1. Consider a mixture of a $N(\mu, 1)$ distribution and a $N(0, 1)$ distribution.

$$f(y; \tau, \mu) = \tau(\frac{1}{\sqrt{2\pi}} e^{-(y-\mu)^2/2}) + (1-\tau)(\frac{1}{\sqrt{2\pi}} e^{-y^2/2})$$

where τ is the unknown mixing parameter and μ is the unknown mean of the first subpopulation.

Suppose we collect a sample $y_1, y_2, ..., y_n$ from $f(y; \tau, \mu)$. Write an EM algorithm to estimate the parameters τ and μ. Be clear about what is computed in the E-step and in the M-step.

A hint is to complete the data with the missing classification z. Let $z_i = 1$ if the ith case was drawn from $N(\mu, 1)$ and let $z_i = 0$ if y_i was drawn from $N(0, 1)$. The E-step will involve finding the expected values of these given the observed data and current parameter values. The complete data likelihood can be written as

$$L(\tau, \mu; \mathbf{y}, \mathbf{z}) = \prod_{i=1}^{n} (\frac{\tau}{\sqrt{2\pi}} e^{-(y_i-\mu)^2/2})^{z_i} (\frac{1-\tau}{\sqrt{2\pi}} e^{-y_i^2/2})^{(1-z_i)}$$

2. Generate a dataset with $\tau = 0.5$, $\mu = 1$ and $n = 100$ and run your EM algorithm in R. Show that it converged and give the parameter estimates.