8 Machine Learning and Bayesian Inference (SBH)

Evil Robot has decided to become a gambling cheat. He has a biased coin with $\Pr(\text{head}) = p$ and two dice. The first die is biased with $\Pr(n) = p_n$ for the nth outcome with $n \in \{1, 2, 3, 4, 5, 6\}$. The second die is also biased, and has different numbers: its distribution is $\Pr(n) = q_n$ with $n \in \{4, 5, 6, 7, 8, 9\}$.

Evil Robot flips the coin. If he gets a head then he rolls the first die, otherwise he rolls the second. He then tells you the outcome. You see only the number obtained and nothing else. He does this m times, so you observe a sequence of m numbers in the range 1 to 9. Your aim is to estimate p and the distributions of each die, given the m numbers. In the following, $\mathbf{n}$ is the vector of m observed numbers $(n_1 \quad \cdots \quad n_m)^T$, θ is the set of parameters $\{p, p_1, \dots, p_6, q_4, \dots, q_9\}$ and we define $q = 1 - p$.

(a) Write down an expression for the distribution $\Pr(n|\theta)$ where $n \in \{1, \dots, 9\}$. [2 marks]

(b) Define the variable

$$z_i = \begin{cases} 1 & \text{if } n_i \text{ was obtained by rolling die 1} \\ 0 & \text{otherwise.} \end{cases}$$

and let $\mathbf{z}$ denote the corresponding vector with m values. Write down an expression for $\log \Pr(\mathbf{n}, \mathbf{z}|\theta)$. [3 marks]

(c) Describe the EM algorithm for maximizing likelihood in a problem involving *latent variables.* [3 marks]

(d) Show that, with the distribution $\Pr(\mathbf{z}|\mathbf{n}, \theta)$,

$$E(z_i) = \begin{cases} 1 & \text{if } n_i \in \{1, 2, 3\} \\ 0 & \text{if } n_i \in \{7, 8, 9\} \\ \frac{pp_{n_i}}{pp_{n_i} + qq_{n_i}} & \text{otherwise.} \end{cases}$$

[4 marks]

(e) Define $\gamma_i = E(z_i)$ as in Part (d). By applying the EM algorithm to this problem, show that you can estimate the parameters in θ using the following updates

$$p = \frac{\gamma}{m}$$
$$p_n = \frac{1}{\gamma} \sum_{\{i|n=n_i\}} \gamma_i$$
$$q_n = \frac{1}{m - \gamma} \sum_{\{i|n=n_i\}} (1 - \gamma_i)$$

where $\gamma = \sum_{i=1}^{m} \gamma_i$. [8 marks]