The EM algorithm

1. Suppose that the random variable X has a mixture distribution; that is, the Y_i are independently distributed as

$$Y_i \sim \theta g(y) + (1-\theta)h(y), \quad i = 1, \ldots, n,$$

where g and h are known. An EM algorithm can be used to find the ML estimator of θ. Introduce $X_1, \ldots, X_n$, where X_i indicates from which distribution Y_i has been drawn, so $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = \theta$ and

$$Y_i \mid X_i = 1 \sim g(y),$$
$$Y_i \mid X_i = 0 \sim h(y).$$

   (a) Compute the complete data likelihood for this model.

   (b) Compute $\mathbb{E}_\theta(X_i \mid Y_i)$.

   (c) Use (a) and (b) to derive an EM updating formula for θ.