Multilevel Mixture Models

Tihomir Asparouhov
Muthen & Muthen

Bengt Muthen
UCLA

Motivation

- Multilevel Models

- Mixture Models

- Structural Equation Models

- Mplus 3, Mplus 4.1, Comprehensive Modeling Framework

Topics

- The Basic Two-level Mixture Model
- Two-level Latent Transition Analysis
- Between Level Class Variables
- Grade of Membership Models

The Basic Two-level Mixture Model

y^*_{ij} - observed dependent variables

x_{ij} - within level covariates

η_{ij} - continuous latent variables

C_{ij} - categorical latent variables

η_j - between latent variable (random effects)

x_j - between level covariates

The Basic Two-level Mixture Model

$$[y^*_{ij}|C_{ij} = c] = \nu_{cj} + \Lambda_{cj}\eta_{ij} + \varepsilon_{ij}$$

$$[\eta_{ij}|C_{ij} = c] = \mu_{cj} + \Gamma_{cj}\eta_{ij} + B_{cj}x_{ij} + \xi_{ij}$$

$$P(C_{ij} = c) = \frac{\exp(\alpha_{cj} + \beta_{cj}x_{ij})}{\sum_c \exp(\alpha_{cj} + \beta_{cj}x_{ij})}.$$

Within

Between

$$\eta_j = \mu + \Gamma\eta_j + Bx_j + \xi_j.$$

Two-level Latent Transition Analysis

Multiple latent class variable C_1,...,C_T.

Recursive System of Logit Models.

$$P(C_1 = 1) = \frac{exp(\alpha_{1j})}{exp(\alpha_{1j}) + 1}$$

$$P(C_2 = 1|C_1) = \frac{exp(\alpha_{2j} + \gamma I(C_1))}{exp(\alpha_{2j} + \gamma I(C_1)) + 1}$$

$I(C_1)$ - indicator variable for $C_1 = 1$

Two-level Latent Transition Analysis continued

- Random Effects: α_1 and α_2.

- Within Level Logistic Regression: $C_1 \rightarrow C_2$

- Between Level Linear Regression: $\alpha_1 \rightarrow \alpha_2$

Two-level Latent Transition Analysis Example

- Baltimore study of aggressive and disruptive behavior in the classroom. TOCA instruments on binary scale. The substantive research question comes from Nick Ialongo at Johns Hopkins U.

- 2 class model for first grade fall (C_1) and spring (C_2).

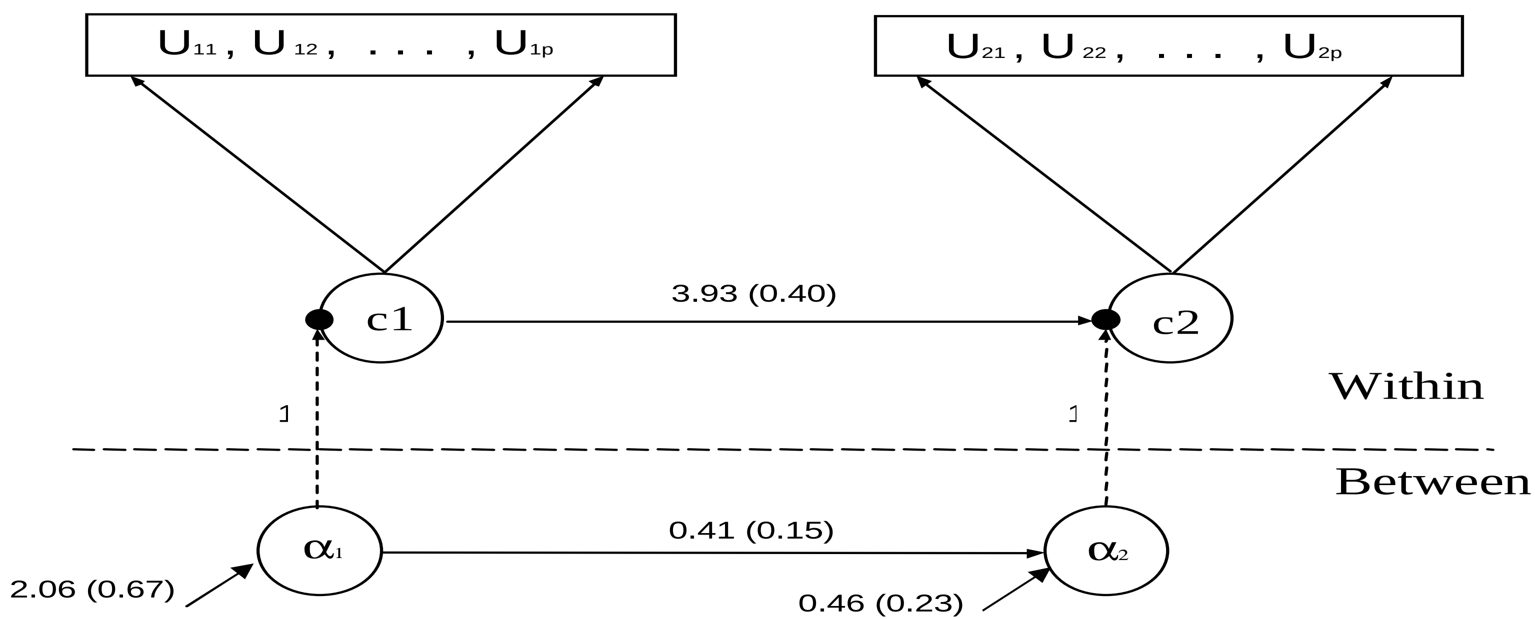
- Disruptive class is 46% in the fall and 52% in the spring.

- Transition probabilities: $P(C_2 = 1|C_1 = 2) = 18\%$, $P(C_2 = 2|C_1 = 1) = 7\%$

Two-level Latent Transition Analysis Example continued

class	$C_t = 1$	$C_t = 2$
Stubborn	0.92	0.36
Break Rules	0.96	0.29
Harm Others	0.73	0.03
Break Things	0.59	0.03
Yells at Others	0.82	0.18
Take Others' Property	0.78	0.07
Fights	0.73	0.08
Lies	0.81	0.10
Tease Classmates	0.90	0.24
Trouble Accepting Authority	0.78	0.12

Two-level Latent Transition Analysis Example continued



Two-level Latent Transition Analysis Example continued

- $\alpha_1 \rightarrow C_1$ $R^2 = 39\%$

- $\alpha_1, \alpha_2, C_1 \rightarrow C_2$ $R^2 = 65\%$
  - $\alpha_1 \rightarrow C_2$ 35%
  - residual $C_1 \rightarrow C_2$ 25% (total explained var 41%)
  - residual $\alpha_2 \rightarrow C_2$ 5% (additional explained var 24%)

Between Level Class Variables

- Between level heterogeneity

- $C_{ij} = C_j$

- How are between level class variables identified and interpreted?

- Is between level sample size important?

Between Level Class Variables - Simulation

$$Y_{ij} = \mu_{cj} + \beta_{cj} X_{ij} + \varepsilon_{ij}$$

$$Var(\mu_{cj}) = Var(\beta_{cj}) = v$$

	C	v	MSE of α_1
Model 1	within	0	0.31
Model 2	between	0	0.10
Model 3	between	0.1	0.26
Model 4	between	0.2	0.44

Level Class Variables Status

Is Population Heterogeneity a Between Level or a Within Level Phenomenon? Modeling options are

- Within Latent Class Variable: $Var(\alpha_{cj}) = 0$
- Between Latent Class Variable: $Var(\alpha_{cj}) = huge$
- Within-Between Latent Class Variable: $Var(\alpha_{cj}) > 0$

Level Class Variables Status - Simulation

Parameter	True Value	Between C	Within-Between C
μ_1	1	0.90(0.14)	0.90(0.13)
μ_2	0	0.13(0.13)	0.12(0.12)
β_1	0.2	0.25(0.07)	0.25(0.06)
β_2	0.8	0.81(0.07)	0.81(0.06)
θ	1	1.01(0.04)	1.01(0.04)
v	0.2	0.22(0.08)	0.22(0.08)
LL		-1517.6	-1517.6

$$ICC(C) = \frac{Var(\alpha_1)}{Var(\alpha_1)+\pi^2/3} = 0.995$$

The Grade of Membership Model

- Extension of LCA - Latent class variable for each item

- Partial class membership

- Individuals in transitional state: in between classes.

- Some measurements indicate one class, other measurements indicate another class

The Grade of Membership Model continued

- LCA

$$P(Y_{ij} = 1 | C_j = c) = \Phi(\tau_{ic})$$

$$P(C_j = 1) = \frac{exp(\alpha_1)}{1 + exp(\alpha_1)}$$

- GoM

$$P(Y_{ij} = 1 | C_{ij} = c) = \Phi(\tau_{ic})$$

$$P(C_{ij} = 1) = \frac{exp(\alpha_{1j})}{1 + exp(\alpha_{1j})}$$

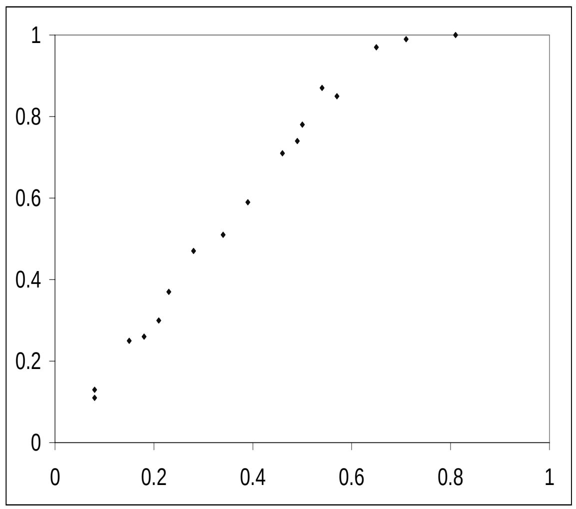
The Grade of Membership Model Example

Antisocial Behavior (ASB) data from the National Longitudinal Survey of Youth (NLSY). 17 binary items: 8 property offense items, 5 personal offense items and 4 drug offense items.

Model	2 class	3 class
LCA	-42625.7 (35)	-41713.1 (53)
GoM	-42159.1 (36)	-41554.6 (55)

The Grade of Membership Model Example

17 Items Probability Profiles for the Offense Prone Class for GoM v.s. LCA models. Correlation 99%.



Combining LCA, IRT and GoM Models: GoM-FMA

- FMA model

$$P(Y_{ij} = 1 | C_j = c) = \Phi(\tau_{ic} + \lambda_{ic}\eta_j)$$

$$P(C_j = 1) = \frac{exp(\alpha_1)}{1 + exp(\alpha_1)}$$

- GoM-FMA model

$$P(Y_{ij} = 1 | C_{ij} = c) = \Phi(\tau_{ic} + \lambda_{ic}\eta_j)$$

$$P(C_{ij} = 1) = \frac{exp(\alpha_{1j})}{1 + exp(\alpha_{1j})}$$

GoM-FMA Example

UCLA clinical sample of 425 males with ADHD diagnosis. The data consists of 9 inattentiveness items and 9 hyperactivity items, all dichotomous.

2 Class Model	Log-Likelihood	Number of Parameters
LCA	-3650.0	37
FMA	-3502.4	56
FMA-GoM	-3501.7	57

ICC(C)=86%

Multilevel Mixture Models with Survey Data

- Sampling Weights on the Within and the Between Level
- Stratification
- Multistage Sampling: PSU and SSU

Multilevel Mixture Models with Non-Normal Data

- Binary, Polytomous, Nominal, Poisson, Censored, Two-Part

- Survival Analysis: Cox regression, Finite Mixtures of Frailty Models

Technical Aspects

- Random Starts
  - Starting values from simple models
  - Replicating the best solution

- Numerical Integration
  - Adaptive and Non-Adaptive Integration
  - Gauss-Hermite, Regular, Monte-Carlo Integration
  - Parametrization, Cholesky option

Conclusions

- Mplus 4.1

- Mplus 4.2

  - Multiple Class Variables for Two-Level Models

  - Between Level Class Variables

  - Multiple Within, Between and Within-Between Class Variables