sum_{k=1}^{3} P(Z_i = k \mid X_i, \theta^\circ)[X_i \log(\lambda_k) - \lambda_k - \log(X_i!) + \log \pi_k]$$

where $\pi_3 = 1 - \pi_1 - \pi_2$

b) Let $\theta^\circ = (\pi_1^\circ, \pi_2^\circ, \lambda_1^\circ, \lambda_2^\circ, \lambda_3^\circ)$

$$P(Z_i = k \mid X_i, \theta^\circ) = \frac{P(Z_i = k, X_i \mid \theta^\circ)}{P(X_i \mid \theta^\circ)}$$

$$= \frac{P(X_i \mid Z_i = k, \theta^\circ)\, P(Z_i = k \mid \theta^\circ)}{\sum_{k'=1}^{3} P(X_i \mid Z_i = k', \theta^\circ)\, P(Z_i = k' \mid \theta^\circ)}$$

$$P(Z_i = 1 \mid X_i, \theta^\circ) = \frac{P(X_i \mid Z_i = 1, \theta^\circ)\, P(Z_i = 1 \mid \theta^\circ)}{P(X_i \mid Z_i = 1, \theta^\circ)\, P(Z_i = 1 \mid \theta^\circ) + P(X_i \mid Z_i = 2, \theta^\circ)\, P(Z_i = 2 \mid \theta^\circ) + P(X_i \mid Z_i = 3, \theta^\circ)\, P(Z_i = 3 \mid \theta^\circ)}$$

$$P(Z_i = 2 \mid X_i, \theta^\circ) = \frac{P(X_i \mid Z_i = 2, \theta^\circ)\, P(Z_i = 2 \mid \theta^\circ)}{P(X_i \mid Z_i = 1, \theta^\circ)\, P(Z_i = 1 \mid \theta^\circ) + P(X_i \mid Z_i = 2, \theta^\circ)\, P(Z_i = 2 \mid \theta^\circ) + P(X_i \mid Z_i = 3, \theta^\circ)\, P(Z_i = 3 \mid \theta^\circ)}$$

$$P(Z_i = 3 \mid X_i, \theta^\circ) = 1 - P(Z_i = 1 \mid X_i, \theta^\circ) - P(Z_i = 2 \mid X_i, \theta^\circ)$$

where $P(X_i \mid Z_i = 1, \theta^\circ) = \dfrac{\lambda_1^{x_i} e^{-\lambda_1}}{x_i!}$, $P(X_i \mid Z_i = 2, \theta^\circ) = \dfrac{\lambda_2^{x_i} e^{-\lambda_2}}{x_i!}$, $P(Z_i = 1 \mid \theta^\circ) = \pi_1$, $P(Z_i = 2 \mid \theta^\circ) = \pi_2$

c) Let $\frac{\partial Q(\theta,\theta^\circ)}{\partial \pi_1}$

$$= \sum_{i=1}^{n} \left[\frac{P(z_i=1|x_i,\theta^\circ)}{\pi_1} - \frac{P(z_i=3|x_i,\theta^\circ)}{1-\pi_1-\pi_2} \right]$$

$$\Rightarrow \frac{(1-\pi_1-\pi_2)\left[\sum_{i=1}^{n} P(z_i=1|x_i,\theta^\circ)\right] - \pi_1\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]}{\pi_1\ (1-\pi_1-\pi_2)} = 0 \quad (1)$$

$$\frac{\partial Q(\theta,\theta^\circ)}{\partial \pi_2} = \sum_{i=1}^{n} \left[\frac{P(z_i=2|x_i,\theta^\circ)}{\pi_2} - \frac{P(z_i=3|x_i,\theta^\circ)}{1-\pi_1-\pi_2} \right]$$

$$\Rightarrow \frac{(1-\pi_1-\pi_2)\left[\sum_{i=1}^{n} P(z_i=2|x_i,\theta^\circ)\right] - \pi_2\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]}{\pi_2\ (1-\pi_1-\pi_2)} = 0 \quad (2)$$

from (1) $\Rightarrow (1-\pi_1-\pi_2)\left[\sum_{i=1}^{n} P(z_i=1|x_i,\theta^\circ)\right] = \pi_1\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]$ (1)'

from (2) $\Rightarrow (1-\pi_1-\pi_2)\left[\sum_{i=1}^{n} P(z_i=2|x_i,\theta^\circ)\right] = \pi_2\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]$ (2)'

(1)' + (2)' $\Rightarrow (1-\pi_1-\pi_2)\left[\sum_{i=1}^{n}\left(P(z_i=1|x_i,\theta^\circ) + P(z_i=2|x_i,\theta^\circ)\right)\right]$

$$= (\pi_1+\pi_2)\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]$$

$$1-\pi_1-\pi_2 = \frac{\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)}{\sum_{i=1}^{n}\left(P(z_i=1|x_i,\theta^\circ)+P(z_i=2|x_i,\theta^\circ)+P(z_i=3|x_i,\theta^\circ)\right)}$$

$$= \frac{\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)}{n}$$

from (1)' $\Rightarrow \hat{\pi}_1 = (1-\pi_1-\pi_2)\left[\sum_{i=1}^{n} P(z_i=1|x_i,\theta^\circ)\right]$

$$= \frac{\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)}{n} \frac{\left[\sum_{i=1}^{n} P(z_i=1|x_i,\theta^\circ)\right]}{\left[\sum_{i=1}^{n} P(z_i=3|x_i,\theta^\circ)\right]}$$

$$= \frac{\sum_{i=1}^{n} P(z_i=1|x_i,\theta^\circ)}{n}$$

from (2)' $\Rightarrow \hat{\pi}_2 = \frac{\sum_{i=1}^{n} P(z_i=2|x_i,\theta^\circ)}{n}$ $\qquad \hat{\pi}_3 = 1-\hat{\pi}_1-\hat{\pi}_2$

Let $\frac{\partial Q(\theta,\theta^\circ)}{\partial \lambda_k} = \sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)\left[-1+\frac{x_i}{\lambda_k}\right]$

$$= \frac{\sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)\,x_i - \lambda_k \sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)}{\lambda_k} = 0$$

$$\Rightarrow \lambda_k \sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ) = \sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)\,x_i$$

$$\Rightarrow \hat{\lambda}_k = \frac{\sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)\,x_i}{\sum_{i=1}^{n} P(z_i=k|x_i,\theta^\circ)}$$

2. $Z_i \sim \text{categorical}(\pi_1, \pi_2, 1-\pi_1-\pi_2)$

$$(X_i \mid Z_i = 1) \sim \text{Poisson}(\lambda_1)$$
$$(X_i \mid Z_i = 2) \sim \text{Poisson}(\lambda_2) \qquad \text{for } i = 1, \dots, 300$$
$$(X_i \mid Z_i = 3) \sim \text{Poisson}(\lambda_3)$$

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad X_i \sim \text{Poisson}(\lambda_2) \text{ for } i = 301, \dots, 400$$

a) Let $X = (X_1, \dots, X_{400})$ and $Z = (Z_1, \dots, Z_{300})$

Derive $Q(\theta, \theta^0) = E_{Z|X,\theta^0}[\log(P(X,Z|\theta))]$

Let $n = 300$

$$P(X,Z \mid \theta) = \prod_{i=1}^{n} \left[P(X_i \mid Z_i, \theta)\, P(Z_i \mid \theta)\right] \prod_{i=301}^{400} P(X_i)$$
$$= \prod_{i=1}^{n} \prod_{k=1}^{3} \left[P(X_i \mid Z_i = k, \theta)\, P(Z_i = k \mid \theta)\right]^{I(Z_i = k)} \prod_{i=301}^{400} P(X_i)$$

$$\log P(X,Z \mid \theta) = \sum_{i=1}^{n} \sum_{k=1}^{3} \left(I(Z_i = k)\left[\log P(X_i \mid Z_i = k, \theta) + \log P(Z_i = k \mid \theta)\right]\right) + \sum_{i=301}^{400} \log P(X_i)$$

$$Q(\theta, \theta^0) = \sum_{i=1}^{n} \sum_{k=1}^{3} P(Z_i = k \mid X_i, \theta^0)\left[\log P(X_i \mid Z_i = k, \theta) + \log P(Z_i = k \mid \theta)\right] + \sum_{301}^{400} \log P(X_i)$$
$$= \sum_{i=1}^{n} \sum_{k=1}^{3} P(Z_i = k \mid X_i, \theta^0)\left[X_i \log(\lambda_k) - \lambda_k - \log(X_i!) + \log \pi_k\right] + \sum_{301}^{400}\left[X_i \log(\lambda_2) - \lambda_2 - \log(X_i!)\right]$$

where $\pi_3 = 1 - \pi_1 - \pi_2$

b) E step:

Let $\theta^\circ = (\pi_1^\circ, \pi_2^\circ, \lambda_1^\circ, \lambda_2^\circ, \lambda_3^\circ)$

$$P(z_i = k \mid X_i, \theta^\circ) = \frac{P(z_i = k, X_i \mid \theta^\circ)}{P(X_i \mid \theta^\circ)}$$

$$= \frac{P(X_i \mid z_i = k, \theta^\circ)\, P(z_i = k \mid \theta^\circ)}{\sum_{k'=1}^{3} P(X_i \mid z_i = k', \theta^\circ)\, P(z_i = k' \mid \theta^\circ)}$$

$$P(z_i = 1 \mid X_i, \theta^\circ) = \frac{P(X_i \mid z_i = 1, \theta^\circ)\, P(z_i = 1 \mid \theta^\circ)}{P(X_i \mid z_i = 1, \theta^\circ) P(z_i = 1 \mid \theta^\circ) + P(X_i \mid z_i = 2, \theta^\circ) P(z_i = 2 \mid \theta^\circ) + P(X_i \mid z_i = 3, \theta^\circ) P(z_i = 3 \mid \theta^\circ)}$$

$$P(z_i = 2 \mid X_i, \theta^\circ) = \frac{P(X_i \mid z_i = 2, \theta^\circ)\, P(z_i = 2 \mid \theta^\circ)}{P(X_i \mid z_i = 1, \theta^\circ) P(z_i = 1 \mid \theta^\circ) + P(X_i \mid z_i = 2, \theta^\circ) P(z_i = 2 \mid \theta^\circ) + P(X_i \mid z_i = 3, \theta^\circ) P(z_i = 3 \mid \theta^\circ)}$$

$$P(z_i = 3 \mid X_i, \theta^\circ) = 1 - P(z_i = 1 \mid X_i, \theta^\circ) - P(z_i = 2 \mid X_i, \theta^\circ)$$

where $P(X_i \mid z_i = 1, \theta^\circ) = \dfrac{\lambda_1^{x_i} e^{-\lambda_1}}{x_i!}$, $P(X_i \mid z_i = 2, \theta^\circ) = \dfrac{\lambda_2^{x_i} e^{-\lambda_2}}{x_i!}$, $P(z_i = 1 \mid \theta^\circ) = \pi_1$, $P(z_i = 2 \mid \theta^\circ) = \pi_2$

M step:

The new summation is removed from all partial derivatives except for $\dfrac{\partial Q(\theta, \theta^\circ)}{\partial \lambda_2}$ so we copy our answers from 1c

$$\Rightarrow \hat{\pi}_1 = \frac{\sum_{i=1}^{n} P(z_i = 1 \mid X_i, \theta^\circ)}{n}$$

$$\Rightarrow \hat{\pi}_2 = \frac{\sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ)}{n} \qquad \hat{\pi}_3 = 1 - \hat{\pi}_1 - \hat{\pi}_2$$

Let $$\frac{\partial Q(\theta, \theta^\circ)}{\partial \lambda_k} = \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ)\left[-1 + \frac{X_i}{\lambda_k}\right] \qquad \text{for } k \neq 2$$

$$= \frac{\sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) X_i - \lambda_k \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ)}{\lambda_k}$$

$$\lambda_k \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) = \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) X_i$$

$$\Rightarrow \hat{\lambda}_k = \frac{\sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) X_i}{\sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ)} \qquad \text{for } k \neq 2$$

$$\frac{\partial Q(\theta, \theta^\circ)}{\partial \lambda_2} = \sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ)\left[-1 + \frac{X_i}{\lambda_2}\right] + \sum_{i=301}^{400}\left[-1 + \frac{X_i}{\lambda_2}\right]$$

$$= \frac{\sum_{i=1}^{n}\left[P(z_i = 2 \mid X_i, \theta^\circ) X_i - \lambda_k \sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ)\right] + \sum_{i=301}^{400}\left[X_i - \lambda_2\right]}{\lambda_2} = 0$$

$$= \frac{\sum_{i=1}^{n}\left[P(z_i = 2 \mid X_i, \theta^\circ) X_i - \lambda_2 \sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ)\right] + \sum_{i=301}^{400} X_i - 100\lambda_2}{\lambda_2} = 0$$

$$\Rightarrow \lambda_2 \sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ) + 100\lambda_2 = \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) X_i + \sum_{i=301}^{400} X_i$$

$$\Rightarrow \lambda_2 \left(\sum_{i=1}^{n}\left[P(z_i = 2 \mid X_i, \theta^\circ)\right] + 100\right) = \sum_{i=1}^{n} P(z_i = k \mid X_i, \theta^\circ) X_i + \sum_{i=301}^{400} X_i$$

$$\Rightarrow \hat{\lambda}_2 = \frac{\sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ) X_i + \sum_{i=301}^{400} X_i}{\sum_{i=1}^{n} P(z_i = 2 \mid X_i, \theta^\circ) + 100}$$