1) [CA session] Problem 1: The Bayesian Bag of Candies Model

You have a lot of candy left over from Halloween, and you decide to give them away to your friends. You have four types of candy: Apple, Banana, Caramel, Dark-Chocolate. You decide to prepare candy bags using the following process.

- For each candy bag, you first flip a (biased) coin $Y$ which comes up heads ($Y = H$) with probability $\lambda$ and tails ($Y = T$) with probability $1 - \lambda$.

- If $Y$ comes up heads ($Y = H$), you make a Healthy bag, where you:
  (a) Add one Apple candy with probability $p_1$ or nothing with probability $1 - p_1$;
  (b) Add one Banana candy with probability $p_1$ or nothing with probability $1 - p_1$;
  (c) Add one Caramel candy with probability $1 - p_1$ or nothing with probability $p_1$;
  (d) Add one Dark-Chocolate candy with probability $1 - p_1$ or nothing with probability $p_1$.

- If $Y$ comes up tails ($Y = T$), you make a Tasty bag, where you:
  (a) Add one Apple candy with probability $p_2$ or nothing with probability $1 - p_2$;
  (b) Add one Banana candy with probability $p_2$ or nothing with probability $1 - p_2$;
  (c) Add one Caramel candy with probability $1 - p_2$ or nothing with probability $p_2$;
  (d) Add one Dark-Chocolate candy with probability $1 - p_2$ or nothing with probability $p_2$.

For example, if $p_1 = 1$ and $p_2 = 0$, you would deterministically generate: Healthy bags with one Apple and one Banana; and Tasty bags with one Caramel and one Dark-Chocolate. For general values of $p_1$ and $p_2$, bags can contain anywhere between 0 and 4 pieces of candy.

Denote $A, B, C, D$ random variables indicating whether or not the bag contains candy of type Apple, Banana, Caramel, and Dark-Chocolate, respectively.
(a)

(i) Draw the Bayesian network corresponding to process of creating a single bag.

(ii) What is the probability of generating a Healthy bag containing Apple, Banana, Caramel, and not Dark-Chocolate? For compactness, we will use the following notation to denote this possible outcome:

$$(\text{Healthy}, \{\text{Apple, Banana, Caramel}\}).$$

(iii) What is the probability of generating a bag containing Apple, Banana, Caramel, and not Dark-Chocolate?

(iv) What is the probability that a bag was a Tasty one, given that it contains Apple, Banana, Caramel, and not Dark-Chocolate?
(b)

You realize you need to make more candy bags, but you’ve forgotten the probabilities you used to generate them. So you try to estimate them looking at the 5 bags you’ve already made:

\[\begin{align*}
\text{bag } 1 & : \quad \text{(Healthy, \{Apple, Banana\})} \\
\text{bag } 2 & : \quad \text{(Tasty, \{Caramel, Dark-Chocolate\})} \\
\text{bag } 3 & : \quad \text{(Healthy, \{Apple, Banana\})} \\
\text{bag } 4 & : \quad \text{(Tasty, \{Caramel, Dark-Chocolate\})} \\
\text{bag } 5 & : \quad \text{(Healthy, \{Apple, Banana\})}
\end{align*}\]

Estimate \(\lambda, p_1, p_2\) by maximum likelihood.

Estimate \(\lambda, p_1, p_2\) by maximum likelihood, using Laplace smoothing with parameter 1.
You find out your little brother had been playing with your candy bags, and had mixed them up (in a uniformly random way). Now you don’t even know which ones were Healthy and which ones were Tasty. So you need to re-estimate $\lambda, p_1, p_2$, but now without knowing whether the bags were Healthy or Tasty.

\[\text{bag 1 : (, \{Apple, Banana, Caramel\})}\]
\[\text{bag 2 : (, \{Caramel, Dark-Chocolate\})}\]
\[\text{bag 3 : (, \{Apple, Banana, Caramel\})}\]
\[\text{bag 4 : (, \{Caramel, Dark-Chocolate\})}\]
\[\text{bag 5 : (, \{Apple, Banana, Caramel\})}\]

You remember the EM algorithm is just what you need. Initialize with $\lambda = 0.5, p_1 = 0.5, p_2 = 0$, and run one step of the EM algorithm.

(i) E-step:

(ii) M-step:
You decide to make candy bags according to a new process. You create the first one as described above. Then with probability $\mu$, you create a second bag of the same type as the first one (Healthy or Tasty), and of different type with probability $1 - \mu$. Given this type, the bag is filled with candy as before. Then with probability $\mu$, you create a third bag of the same type as the second one (Healthy or Tasty), and of different type with probability $1 - \mu$. And so on, you repeat the process $M$ times. Denote $Y_i, A_i, B_i, C_i, D_i$ the variables at each time step, for $i = 0, \ldots, M$. Let $X_i = (A_i, B_i, C_i, D_i)$.

Now you want to compute:

$$\mathbb{P}(Y_i = \text{Healthy} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0))$$

exactly for all $i = 0, \ldots, M$, and you decide to use the forward-backward algorithm. Suppose you have already computed the marginals:

$$f_i = \mathbb{P}(Y_i = \text{Healthy} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0))$$

for some $i \geq 0$. Recall the first step of the algorithm is to compute an intermediate result proportional to

$$\mathbb{P}(Y_{i+1} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0), X_{i+1} = (1, 1, 1, 0))$$

(i) Write an expression that is proportional to

$$\mathbb{P}(Y_{i+1} = \text{Healthy} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0), X_{i+1} = (1, 1, 1, 0))$$

in terms of $f_i$ and the parameters $p_1, p_2, \lambda, \mu$. 


(ii) Write an expression that is proportional to
\[ P(Y_{i+1} = \text{Tasty} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0), X_{i+1} = (1, 1, 1, 0)) \]
in terms of \( f_i \) and the parameters of the model \( p_1, p_2, \lambda, \mu \). The proportionality constant should be the same as in (i).

(iii) Let \( h \) be the answer for part (i), and \( t \) for part (ii). Write an expression for
\[ P(Y_{i+1} = \text{Healthy} \mid X_0 = (1, 1, 1, 0), \ldots, X_i = (1, 1, 1, 0), X_{i+1} = (1, 1, 1, 0)) \]
in terms of \( h, t \) and the parameters of the model \( p_1, p_2, \lambda, \mu \).