Expectation Maximization pour les mélanges de lois de Poisson

Considérons un mélange de K lois de Poisson. Pour $1 \leq k \leq K$, nous noterons $\lambda_k > 0$ le paramètre de la k-ème composante et $\pi_k \in (0,1)$ son poids. Notons $\theta = (\pi_1, \dots, \pi_K, \lambda_1, \dots, \lambda_K)$ et

$$\Theta = \left\{ \theta = (\pi_1, \dots, \pi_K, \lambda_1, \dots, \lambda_K)\,;\, \forall k \in \{1, \dots, K\}\,,\, \pi_k \in (0,1)\,,\, \lambda_k > 0\,,\, \sum_{k=1}^{K} \pi_k = 1 \right\}.$$

1. Soit $\theta \in \Theta$, expliquer comment construire une variable aléatoire X suivant un mélange de lois de Poisson paramétré par θ.

2. Notons $\mathbb{P}_\theta$ la loi de X. Pour tout $j \geq 0$, calculer $\mathbb{P}_\theta(X = j)$.

3. Soit $\theta \in \Theta$ et $(x_1, \dots, x_n) \in \mathbb{N}^n$. Calculer $\log \mathbb{P}_\theta(X_{1:n} = x_{1:n})$ où les $(X_i)_{1 \leq i \leq n}$ sont i.i.d de même loi que X.

4. Puisque nous ne pouvons pas maximiser la logvraisemblance explicitement, nous allons utiliser l'algorithme Expectation Maximization.

   (a) Pour tout $\theta \in \Theta$ et tout $k \in \{1, \dots, K\}$, calculer $\mathbb{P}_\theta(Z = k | X = j)$.

   (b) Calculer la logvraisemblance complète des données.

   (c) Calculer la quantité intermédiaire de l'algorithme EM.

   (d) En déduire la mise à jour d'une itération de l'algorithme EM.

   (e) Détailler le fonctionnement complet de l'algorithme EM

   (f) Cet algorithme converge t'il vers le maximum de vraisemblance ?