# Homework problem set 6

Your homework must be handed in **electronically via Moodle before Wednesday October 12th, 22:00h**. This deadline is strict and late submissions are graded with a 0. At the end of the course, the lowest of your 7 weekly homework grades will be dropped. You are strongly encouraged to work together on the exercises, including the homework. However, after this discussion phase, you have to write down and submit your own individual solution. Numbers alone are never sufficient, always motivate your answers.

## Problem 1

Let $Y$ be a categorical random variable over $n$ components:

$$Y \sim \text{Cat}(w_1, \dots, w_n),$$

where of course $\sum_i w_i = 1$.

## Problem 2: EM in a mixture of Poissons

Suppose you observe the data set

$$x_1^N = \{1, 5, 5, 1, 5, 7, 7, 2, 2, 5\}$$

You decide to model the data with a mixture model with 3 unobserved components, whose weights you will have to estimate in this exercise using Expectation Maximization. (As for notation; we have $N = 10$ datapoints $x_1, \dots, x_N$ which will be indexed by $i$ and we have $M = 3$ mixture components $c_1, c_2, c_3$ indexed by $j$.)

Our model assumes that for every datapoint $X_i$, there is a unobserved categorical variable $Y_i$ that takes as its value one of the three components $c_1, c_2, c_3$ with (unknown) probabilities $w_1, w_2$ and $w_3$ respectively. So $p(Y_i = c_j) = w_j$. For every component there then exists a Poisson distribution where the observed $x_i$ is then drawn from. The three Poisson distributions have (unknown) parameters $\lambda_{c_1}, \lambda_{c_2}$ and $\lambda_{c_3}$. We will collect all parameters in $\theta = (w_1, w_2, w_3, \lambda_{c_1}, \lambda_{c_2}, \lambda_{c_3})$, so we can concisely summarize all this as

$$
\begin{aligned}
X_i \mid Y_i = c_1, \Theta = \theta &\sim \text{Poisson}(\lambda_{c_1}) \\
X_i \mid Y_i = c_2, \Theta = \theta &\sim \text{Poisson}(\lambda_{c_2}) \\
X_i \mid Y_i = c_3, \Theta = \theta &\sim \text{Poisson}(\lambda_{c_3}) \\
Y_i \mid \Theta = \theta &\sim \text{Categorical}(w_1, w_2, w_3).
\end{aligned}
$$

Furthermore, we assume independence between the mixture components. Thus, your model is

$$P(X_1^N = x_1^N, Y_1^N = y_1^N \mid w)$$

$$= \prod_{i=1}^N P(Y_i = y_i) \cdot P(X_i = x_i \mid Y_i = y_i, w).$$

Again, the $X_i$ are observed Poisson variables, with natural numbers as outcomes and the $Y_i$ are unobserved categorical random variables that take values $y_i \in \{c_1, c_2, c_3\}$.

Expectation-Maximization starts from some initial guesses $\theta^{(0)}$ of the model parameters $\theta$. We use the following:

$$w_1^{(0)} = w_2^{(0)} = w_3^{(0)} = \frac{1}{3}, \quad \lambda_{c_1} = 0.5, \quad \lambda_{c_2} = 1, \quad \lambda_{c_3} = 1.5,$$

or equivalently $\theta^{(0)} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, .5, 1, 1.5\right)$.

**(a)** Compute the marginal log-likelihood of the data under the initial parameter settings, i.e. compute $P(X_1^N = x_1^N \mid \Theta = \theta^{(0)})$.

**(b)** Compute the posterior probabilities for the mixture components per data point, i.e, compute

$$P\big(Y_i = c_j \mid X_i = x_i, \Theta = \theta^{(0)}\big) \quad \text{for } j = 1, 2, 3$$

**(c)** With the posterior probabilities in place, we can perform the E-step. Report the expected sufficient statistics $\mathbb{E}[c = c_j \mid w^{(0)}]$ for each $1 \leq i \leq 3$.

**(d)** Report the expected sufficient statistics $\mathbb{E}[x = v \mid c_i, \lambda]$ for $v \in \{2, 3, 4, 5, 6, 7, 8\}$. Use a table for this exercise!

**(e)** Perform the M-step. Report the new parameters of the distributions $P_Y$ and $P_{X|Y_i}$ for $1 \leq i \leq 5$.

**(f)** Compute the marginal log-likelihood of $x_1^N$ under the new parameter settings $\theta^{(1)}$.