ML HW4

姓名：陳以瑄 學號：313551047

Part 3 EM algorithm (Handwritten)

Given two coins (C_0, C_1) and in each trial, a certain coin is chosen and tossed for three times. There are totally three trial outcomes: {HHH, HHT, TTT}. Note that the chance of choosing C_0 is k, the chance of C_0 showing H is P_0, and the chance of C_1 showing H is P_1. Use EM algorithm to update the estimates of the parameters for just one round. You should show the process step by step with the initial values of parameters: k = 0.5, P_0 = 0.6, P_1 = 0.1

$x = \{ HHH, HHT, TTT \}$

$w = \{ w_1, w_2, w_3 \}$

$\theta^{(0)} = \{ k^{(0)} = 0.5, \; P_0^{(0)} = 0.6, \; P_1^{(0)} = 0.1 \}$

the trial ~ $Bin(N, p)$. the pdf $Bin(m \mid N, p) = \binom{N}{m} p^m (1-p)^{N-m}$

in this example . $N=3$, $m = \{m_1, m_2, m_3\} = \{3, 2, 0\}$

The joint likelihood $P = \prod_i \left(k \binom{N}{m_i} P_0^{m_i} (1-P_0)^{N-m_i}\right)^{w_i} \left((1-k) \cdot \binom{N}{m_i} P_1^{m_i} (1-P_1)^{N-m_i}\right)^{1-w_i}$

weighted log likelihood $J = \sum_i w_i \left[\log k + \log \binom{N}{m_i} + m_i \log P_0 + (N-m_i) \log(1-P_0)\right] + (1-w_i)\left[\log(1-k) + \log\binom{N}{m_i} - m_i \log P_1 + (N-m_i)\log(1-P_1)\right]$

To find $\underset{\theta}{\arg\max}\, J(\theta)$

$\frac{\partial J}{\partial k} = \frac{\partial}{\partial k}\left(\sum_i w_i \log k + \sum_i (1-w_i) \log(1-k)\right)$

$= \frac{\partial}{\partial k} \log k \sum_i w_i + \frac{\partial}{\partial k} \log(1-k) \sum_i (1-w_i)$

$= \frac{1}{k} \cdot \sum_i w_i + \frac{-1}{1-k} \sum_i (1-w_i) = 0$

$k \cdot \sum_i (1-w_i) = (1-k) \sum_i w_i = \sum_i w_i - k \sum_i w_i$ $\quad (1-w_i)\left[\log(1-k) + \log\binom{N}{m_i} - m_i \log P_1 + (N-m_i)\log(1-P_1)\right]$

$\sum_i w_i = k \sum_i 1 - w_i + w_i = Nk$

$k = \frac{\sum_i w_i}{N}$

$\frac{\partial J}{\partial P_0} = \frac{\partial}{\partial P_0}\left(\sum_i w_i \cdot \left[m_i \log P_0 + (N-m_i)\log(1-P_0)\right]\right)$

$= \frac{\partial}{\partial P_0} \log P_0 \sum_i w_i m_i + \frac{\partial}{\partial P_0} \log(1-P_0) \sum_i w_i (N-m_i)$

$= \frac{1}{P_0} \sum_i w_i m_i + \frac{-1}{1-P_0} \sum_i w_i (N-m_i) = 0$

$P_0 \cdot \sum_i w_i (N-m_i) = (1-P_0) \sum_i w_i m_i = \sum_i w_i m_i - P_0 \sum_i w_i m_i$

$\sum_i w_i m_i = P_0 \sum_i w_i (N - \cancel{m_i}) + \cancel{w_i m_i} = P_0 \sum_i w_i N$

$P_0 = \frac{\sum_i w_i m_i}{\sum_i w_i N}$

$\frac{\partial J}{\partial P_1} = \frac{\partial}{\partial P_1}\left(\sum_i (1-w_i)\left[m_i \log P_1 + (N-m_i)\log(1-P_1)\right]\right)$

$= \frac{\partial}{\partial P_1} \log P_1 \sum_i (1-w_i) m_i + \frac{\partial}{\partial P_1} \log(1-P_1) \sum_i (1-w_i)(N-m_i)$

$= \frac{1}{P_1} \sum_i (1-w_i) m_i + \frac{-1}{1-P_1} \sum_i (1-w_i)(N-m_i) = 0$

$P_1 \sum_i (1-w_i)(N-m_i) = (1-P_1) \sum_i (1-w_i) m_i = \sum_i (1-w_i) m_i - P_1 \sum_i (1-w_i) m_i$

$\sum_i (1-w_i) m_i = P_1 \sum_i (1-w_i)(N - \cancel{m_i}) + \cancel{(1-w_i) m_i} = P_1 \sum_i (1-w_i) N$

$P_1 = \frac{\sum_i (1-w_i) m_i}{\sum_i (1-w_i) N}$

$x = \{HHH, HHT, TTT\}$

$N=3$, $m = \{m_1, m_2, m_3\} = \{3, 2, 0\}$

$\theta^{(0)} = \{ k^{(0)} = 0.5,\ P_0^{(0)} = 0.6,\ P_1^{(0)} = 0.1 \}$

$k^{(n+1)} = \frac{\sum_i w_i^{(n)}}{N}$, $P_0^{(n+1)} = \frac{\sum_i w_i^{(n)} m_i}{N \sum_i w_i^{(n)}}$, $P_1^{(n+1)} = \frac{\sum_i (1-w_i^{(n)}) m_i}{N \sum_i (1-w_i^{(n)})}$

E-step:

$P(Z_i = C_0, m_i = 3 \mid \theta^{(0)}) = \binom{3}{3} 0.6^3 (1-0.6)^0 = 0.216$

$P(Z_i = C_1, m_i = 3 \mid \theta^{(0)}) = \binom{3}{3} 0.1^3 (1-0.1)^0 = 0.001$

$w_1 = \frac{0.216}{0.216 + 0.001} = 0.9954$

$P(Z_i = C_0, m_i = 2 \mid \theta^{(0)}) = \binom{3}{2} 0.6^2 (1-0.6)^1 = 0.432$

$P(Z_i = C_1, m_i = 2 \mid \theta^{(0)}) = \binom{3}{2} 0.1^2 (1-0.1)^1 = 0.027$

$w_2 = \frac{0.432}{0.432 + 0.027} = 0.9412$

$P(Z_i = C_0, m_i = 0 \mid \theta^{(0)}) = \binom{3}{0} 0.6^0 (1-0.6)^3 = 0.064$

$P(Z_i = C_1, m_i = 0 \mid \theta^{(0)}) = \binom{3}{0} 0.1^0 (1-0.1)^3 = 0.729$

$w_3 = \frac{0.064}{0.064 + 0.729} = 0.0807$

M-Step:

$\sum_i w_i = 0.9954 + 0.9412 + 0.0807 = 2.0173$, $\sum_i^3 (1-w_i) = 3 - \sum_i w_i = 0.9827$

$\sum_i w_i m_i = 0.9954 \times 3 + 0.9412 \times 2 + 0.0807 \times 0 = 4.8686$

$\sum_i (1-w_i) m_i = 0.0046 \times 3 + 0.0588 \times 2 + 0.9193 \times 0 = 0.1314$

$k^{(1)} = \frac{\sum_i w_i^{(0)}}{N} = \frac{2.0173}{3} = 0.6724$

$P_0^{(1)} = \frac{\sum_i w_i^{(0)} m_i}{N \sum_i w_i^{(0)}} = \frac{4.8686}{3 \times 2.0173} = 0.8045$

$P_1 = \frac{\sum_i (1-w_i^{(0)}) m_i}{N \sum_i (1-w_i^{(0)})} = \frac{0.1314}{3 \times 0.9827} = 0.0446$