Confirmatory mixture models in a developmental context

Karen Draney & Mark Wilson
UC Berkeley

Talk Outline

- "Exploratory" versus "Confirmatory" models
- Theory and purpose behind the saltus model
- Specifics of the saltus model
- Example of a saltus analysis: Measuring deductive reasoning

One way to look at mixture models

- Models with more exploratory features
  - Rost's mixed Rasch model (other models mentioned are formally submodels of this)
- Models with more confirmatory features
  - Mislevy and Verhelst's LLTM-based model
  - The saltus model (Wilson, Draney)

The saltus model

- More confirmatory in nature
- Originally developed to investigate developmental theories
  - E.g. Piagetian/neo-Piagetian
- Most useful in strongly theoretical contexts

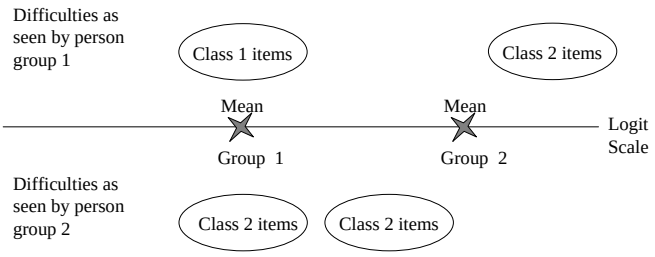
Model structure

- H groups of persons
  - Person group membership is latent
  - Groups are ordered from lower to higher (developmentally)
- H classes of items
  - Item group membership known *a priori*
  - Items represent the first group at which a person has all of the skills to correctly answer the item
  - Not required that there be the same number of groups and classes, but it is commonly the case

Model structure, cont'd

- Items within a class can vary in overall difficulty
- Persons within a group can vary in overall proficiency
- Classes of items vary in relative difficulty for different groups (often becoming relatively easier for higher groups)

An illustration



The formal model

- The probability of person n with proficiency θ_n responding in category j to item i (with difficulty vector β_i), given that the person is a member of group h, is given by:

$$P\left(X_{ni} = x_{ni} \mid \theta_n, \phi_{nh} = 1, \beta_i, \tau\right) = \frac{\exp \sum_{j=0}^{x} \left(\theta_n - \beta_{ij} + \tau_{hk}\right)}{\sum_{t=0}^{J_i} \exp \sum_{j=0}^{t} \left(\theta_n - \beta_{ij} + \tau_{hk}\right)}.$$

- where τ_{hk} is the effect of being in group h on items in class k

Constraints on the model

- Item difficulties are centered on zero.
- Difficulty parameters of class 1 items are the same for all person groups (i.e. $\tau_{h1} = 0$ for all h).
- Shifts in difficulty parameters for person groups 2,…,k are all defined relative to the difficulties as seen by person group 1 (i.e. $\tau_{1k} = 0$ for all k).

A brief history

- Dichotomous saltus model first developed by Wilson (1984)
- Polytomous saltus model developed by Draney (1996)
- Most prominent applications
  - Siegler's balance scale data (1984)
  - Noelting's juice mixtures data (1996)
- Fixed saltus model applied by De Boeck et al (2000)

Our example application

- Data collected by Spiel, Glück, & Gössler (2001)
- Instrument measuring deductive reasoning
- Contains items of the following types (types crossed to produce 24 items total)
  - Modus Ponens, Modus Tonens, Negation of Antecedent, Affirmation of Consequent
  - Concrete, Abstract, Counterfactual
  - With and without negation
- Possible responses: Yes, Perhaps, No

Structure of items, part 1

General Form	*Example*	*Types of Inference*
A, therefore B.	Klaus is ill. Correct inference: Klaus is lying in his bed.	Affirmation of the Antecedent = Modus Ponens (MP)
Not A, therefore B or not B.	Klaus is not ill. Correct inference: Perhaps Klaus is lying in his bed, perhaps not.	Negation of the Antecedent (NA)
B, therefore A or not A.	Klaus lies in his bed. Correct inference: Perhaps Klaus is ill, perhaps not.	Affirmation of the Consequent (AC)
Not B, therefore not A.	Klaus does not lie in his bed. Correct inference: Klaus is not ill.	Negation of the Consequent = Modus Tollens (MT)

From Spiel, C. & Glück, J. (in press). A computer based test of competence profile and competence level in deductive reasoniing E. Klieme & D. Leutner (Eds), Assessment of Competencies in educational contexts: State of the art and future prospects. Göttingen: Hogrefe.

Structure of items, part 2

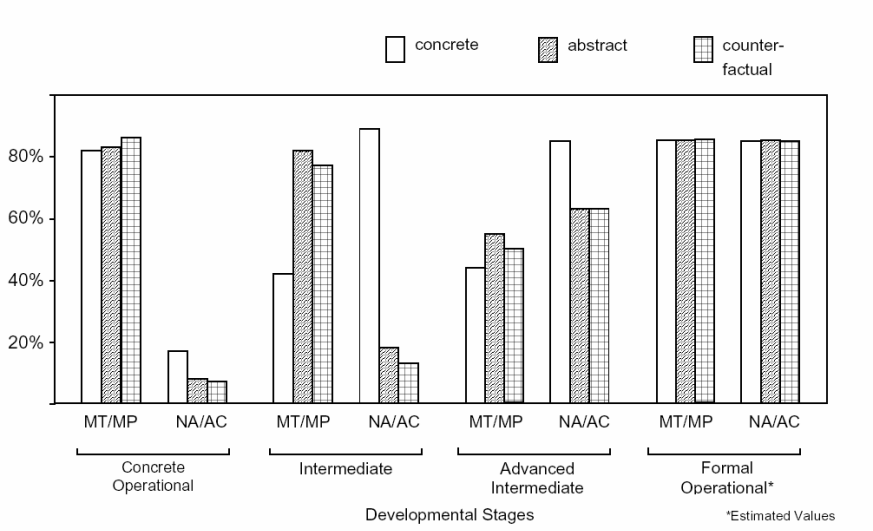
	Concrete	Abstract	Counterfactual
Without negation of antecedent	If the sun shines, Tina wears a red skirt.	If Y belongs to group F, Y has attitude g.	If it is evening, the sun rises.
With negation of antecedent	If the sun does not shine, Peter wears blue trousers.	If X does not belong to group B, X has attitude c.	If it is not evening, the sun sets.

From Spiel, C. & Glück, J. (in press). A computer based test of competence profile and competence level in deductive reasoniing E. Klieme & D. Leutner (Eds), Assessment of Competencies in educational contexts: State of the art and future prospects. Göttingen: Hogrefe.

Results of prior mixture analyses

- 4 latent groups:
  - Concrete:
    - Tend to correctly solve MP and MT items, no others
  - 2 intermediate:
    - Tend to correctly solve concrete-level Fallacy (i.e. NA and AC) items, but have difficulty with concrete MP and MT items
    - Pattern is reversed for Abstract and Counterfactual items
    - Advanced intermediate similar, but more likely to correctly solve items in general
  - Formal
    - Tend to correctly solve most items

Graphic from prior analyses



From Spiel, C. & Glück, J. (in press). A computer based test of competence profile and competence level in deductive reasoniing E. Klieme & D. Leutner (Eds), Assessment of Competencies in educational contexts: State of the art and future prospects. Göttingen: Hogrefe.

We fit two saltus models

- Model 1: Two-level
  - Concrete person group represented by MP/MT items
  - Formal person group represented by Fallacy items
- Model 2: Three-level
  - Concrete person group represented by MP/MT items
  - Intermediate person group represented by Concrete Fallacy items
  - Formal person group represented by Abstract & Counterfactual Fallacy items

Results from Model 1

- Item difficulties
  - Ranged from -3.15 to -0.85 for Class 1
  - Ranged from 0.17 to 3.52 for Class 2
- τ parameter = 4.28 (0.05)
- Means (standard deviations)
  - Class 1: -0.39 (0.41)
  - Class 2: -1.62 (1.05)
- Proportions in class
  - Class 1: 0.43
  - Class 2: 0.57

Model 1, continued

- No one who scored over 5 on the fallacy items was classified into class 2; most scored 0, 1, or 2.
- No one who scored under 6 on the MP/MT items was classified into class 1; most scored 9 or above.
- The persons who scored low on both sets of items were classified into class 2. This helps to account for higher variance and lower mean of this class.

Example persons

Response string	P(class 1)	P(class 2)
111101111111000000000000	1.00	0.00
111111101100111101111110	0.00	1.00
000000000000111111111000	0.00	1.00
000010001100100000000000	0.01	0.99

Results from Model 2

- Item difficulties similar to Model 1

Saltus parameters (standard errors)

0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
0.000 (0.000)	4.878 (0.819)	7.842 (0.264)
0.000 (0.000)	3.312 (5.645)	7.199 (0.694)

	CLASS 1	CLASS 2	CLASS 3
MEANS	-0.285	-1.645	-2.749
SDs	0.775	0.833	1.614
PROPORTIONS	0.307	0.509	0.184

Model 2, continued

- Interpretation here more complex
- Persons who scored low on both class 2 and class 3 items and high on class 1 items were in class 1
- Persons who scored high on class 3 items (regardless of other scores) were in class 3
- Other persons were in class 2, a mixed class

Summary

- Large changes in average difficulty for groups of items, based on class membership
- Not simple developmental increases in proficiency -- the class 1 items actually become harder (this is typical of developmental studies)
- An LLTM-based saltus model would be helpful in such cases.