Exercise 11-1

Compute the maximum θ^{new} of the function $Q(\theta, \theta^{old}) = \sum_{z \in heads, tails} P(Z=z|X=x, \theta^{old}) \log P(X=x, Z=z|\theta)$

We split z into z_1 and z_2, where z_1 = heads and z_2 = tails

$$Q(\theta, \theta^{old}) = \sum_{i \in z_1} P(Z=heads|X=x_i, \theta^{old}) \log P(X=x_i, Z=heads|\theta)$$

$$+ \sum_{k \in z_2} P(Z=tails|X=X_k, \theta^{old}) \log P(X=X_k, Z=tails|\theta)$$

$$= \sum_{i \in z_1} P(Z=head|X=x_i, \theta^{old}) \log(\lambda \prod_{j=1}^{m} P(X_j = X_j^{(i)})|Z=heads, \theta)$$

$$+ \sum_{k \in z_2} P(Z=tails|X=X_k, \theta^{old}) \log(1-\lambda) \prod_{l=1}^{m} P(X_k = X_k^{l})|Z=tails, \theta)$$

$$= \sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old})(\log(\lambda) + \log \prod_{j=1}^{m} P(X_j = X_j^{(i)})|Z=heads, \theta)$$

$$+ \sum_{k \in z_2} P(Z=tails|X=X_k, \theta^{old})(\log(1-\lambda) + \log(\prod_{l=1}^{m} P(X_k = X_k^{l})|Z=tails, \theta))$$

$$\frac{\partial}{\partial \lambda} Q(\theta, \theta^{old}) = \sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old}) \frac{1}{\lambda} + \sum_{k \in z_2} P(Z=tails|X=X_k, \theta^{old}) \frac{1}{1-\lambda} = 0$$

$$0 = \sum_{i \in z_1} P(Z=heads|X=x_i, \theta^{old}) \frac{1}{\lambda} + \sum_{k \in z_2} P(Z=tails|X=X_k, \theta^{old}) \frac{1}{1-\lambda}$$ to find the maximum

$$\hat{\lambda} := \frac{\sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old})}{\sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old}) \sum_{k \in z_2} P(Z=tails|X=X_k|\theta^{old})}$$

Compute $\hat{P}_1$ with $\frac{\partial}{\partial P_1} Q(\theta, \theta^{old}) = 0$ We denote the numbers of times heads came up in x_i by h_i and the number of tails by t_i

$$\sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old}) \log(\lambda P_1^{h_i}(1-P_1)^{t_i})$$

$$\sum_{i \in z_1} P(Z=heads|X=x_i, \theta^{old}) \log(\lambda) + h_i \log(P_1) + t_i \log(1-P_1))$$

$$\frac{\partial}{\partial P_1} Q(\theta, \theta^{old}) = \sum_{i \in z_1} P(Z=heads|X=X_i, \theta^{old}) \left(\frac{h_i}{P_1} + \frac{t_i}{1-P_1}\right) = 0$$

$$0 = \sum_{i \in z_1} P(Z=heads|X=x_i, \theta^{old}) \left(\frac{h_i}{P_1} + \frac{t_i}{1-P_1}\right)$$

$$0 = \sum_{i \in z_1} P(Z=head|X=x_i, \theta^{old}) \frac{h_i}{P_1} + \sum_{i \in z_1} P(Z=head|X=x_i, \theta^{old}) \frac{t_i}{1-P_1}$$

$$\hat{p}_1 := \frac{\sum_{i \in Z_1} P(Z=\text{heads} \mid X=x_i, \theta^{old}) \cdot h_i}{\sum_{i \in Z_1} P(Z=\text{heads} \mid X=x_i, \theta^{old})(h_i - t_i)}$$

$$\hat{p}_2 = \frac{\sum_{i \in Z_2} P(Z=\text{tails} \mid X=x_i, \theta^{old}) h_i}{\sum_{i \in Z_2} P(Z=\text{tails} \mid X=x_i, \theta^{old})(h_i - t_i)}$$

$\theta^{new} = (\hat{\lambda}, \hat{p}_1, \hat{p}_2)$ is the global maximum of Q