

1. 에너지원을 화석연료, 석탄, 수력, 원자력으로 분류하여 매년 소비되는 에너지 유형을 조사한다. 관측값은  $X = (X_1, X_2, X_3, X_4)$ 로 나타낼 수 있으며  $p = \left(\frac{2(1+\theta)}{9}, \frac{3-2\theta}{9}, \frac{2}{9}, \frac{2}{9}\right)$ 를 모수로 가지는 다항분포를 따른다고 가정하자. 여기서  $0 \leq \theta \leq 1$ 이다. 에너지원 중 화석연료는 천연가스와 석유로 나뉘지만 이는 조사되지 않는다고 하자. 실제 관측값이  $(x_1, x_2, x_3, x_4) = (44, 19, 3, 6)$ 일 때 EM 알고리즘을 이용하여  $\theta$ 에 대한 최대가능도추정값을 구하시오.
2. 파라미터가  $\lambda_1$ 과  $\lambda_2$ 인 두 지수분포의 혼합분포

$$f(x) = \alpha \lambda_1 e^{-\lambda_1 x} + (1 - \alpha) \lambda_2 e^{-\lambda_2 x}$$

에  $\alpha$ 와  $\lambda_1, \lambda_2$ 의 최대 가능도 추정량을 구하는 EM 알고리즘을 유도하고 R 함수로 작성하시오.  $\alpha = 0.4$ ,  $\lambda_1 = 2$ ,  $\lambda_2 = 3$ 인 경우 혼합분포  $f(x)$ 에서  $N = 1000$ 개의 난수를 발생하고 EM 알고리즘에 의해 모수  $\alpha$ 와  $\lambda_1, \lambda_2$ 를 추정하여 참값과 비교하시오.