Application of Latent Class with Random Effects Models to Longitudinal Data

Ken Beath
Macquarie University

Introduction

- Latent trajectory is a method of classifying subjects based on longitudinal data
- Original applications were for continuous and count data. An example is criminology, based on number of arrests in each time period can classify subjects as possibly
  - non criminal,
  - those that experiment with crime in their teens and cease
  - long term criminals

Application to Binary Data

- Croudace (2003) used standard latent class to classify subjects using longitudinal binary data, and this methodology has been used by a number of authors
- An assumption of latent class analysis is homogenous classes
- Heterogeneous classes is a more realistic assumption
- Example used is reanalysis of Croudace data using latent class with random effects.
- One problem is the assumption of normally distributed random effect.

Latent Class

- Binary outcomes.
- Categorical Latent variable.
- Probabilities for each outcome are determined by the class $c = 1, ..., C$
- Form of finite mixture distribution

Latent Class Model

- Probability for each subject is simply the Bernoulli probability, conditional on the class

$$\begin{aligned} \Pr\left(\mathrm{y_{i1}}, \mathrm{y_{i2}}, \ldots, \mathrm{y_{ik}} | \mathrm{c}\right) &= \prod_{j=1}^{k} \pi_{cj}^{y_{ij}} \left(1 - \pi_{cj}\right)^{1-y_{ij}} \\ \pi_{cj} &= \text{probability of outcome } j = 1 \\ & \text{for subject in class } c \end{aligned}$$

- Likelihood can be obtained by a weighted sum over classes and optimised using standard methods

Hybrid Models

- Latent class model is usually only an approximation, as it assumes no heterogeneity
- Problems resulting from ignoring heterogeneity
  - biased estimates
  - additional latent classes, especially as sample size increases
- Solution for diagnostic testing was to incorporate a random effect (latent variable) into the latent class model
- Model is effectively a mixture of Item Response Theory or Rasch models, depending on whether the random effect loading is allowed to vary by time
- Could be argued that these are always an appropriate model to fit instead of latent class

Latent Class with Random Effect Model

- Probability for each subject is now also conditional on the random effect $\lambda \sim N(0,1)$

$$
\begin{aligned}
\Pr\left(\mathrm{y}_{\mathrm{i1}}, \mathrm{y}_{\mathrm{i2}}, ..., \mathrm{y}_{\mathrm{ik}} | \mathrm{c}, \lambda\right) &= \prod_{j=1}^{k} \pi_{cj}^{y_{ij}} \left(1 - \pi_{cj}\right)^{1-y_{ij}} \\
\pi_{cj} &= \frac{\exp\left(\beta_{cj} + \ell_{cj}\lambda\right)}{1 + \exp\left(\beta_{cj} + \ell_{cj}\lambda\right)} \\
\ell_{cj} &= \text{loading, may be constant or} \\
& \quad \text{allowed to vary by } j \text{ and/or } c
\end{aligned}
$$

- Likelihood can be obtained by integrating over λ,a weighted sum over classes and optimised using standard methods

Nighttime Bladder Control Data

- 1946 British National Birth Cohort
- Singleton births to married parents
- 5362 subjects
- Data on bedwetting in the previous month recorded at 4, 6, 8, 9, 11 and 15 years. For this analysis recoded to presence or absence.
- Only subjects with complete data available used in analysis - 3272 subjects
- Analysis previously published by Croudace et al (2003) using standard latent class.

Methods

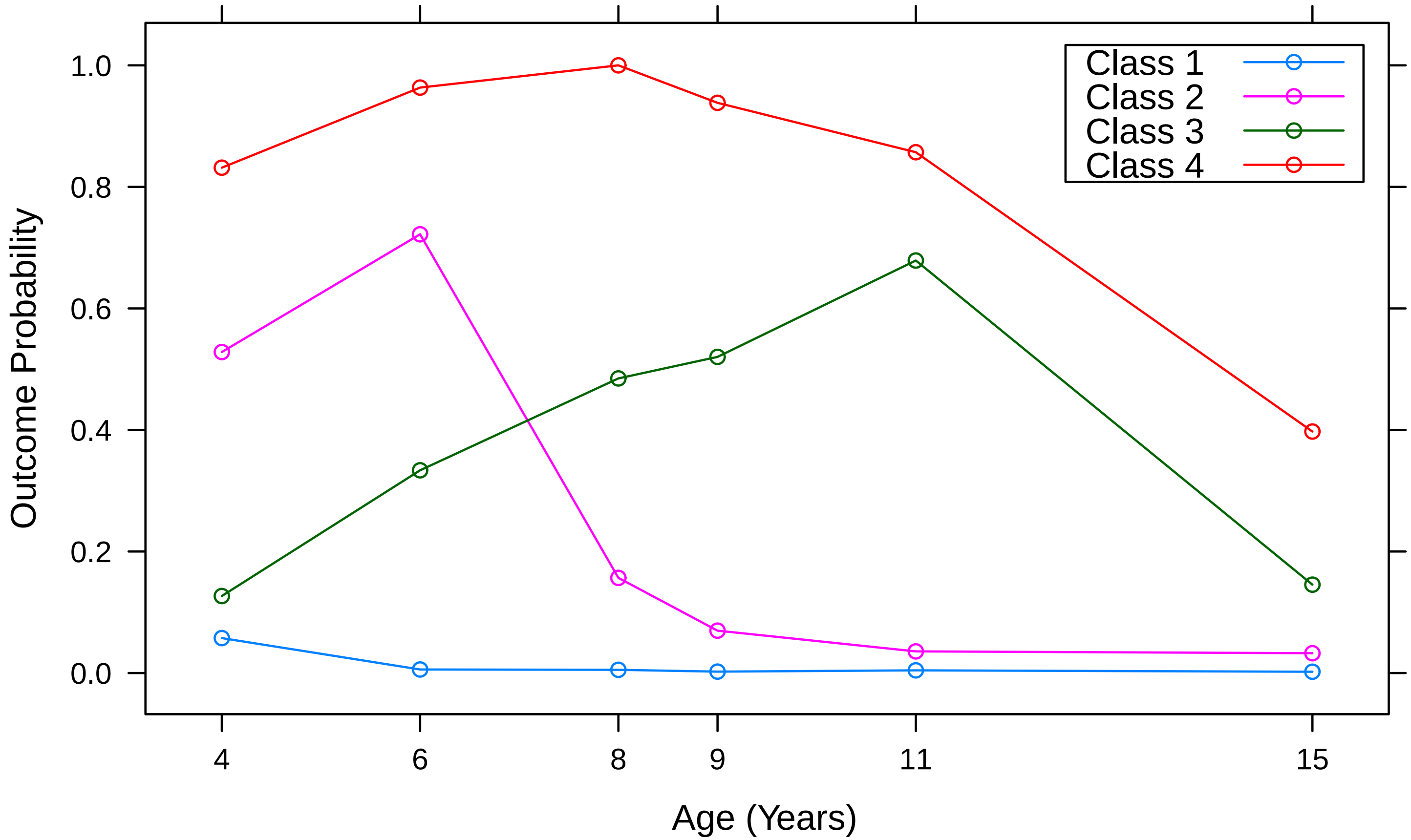
- Models fitted for Latent Class and with Random effect
- Allowing the random effect loading to vary by time period wasn't able to be fitted due to the large number of parameter estimates
- Fitted with commercial package `Latent GOLD Syntax Module` but the standard models can be fitted with `randomLCA` package for `R`
- Select number of classes based on BIC

BIC by Class and Model

		LC RE	
		Constant Loading	
Classes	LC	Common	by Class
1	9093.6	6710.5	6710.5
2	6794.1	**6568.7**	**6576.4**
3	6649.1	6585.9	6596.1
4	**6640.9**	6621.4	6638.3
5	6677.9	6669.8	6698.9

- Random Effects models better fit than standard Latent Class
- For standard Latent Class 4 classes is selected
- For random effects models 2 classes (normal and delayed) with no improvement by varying loading between classes

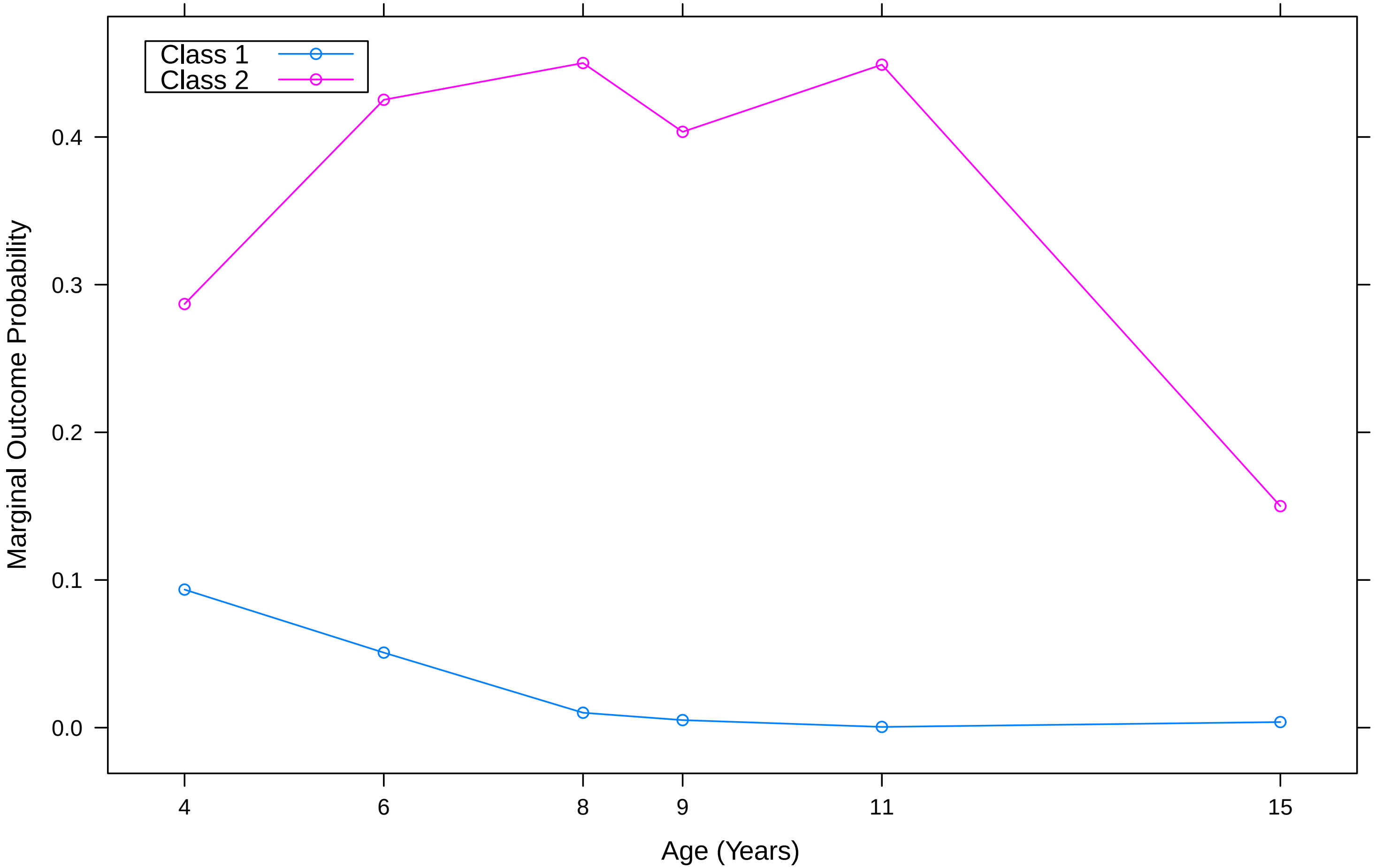
4 Class LC Model



Class probabilities

Class 1	Class 2	Class 3	Class 4
0.87	0.07	0.03	0.03

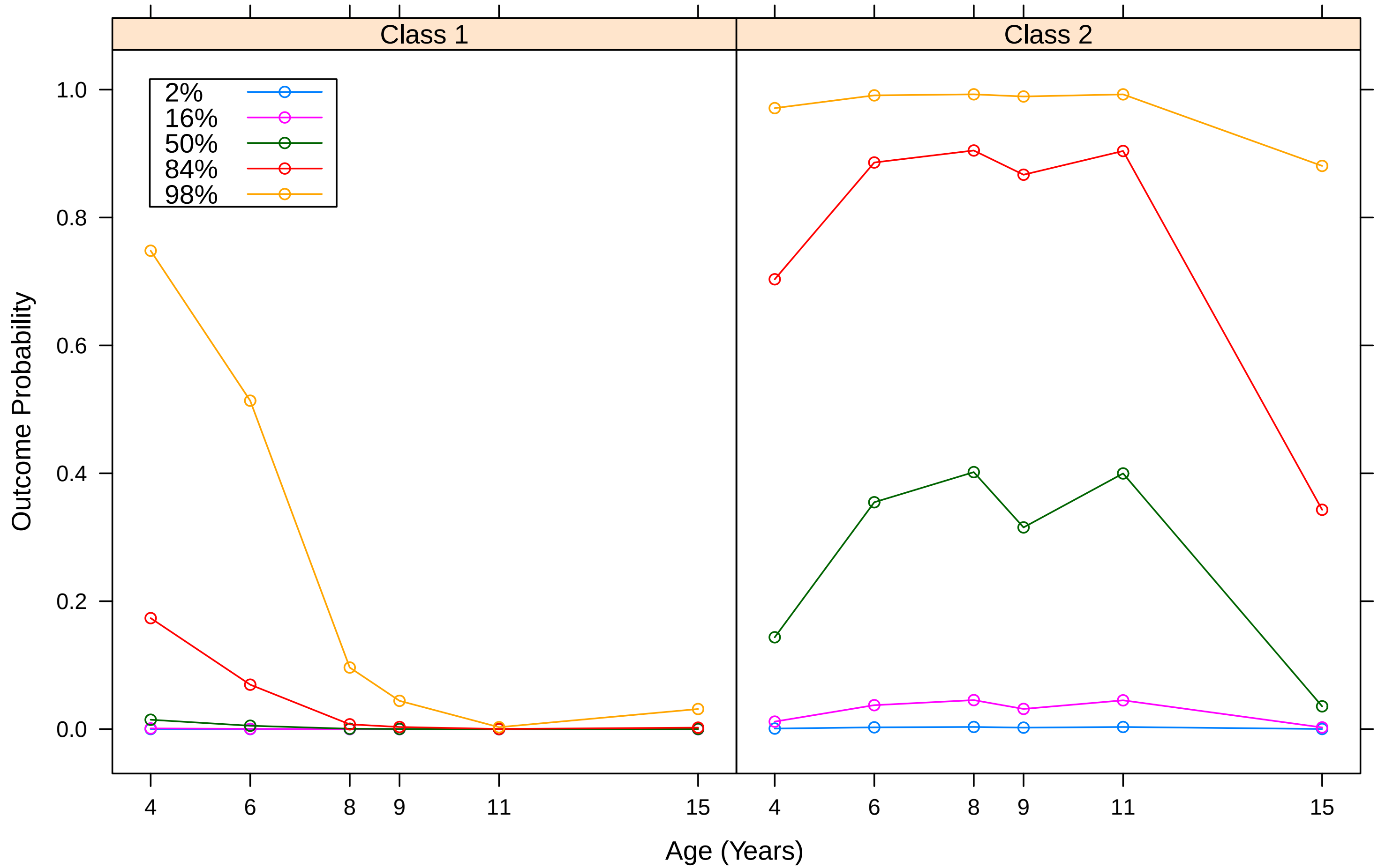
2 Class LCRE Model (Marginal Probabilities)



Class probabilities

Class 1	Class 2
0.89	0.11

2 Class LCRE Model (Conditional Probabilities)



Possibly some problems with assumption of constant loading for first year.

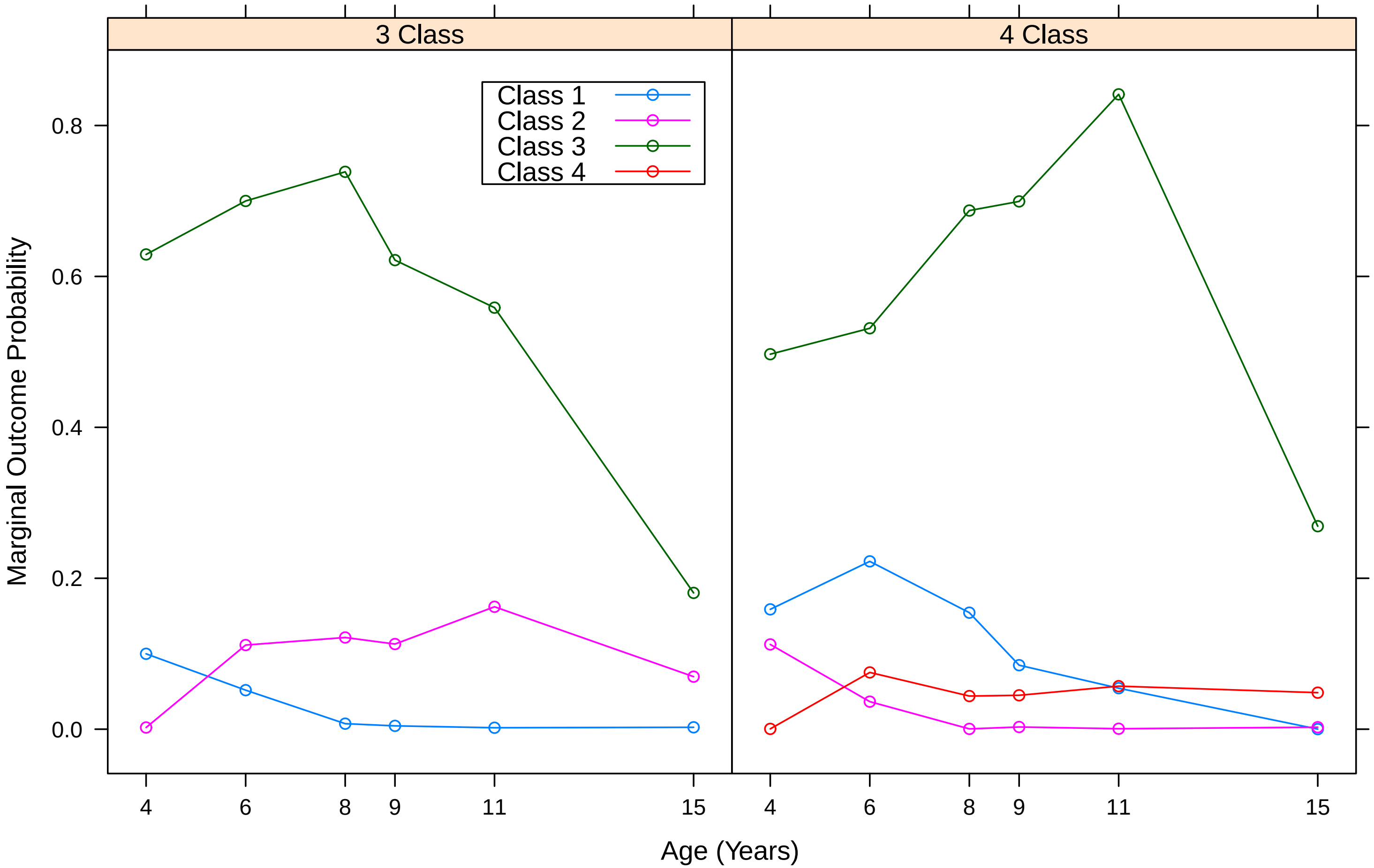
2 Class LCRE Model (Posterior Class Probabilities)

Pattern	Freq	4 Class				2 Class Random Effects	
		1	2	3	4	1	2
000000	2673	0.991	0.008	0.001	0.000	0.956	0.044
000001	7	0.818	0.103	0.079	0.000	0.701	0.299
000010	20	0.608	0.040	0.352	0.000	0.012	0.988
000100	11	0.551	0.140	0.309	0.000	0.212	0.788
001000	21	0.669	0.187	0.144	0.000	0.298	0.702
010000	72	0.216	0.762	0.022	0.000	0.837	0.163
100000	186	0.870	0.128	0.003	0.000	0.979	0.021
⋮							
111110	36	0.000	0.001	0.013	0.987	0.004	0.996
111111	31	0.000	0.000	0.003	0.997	0.005	0.995

Comments on Posterior Probability

- For random effects model a subject with no episodes has a 4.4% probability of being in the delayed class, compared to 0.9% for the standard latent class.
- Subjects with low incidence are unlikely to be detected because of only 1 month periods of observation. Epidemiology literature confirms that there are many subjects where bedwetting occurs less frequently than monthly.
- It is likely that posterior class probabilities are very dependent on distributional assumptions on random effects especially for extreme values of random effect.
- There will also be uncertainty in the posterior probabilities.
- Some of the unusual outcome patterns may be due to data entry errors.

3 and 4 Class LCRE Model (Marginal Probabilities)



- Appears that additional classes are simply modelling the normal class, and some of the delayed.

Robust Latent Class with Random Effect Model

- Allow for additional heterogeneity by allowing a mixture of random effects.
- Similar to "robust Bayesian" methods, which allow robustness and still use likelihood.
- Probability for each subject is now also conditional on the class of the random effect variance e where $\lambda_{ce} \sim N(0, \sigma^2_{ce})$ where $\sigma^2_{c1} = 1$ and σ^2_{c2} may be constrained to be equal across classes

$$
\begin{aligned}
\Pr(\mathrm{y_{i1}, y_{i2}, \ldots, y_{ik} | c, e, \lambda_e}) &= \prod_{j=1}^{k} \pi_{cej}^{y_{ij}} (1 - \pi_{cej})^{1-y_{ij}} \\
\pi_{cej} &= \frac{\exp(\beta_{cj} + \ell_{cj}\lambda_{ce})}{1 + \exp(\beta_{cj} + \ell_{cj}\lambda_{ce})}
\end{aligned}
$$

- Likelihood can be obtained as previously

BIC by Class and Model

Classes	LCRE	Robust LC RE
1	6710.5	6714.1
2	**6568.7**	**6560.7**
3	6585.9	6596.9
4	6621.4	6635.8
5	6669.8	6680.9

No improvement by allowing different error class variance.

Class probabilities for 2 Class Robust model

Class 1	Class 2
0.855	0.145

2 Class Robust LCRE Model (Posterior Class Probabilities)

Pattern	Freq	2 Class Random Effects		2 Class Robust Random Effects	
		1	2	1	2
000000	2673	0.956	0.044	0.901	0.099
000001	7	0.701	0.299	0.000	1.000
000010	20	0.012	0.988	0.024	0.976
000100	11	0.212	0.788	0.085	0.915
001000	21	0.298	0.702	0.174	0.826
010000	72	0.837	0.163	0.750	0.250
100000	186	0.979	0.021	0.983	0.017
⋮					
111110	36	0.004	0.996	0.869	0.131
111111	31	0.005	0.995	0.227	0.773

PROBLEM: Class determined almost completely by shape but not level.

Problem

- Normal distribution limits the extent of each class to region around mode.
- Allowing heavier tails doesn’t allow classes to be constrained, so may be fitting to other aspects of model misspecification.
- As flexibility of random effects distribution increases, identifiability is reduced.
- May be worthwhile where classes are better separated.

Conclusion

- Use of models incorporating random effects produces a better fit to the data.
- Use of standard latent class may result in detection of spurious classes.
- If class was known then random effects model would be used, so seems appropriate to use when class is not known.
- May be optimistic to hope for something better than the normal distribution for random effects. Aim is to divide population into classes, so provided this is robust then choice of distribution is irrelevant.