Part 2e: Mixture models

<u>Textbook</u>: p. 46-48

Introduction to mixture models

- The distribution function in a mixture model is a convex combination of distribution functions.
- In order to simulate from a mixture model it is convenient to introduce an hierarchical structure.

Definition of a mixture model

Definition 2.6 Let $P_1, \dots, P_k$ be probability distributions on $\mathbb{R}^d$ and let $\theta_1, \dots, \theta_k > 0$ such that $\sum_{a=1}^{k} \theta_a = 1$. Then the *mixture* P_θ of the distributions $P_1, \dots, P_k$ with weights $\theta_1, \dots, \theta_k$ is given by

$$P_\theta(A) = \sum_{a=1}^{k} \theta_a P_a(A)$$

for all $A \subseteq \mathbb{R}^d$.

Lemma 2.8 Assume that $P_1, \dots, P_k$ have densities $f_1, \dots, f_k$. Then the mixture distribution P_θ also has a density which is given by

$$f_\theta = \sum_{a=1}^{k} \theta_a f_a.$$

$$\begin{aligned} F_\theta(x) &= P_\theta((-\infty, x]) \\ &= \sum_{a=1}^{k} \theta_a P_a((-\infty, x]) \\ &= \sum_{a=1}^{k} \theta_a F_a(x) \end{aligned}$$

$$\begin{aligned} f_\theta(x) &= \frac{\partial}{\partial x} F_\theta(x) \\ &= \frac{\partial}{\partial x} \sum_{a=1}^{k} \theta_a F_a(x) \\ &= \sum_{a=1}^{k} \theta_a \frac{\partial}{\partial x} F_a(x) \end{aligned}$$

$$= \sum_{a=1} \theta_a \frac{\partial}{\partial x} f_a(x)$$

$$= \sum_{a=1}^{k} \theta_a f_a(x)$$

Simulating a mixture model

Algorithm 2.9 (mixture distributions)

input:

probability distributions $P_1, \ldots, P_k$

weights $\theta_1, \ldots, \theta_k > 0$ with $\sum_{a=1}^{k} \theta_a = 1$

randomness used:

$Y \in \{1, 2, \ldots, k\}$ with $P(Y = a) = \theta_a$ for all a

samples $X \sim P_a$ for different $a \in \{1, \ldots, k\}$

(X, Y)

output:

$X \sim P_\theta$

generate $Y \in \{1, 2, \ldots, k\}$ with $P(Y = a) = \theta_a$ for all a

generate $X \sim P_Y$

return X

Lemma 2.10 The sample X constructed by algorithm 2.9 is distributed according to the mixture distribution from definition 2.6, that is $X \sim P_\theta$.

$$(X, Y) \sim f_{X,Y}(x,y) = P_Y(y) \cdot f_{X|Y}(x|y) = \theta_y \cdot f_y(x)$$

$$X \sim \sum_{a=1}^{k} f_{X,Y}(x,a) = \sum_{a=1}^{k} \theta_a f_a(x)$$