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

Difficulties as seen by person group 1

Class 1 items

Mean

Group 1

Mean

Group 2

Logit Scale

Difficulties as seen by person group 2

Class 2 items

Class 2 items

The formal model

• The probability of person n with proficiency $\theta_n$ responding in category j to item i (with difficulty vector $\beta_i$), given that the person is a member of group h, is given by:

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

• where $\tau_{hk}$ 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)
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

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

Structure of items, part 2

<table>
<thead>
<tr>
<th></th>
<th>Concrete</th>
<th>Abstract</th>
<th>Counterfactual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without negation of antecedent</td>
<td>If the sun shines, Tina wears a red skirt.</td>
<td>If Y belongs to group F, Y has attitude g.</td>
<td>If it is evening, the sun rises.</td>
</tr>
<tr>
<td>With negation of antecedent</td>
<td>If the sun does not shine, Peter wears blue trousers.</td>
<td>If X does not belong to group B, X has attitude e.</td>
<td>If it is not evening, the sun sets.</td>
</tr>
</tbody>
</table>

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

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
• \( \tau \) 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

<table>
<thead>
<tr>
<th>Response string</th>
<th>P(class 1)</th>
<th>P(class 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11110111111110000000000000</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>111111101100111101111110</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>000000000000000111111111100</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>000010001100100000000000000</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Results from Model 2

- Item difficulties similar to Model 1

Saltus parameters (standard errors)

<table>
<thead>
<tr>
<th></th>
<th>CLASS 1</th>
<th>CLASS 2</th>
<th>CLASS 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000 ( 0.000)</td>
<td>-0.285</td>
<td>-1.645</td>
<td>-2.749</td>
</tr>
<tr>
<td>0.000 ( 0.000)</td>
<td>4.878 ( 0.819)</td>
<td>7.842 (0.264)</td>
<td></td>
</tr>
<tr>
<td>0.000 ( 0.000)</td>
<td>3.312 ( 5.645)</td>
<td>7.199 (0.694)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CLASS 1</th>
<th>CLASS 2</th>
<th>CLASS 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANS</td>
<td>-0.285</td>
<td>-1.645</td>
<td>-2.749</td>
</tr>
<tr>
<td>SDs</td>
<td>0.775</td>
<td>0.833</td>
<td>1.614</td>
</tr>
<tr>
<td>PROPORTIONS</td>
<td>0.307</td>
<td>0.509</td>
<td>0.184</td>
</tr>
</tbody>
</table>
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.