

Exercises for the 6th session

Exercise 1: Simulation experiment

Simulate 500 data points from a mixture of three normal distributions. Assume the four components have equal weights and standard deviation equal 1 and means -2, 0 and 2.

- Fit a Bayesian mixture model with two normal mixtures to the data
- Fit a Bayesian mixture model with three normal mixtures to the data
- Fit a Bayesian mixture model with four normal mixtures to the data

There is no simple way of telling what the “correct” number of mixture components is. One suggestion is to assume a maximum number of components H and the use a Dirichlet prior with parameters $1/H$. Try this approach in the case of four mixture components using $k = H = 4$.

Exercise 2: Galaxy speed data

This examples is concerned with the speed of galaxies. The data consist of the velocity of 82 galaxies in the corona borealis region. Download the data set `galaxy.dat` here

<http://people.math.aau.dk/~kkb/Undervisning/Bayes/data/>

Load the data into R using `galaxy = read.table(file="galaxy.dat",header=TRUE)`

Perform a Bayesian data analysis of the galaxy data assuming a mixture model with normal mixture components. Try using different numbers of mixture components, say 1 to 4.