An introduction for mixture modelling for unsupervised clustering

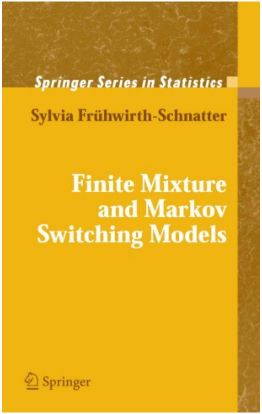
Mini-tutorial

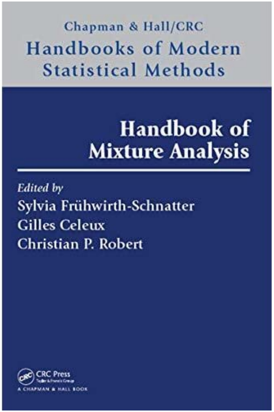
Nicole M White

Australian Centre for Health Services Innovation (AusHSI)
Queensland University of Technology

July 1, 2021

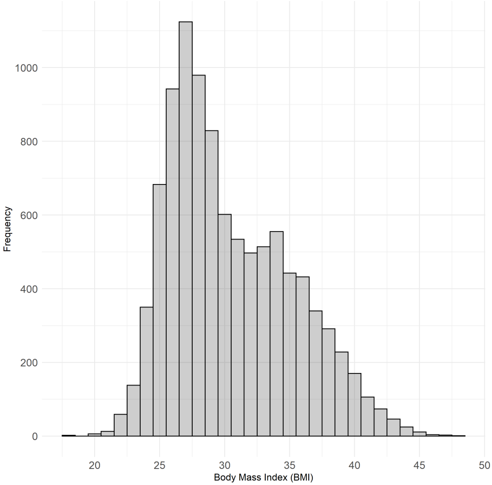
Further reading





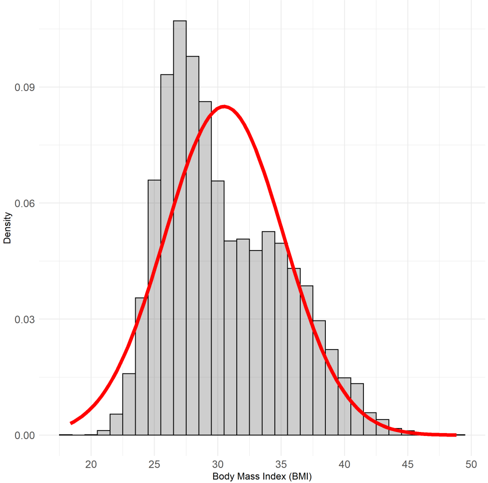
A motivating example

Distribution of body mass index (BMI) for 10,000 participants.



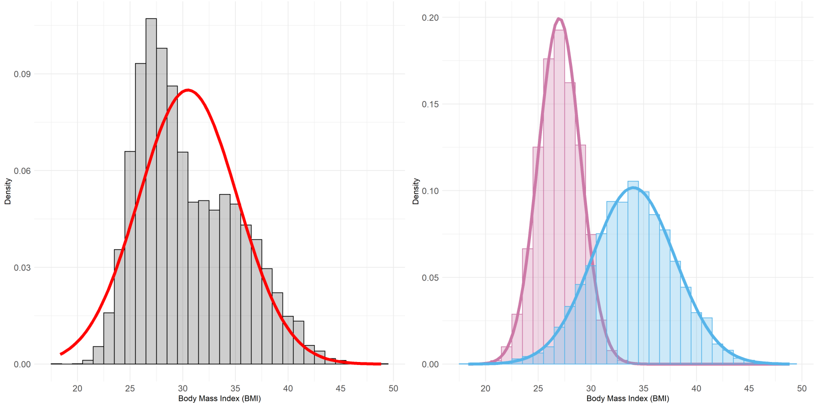
A motivating example

Distribution of body mass index (BMI) for 10,000 participants.



A motivating example

Distribution of body mass index (BMI) for 10,000 participants



Defining clustering

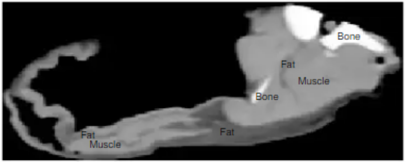
Unsupervised clustering $\leftrightarrow$ Identifying subgroups

Example clustering methods:

- Hierarchical clustering
- K-means
- Mixture models

[todo add figure to demonstrate cluster separation here ;e.g. k-means]

Examples of clustering using mixture models



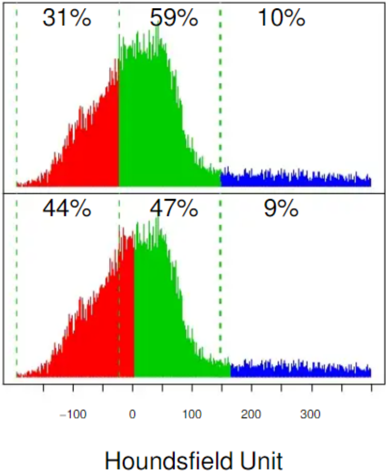


Figure 3: Experimental data, Sheep CT (Alston & Mengersen [ref])

Examples of clustering using mixture models

Spike sorting

- Show unsorted datasets from book chapter
- Show mixture solution with spikes in different colours

Mixture model ingredients

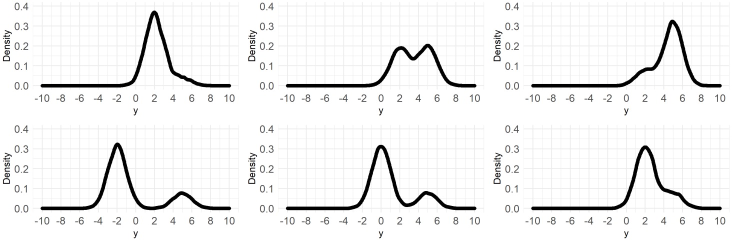
Data are drawn from a *convex combination of components*
For K groups/clusters:

$$\begin{aligned} p(y) &= \eta_1 p(y|\boldsymbol{\theta}_1) + \ldots + \eta_K p(y|\boldsymbol{\theta}_K) \\ &= \sum_{k=1}^{K} \eta_k p(y|\boldsymbol{\theta}_k) \end{aligned}$$

- $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$: Mixture weights; $\sum_{k=1}^{K} \eta_k = 1$
- $p(y|\boldsymbol{\theta}_k)$: k^{th} Mixture component; same parametric family

A simple 2-component mixture model

$$y_i \sim \eta_1 \mathcal{N}(\mu_1, 1) + \eta_2 \mathcal{N}(\mu_2, 1)$$



Mixture model examples

General formulation:

$$p(y_i) = \sum_{k=1}^{K} \eta_k p(y_i|\boldsymbol{\theta}_k)$$

- Latent class analysis (J items)

$$p(\mathrm{y}_i|\boldsymbol{\theta}_k) = \prod_{j=1}^{J} p(y_{ij}|\theta_{jk})$$

Mixture model examples

General formulation:

$$p(y_i) = \sum_{k=1}^{K} \eta_k p(y_i|\boldsymbol{\theta}_k)$$

- Hidden Markov models

[TODO]

Mixture model examples

General formulation:

$$p(y_i) = \sum_{k=1}^{K} \eta_k p(y_i|\boldsymbol{\theta}_k)$$

- Latent class regression: $\eta_k \to \eta_k(\mathrm{x}_i)$

$$\eta_k(\mathrm{x}_i) = \frac{exp\left(\mathrm{x}_i^T \beta_k\right)}{\sum_{l=1}^{K} exp\left(\mathrm{x}_i^T \beta_l\right)}$$
$$\beta_K = 0$$

Mixture model examples

Focus of mini-tutorial: cross-sectional, continuous data

- Finite mixture model
- Dirichlet Process mixture model
- Profile regression

Bayesian approaches to inference: Markov chain Monte Carlo (MCMC)

Setup

Assume a fixed number of components K
Each data point has a probability of belonging to each

Estimating a Finite Mixture Model

Aim is to learn $\boldsymbol{\eta}_{1,\ldots,K}$ and $\boldsymbol{\theta}_{1,\ldots,K}$
Both are conditional on k
Introduce a latent variable, $\boldsymbol{z}$

- One per observation: y_i, z_i
- Each z_i is discrete: $1, \ldots, K$ with $Pr(z_i = k) = \eta_k$ [check thesis]
- y_i belongs to cluster k iff $z_i = k$

Estimating a Finite Mixture Model

$$Pr(z_i = k|y_i, \cdot) = \frac{p(y_i|\boldsymbol{\theta}_k, z_i = k)Pr(z_i = k)}{\sum_{l=1}^{K} p(y_i|\boldsymbol{\theta}_l, z_i = l)Pr(z_i = l)}$$

Frame Title

- Label switching conumdrum
- Unswitching vs. xxx

Frame Title

- AIC, BIC
- variants of DIC

Thanks!

https://www.latexstudio.net/archives/4051.html