Problem Set 4

P4.5

(a)
$$P(x=k|\theta) = \sum_{j=1}^{K} \pi_j \frac{1}{k!} e^{-\lambda_j} \lambda_j^{x_i}$$

$$P(X|\theta) = \prod_i P(x=k_i|\theta)$$

$$= \prod_i \sum_{j=1}^{k} \pi_j \frac{1}{x_i!} e^{-\lambda_j} \lambda_j^{x_i}$$

$$L(X|\theta) = \sum_i \log \sum_{j=1}^{K} \pi_j \frac{1}{x_i!} e^{-\lambda_j} \lambda_j^{x_i}.$$

add hidden value Z, let $P(Z=j) = \pi_j$, $\sum_{j=1}^{k} \pi_j = 1$

$$L(X|\theta) = \sum_i \log \sum_{j=1}^{K} P(Z=j)\, P(x=k_i|Z=j,\theta_j)$$

$$= \sum_i \log \sum_{j=1}^{K} \frac{P(Z=j)\, P(x=k_i|Z=j,\theta_j)}{P(Z=j|x=k_i,\theta_j)} \cdot P(Z=j|x=k_i,\theta_j)$$

According to Jensen inequality $f[E(x)] \geq E[f(x)]$

So let $y = \dfrac{P(Z=j)\, P(x=k_i|Z=j,\theta_j)}{P(Z=j|x=k_i,\theta_j)}$ $\quad f(y) = \log y$

$$f(E(y)) = \log \sum_{j=1}^{k} P(Z=j|x=k_i,\theta_j) \cdot y \geq E(f(y)) = \sum_{j=1}^{k} P(Z=j|x=k_i,\theta_j) \log y$$

Thus: $$L(X|\theta) \geq \sum_i \sum_{j=1}^{K} P(Z=j|x=k_i,\theta_j) \log \frac{P(Z=j)\, P(x=k_i|Z=j,\theta_j)}{P(Z=j|x=k_i,\theta_j)}$$

Let $$Q(\theta,\theta^t) = \sum_i \sum_{j=1}^{K} P(Z=j|x=k_i,\theta_j^t) \log \frac{P(Z=j)\, P(x=k_i|Z=j,\theta_j)}{P(Z=j|x=k_i,\theta_j^t)}$$

and $$P(Z=j|x=k_i,\theta_j) = \frac{P(x=k_i|Z=j,\theta_j)\, P(Z=j|\theta_j)}{P(x=k_i|\theta_j)}$$

$$= \frac{\pi_j \frac{1}{x_i!} e^{-\lambda_j} \lambda_j^{x_i}}{\sum_{j=1}^{K} \pi_j \frac{1}{x_i!} e^{-\lambda_j} \lambda_j^{x_i}} = Z_{i,j}$$

E-step $\quad \hat{z}_{i,j} = P(z=j \mid x=k_i, \theta^t) = \dfrac{\pi_j \frac{1}{x_i!} e^{-\lambda_j^t} \lambda_j^{t\,x_i}}{\sum_j^k \pi_j \frac{1}{x_i!} e^{-\lambda_j^t} \lambda_j^{t\,x_i}}$

M-step

$$Q(\theta, \theta^t) = \sum_i \sum_{j=1}^k \hat{z}_{i,j}^t \log \frac{\pi_j \frac{1}{k_i!} e^{-\lambda_j} \lambda_j^{k_i}}{\hat{z}_{i,j}^t}, \quad \hat{z}_{i,j}^t \text{ has been known can be a constant}$$

$$= \sum_i \sum_{j=1}^k \hat{z}_{ij}^t \left[\log \pi_j + \log \frac{1}{k_i!} - \lambda_j + x_i \log \lambda_j - \log \hat{z}_{i,j}^t\right]$$

For $\hat{\pi}$: $\underset{\pi}{\operatorname{argmax}} = \underset{\pi}{\operatorname{argmax}} \sum_i \sum_j^k \hat{z}_{i,j}^t \log \pi_j$, and $\sum_{j=1}^k \pi_j = 1$

use Lagrange Multipliers, let $g(\pi_j) = (\sum_{j=1}^k \pi_j) - 1 = 0$

$$L(\pi_j, \lambda) = \sum_i \sum_j^k \hat{z}_{ij}^t \log \pi_j + \lambda\left[\sum_{j=1}^k \pi_j - 1\right]$$

$$\frac{\partial L}{\partial \pi_j} = 0 \qquad \frac{\partial L}{\partial \lambda} = 0$$

so $\pi_j = -\dfrac{\sum_i \hat{z}_{ij}^t}{\lambda} \qquad \sum_{j=1}^k \pi_j = -\dfrac{\sum_i \sum_{j=1}^k \hat{z}_{ij}^t}{\lambda} = 1, \quad P(z=j \mid x=k_i, \theta_j^t) = \hat{z}_{i,j}^t$

$$\lambda = -\sum_i 1 = -n \qquad \hat{\pi}_j = \frac{\sum_i \hat{z}_{i,j}^t}{n}$$

For $\hat{\lambda}$ $\quad \underset{\lambda}{\operatorname{argmax}} = \underset{\lambda}{\operatorname{argmax}} \sum_i \sum_j^k \hat{z}_{i,j}^t \left[-\lambda_j + x_i \log \lambda_j\right]$

$$\frac{\partial \operatorname{argmax}}{\partial \lambda_j} = \sum_i \hat{z}_{i,j}^t \left[-1 + \frac{x_i}{\lambda_j}\right] = 0$$

$$\sum_i \hat{z}_{i,j}^t = \sum_i \hat{z}_{i,j}^t \frac{x_i}{\lambda_j}$$

$$\hat{\lambda}_j = \frac{\sum_i \hat{z}_{i,j}^t x_i}{\sum_i \hat{z}_{ij}^t}$$

To ML estimate for Poisson (PS 2-1) $\hat{\lambda} = \frac{\sum_{i=1}^{N} x_i}{N}$

and in EM: $\hat{\lambda}_j = \frac{\sum_i \hat{z}_{ij} x_i}{\sum_i \hat{z}_{ij}}$ and $\sum_i \hat{z}_{ij} = \hat{N}_j$

$\hat{N}_j$ is the number of points assigned to cluster j

$\sum_i \hat{z}_{ij} x_i$ is the sum of x_i value for those points

$\hat{\lambda}_j$ is the MLE for a Poisson in cluster j

b) To the London $K \in \{1, 2, 3, 4, 5\}$

the λ_j is

	1	2	3	4	5
0	0.928819	0.000000	0.000000	0.000000	0.000000
1	0.858622	1.011981	0.000000	0.000000	0.000000
2	0.982007	0.982009	0.740060	0.000000	0.000000
3	0.820118	1.011811	1.011802	0.974591	0.000000
4	0.990750	0.791180	0.990893	0.990880	0.990776

as shown in table.

no matter k, λ_j is near the 1,

Thus we can consider boom hit in London is random

To the Anterwep $K \in \{1, 2, 3, 4, 5\}$

the λ_j is

	1	2	3	4	5
0	0.895833	0.000000	0.000000	0.000000	0.00000
1	0.229790	2.195365	0.000000	0.000000	0.00000
2	0.021582	0.553265	2.374450	0.000000	0.00000
3	0.002849	2.376010	0.530386	0.530345	0.00000
4	2.195365	2.195365	2.195365	2.195365	0.22979

as shown in table

it revealed that some componets mean is near 2. even above 2. which can be considered that there some highly probablity places are the deliberately target.

57558749

王椋辰