

If given Poisson distribution:

$$g(x; \lambda) = \frac{\lambda^x}{x!} \exp\{-\lambda\}$$

The total likelihood

$$\begin{aligned}\mathcal{L}(D; \theta) &= \prod_i \sum_k p_k g_k(D_i; \lambda_k) \\ \ell(D; \theta) &= \log \prod_i \sum_k p_k g_k(D_i; \lambda_k) = \sum_i \log \sum_k p_k g_k(D_i; \lambda_k)\end{aligned}$$

Assume  $Q(z^{(i)} = j) = P(z^{(i)} = j | x^{(i)}; \mathbf{p}, \boldsymbol{\lambda})$ , which means the prob of  $x^{(i)}$  belongs to  $j^{\text{th}}$  Poisson distribution with given  $\mathbf{p}, \boldsymbol{\lambda}$ .

We can add  $Q(z^{(i)} = j)$  in to the likelihood by:

$$\begin{aligned}\ell(D; \theta) &= \sum_i \log \sum_k Q(z^{(i)} = k) \frac{p_k g_k(D_i; \lambda_k)}{Q(z^{(i)} = k)} \geq i \\ \ell(D; \theta) &\geq \sum_i \sum_k Q(z^{(i)} = k) \log \frac{p_k \frac{\lambda_k^{x^{(i)}}}{x^{(i)}!} \exp\{-\lambda_k\}}{Q(z^{(i)} = k)}\end{aligned}$$

according to the Jensen's theory.

Then for E step, still, we calculate the prob:

$$Q(z^{(i)} = j) = \frac{p_j g(x^{(i)}, \lambda_j)}{\sum_{l=1}^K p_l g(x^{(i)}, \lambda_l)}$$

In M step, we should maximize the likelihood by choose the  $\boldsymbol{\phi}, \boldsymbol{\lambda}$ , means that:

$$\arg \max_{\mathbf{p}, \boldsymbol{\lambda}} \ell(D, \mathbf{p}, \boldsymbol{\lambda})$$

Then we can have the minimum by chose the  $\mathbf{p}, \boldsymbol{\lambda}$  where it's derivative equals 0

First, assuming

$$\begin{aligned}\omega_k^{(i)} &= Q(z^{(i)} = k) \\ \sum_i \sum_k \omega_k^{(i)} \log \frac{p_k \frac{\lambda_k^{x^{(i)}}}{x^{(i)}!} \exp\{-\lambda_k\}}{\omega_k^{(i)}} & \\ \nabla_{\lambda_k} \sum_i \sum_k \omega_k^{(i)} \log \frac{p_k \frac{\lambda_k^{x^{(i)}}}{x^{(i)}!} \exp\{-\lambda_k\}}{\omega_k^{(i)}} &= 0\end{aligned} \tag{1}$$

$$\nabla_{p_k} \sum_i^N \sum_k^K \omega_k^{(i)} \log \frac{p_k \frac{\lambda_k^{x^{(i)}}}{x^{(i)}!} \exp\{-\lambda_k\}}{\omega_k^{(i)}} = 0 \quad (2)$$

By solving equation (1), we can have

$$\begin{aligned} \nabla_{\lambda_k} \sum_i^N \sum_k^K \omega_k^{(i)} (x^{(i)} \log \lambda_k - \lambda_k) &= 0 \\ \sum_i^N \omega_k^{(i)} \left( \frac{x^{(i)}}{\lambda_k} - 1 \right) &= 0 \\ \lambda_k &= \frac{\sum_i^N \omega_k^{(i)} x^{(i)}}{\sum_i^N \omega_k^{(i)}} \end{aligned}$$

By solving equation (2), we need to add normalization term

$$\sum_k^K p_k = 1$$

Then we should simplify the origin function with only terms has  $p_k$ :

$$\sum_i^N \sum_k^K \omega_k^{(i)} \log p_k$$

then use Lagrangian function

$$\begin{aligned} \mathcal{L}(p_k) &= \sum_i^N \sum_k^K \omega_k^{(i)} \log p_k + \beta \left( \sum_k^K p_k - 1 \right) \\ \frac{\partial \mathcal{L}(p_k)}{\partial p_k} &= \sum_i^N \frac{\omega_k^{(i)}}{p_k} + \beta \end{aligned}$$

And

$$\begin{aligned} p_k &= \frac{\sum_i^N \omega_k^{(i)}}{\beta} \\ \sum_k^K p_k = 1 &= \sum_k^K \frac{\sum_i^N \omega_k^{(i)}}{\beta} \\ \beta &= N \end{aligned}$$

So

$$p_k = \frac{\sum_i^N \omega_k^{(i)}}{N}$$