An introduction for mixture modelling for unsupervised clustering

Mini-tutorial

Nicole M White

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

June 28, 2021

1. Why mixture models/motivation
2. Finite mixture models
3. Dirichlet Process Mixture models
4. Profile regression
5. Model fitting and inference
6. References

1 Why mixture models/motivation
Definition

2 Finite mixture models

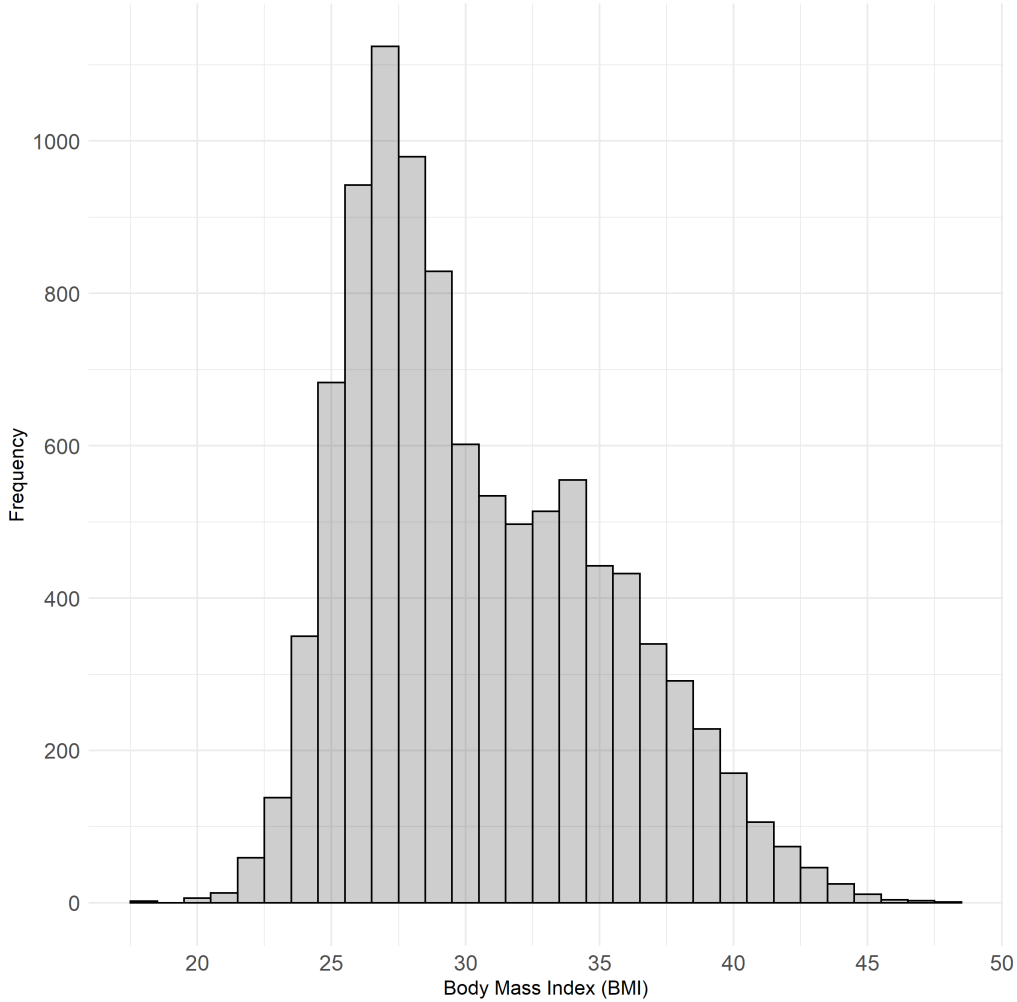
3 Dirichlet Process Mixture models

4 Profile regression

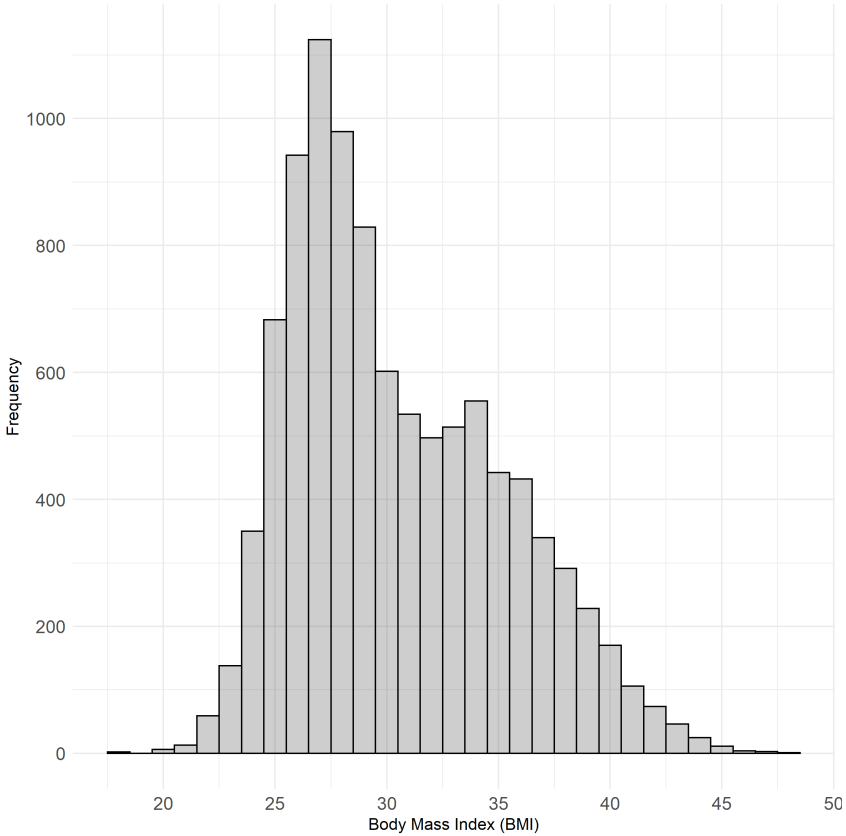
5 Model fitting and inference

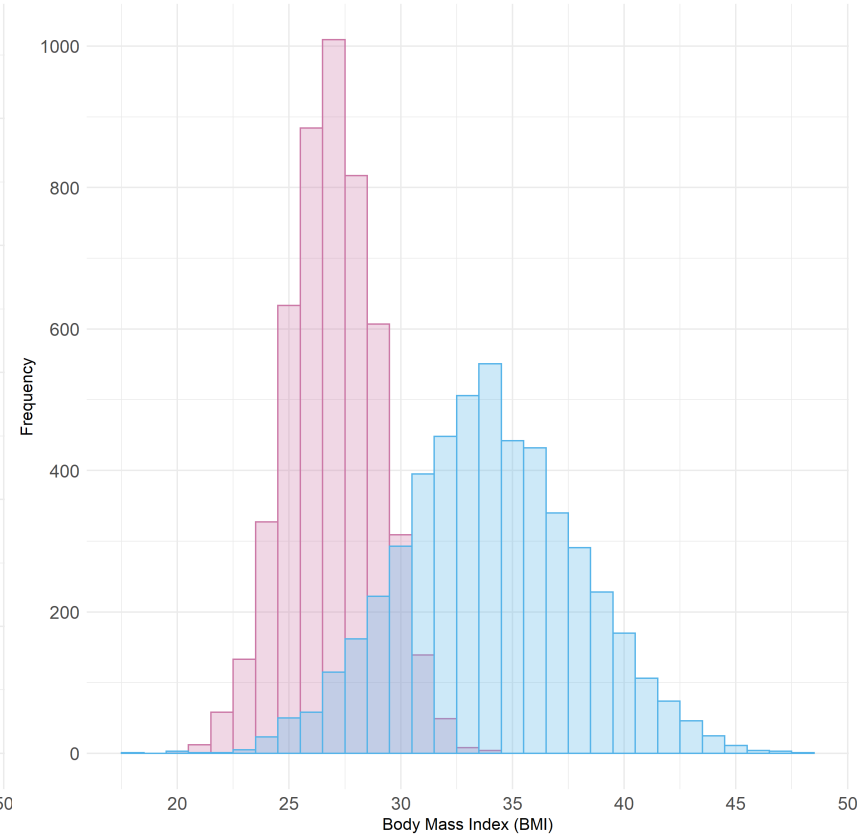
6 References

Motivation


1000
800
600
400
200
0
Frequency
20
25
30
35
40
45
50
Body Mass Index (BMI)

Motivation


Frequency
1000
800
600
400
200
0
20
25
30
35
40
45
50
Body Mass Index (BMI)


Frequency
1000
800
600
400
200
0
20
25
30
35
40
45
50
Body Mass Index (BMI)

Examples of clustering

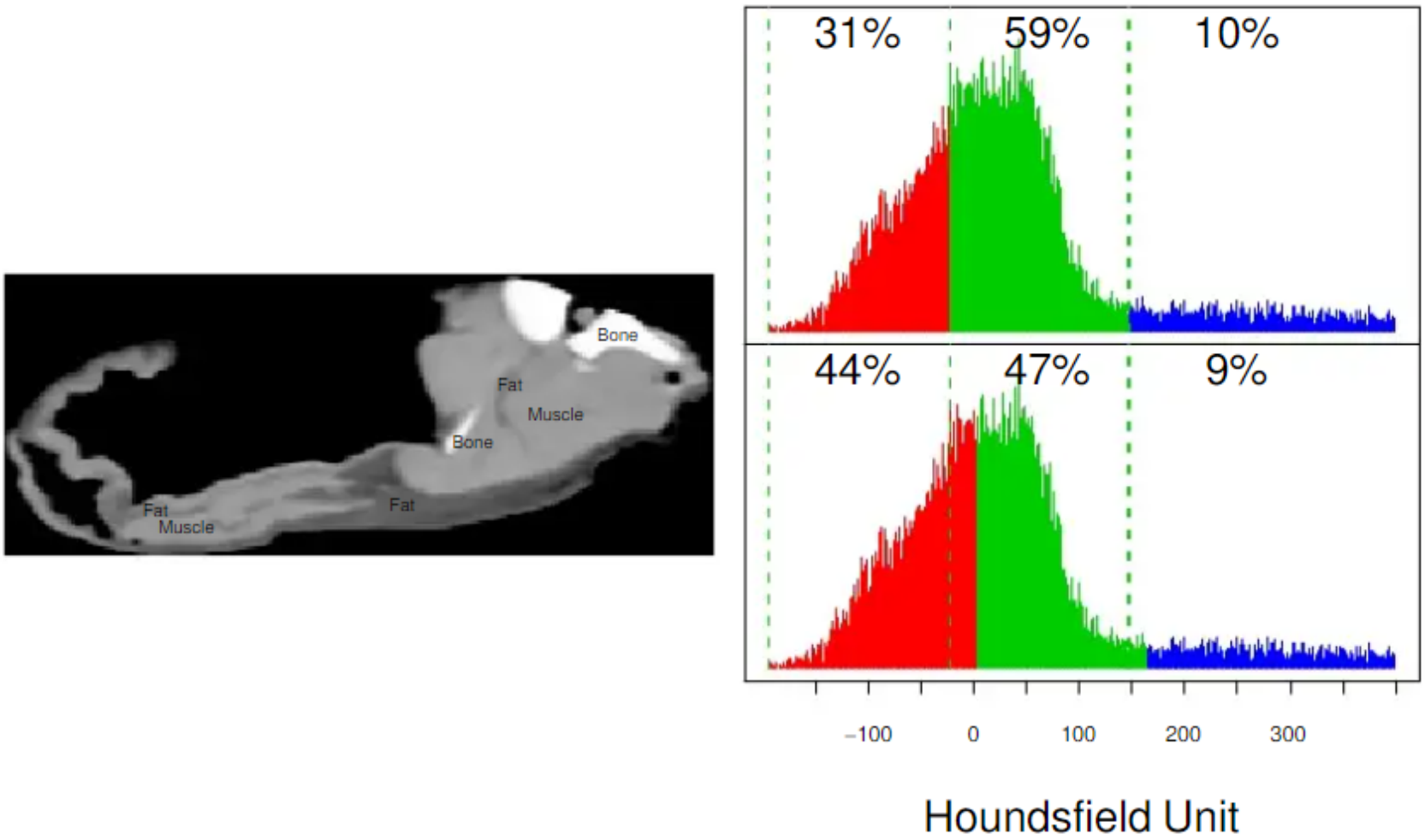


Figure 2: Experimental data, Sheep CT (Alston & Mengersen [ref])

Examples of clustering

Spike sorting

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

1 Why mixture models/motivation
Definition

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference

6 References

Building a mixture model

General formulation: 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

Building a mixture model

Figure here to show flexibility of motivating example - three panels, vary means and/or sds

Mixture model ingredients

Discuss how the basic formulation is flexible and can present different model types Examples

- Latent class analysis
- Latent class regression
- Hidden Markov models

Focus of this tutorial: cross-sectional data

- Finite mixture
- Dirichlet Process mixture (to infinity and beyond)
- Profile regression (joint clustering in x and y)

Bayesian approaches

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference

6 References

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

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

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models
- Stick breaking process
- Polya Urn
- Chinese Restaurant Process

4 Profile regression

5 Model fitting and inference

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models
- Stick breaking process
- Polya Urn
- Chinese Restaurant Process

4 Profile regression

5 Model fitting and inference

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models
- Stick breaking process
- Polya Urn
- Chinese Restaurant Process

4 Profile regression

5 Model fitting and inference

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models
- Stick breaking process
- Polya Urn
- Chinese Restaurant Process

4 Profile regression

5 Model fitting and inference

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference
- R implementation
- Inferring likely clusterings
- Choosing K

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference
- R implementation
- Inferring likely clusterings
- Choosing K

6 References

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference
- R implementation
- Inferring likely clusterings
- Choosing K

6 References

Frame Title

- Label switching conumdrum
- Unswitching vs. xxx

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference
- R implementation
- Inferring likely clusterings
- Choosing K

6 References

Frame Title

- AIC, BIC
- variants of DIC

1 Why mixture models/motivation

2 Finite mixture models

3 Dirichlet Process Mixture models

4 Profile regression

5 Model fitting and inference

6 References

Thanks!

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