Advances in Models for Multivariate Nominal or Ordinal Variables with Latent Variables

IMPS
July 2010
Athens Georgia

Carolyn J. Anderson
Department of Educational Psychology

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Models for Nominal & Ordinal Variables

- Example data set: Espelage et al.
- Review of existing approaches for analyzing such data.
- Log-multiplicative association models.
  - Underlying models that lead to LMA
  - Conditional Specification.
- The State of the Art:
  - Multiple correlated latent variables.
  - Restrictions on scale values for response options and location parameters.
  - Covariates.
  - Estimation.
- Areas for future work.
- Time permitting, a nominal example.


The data consist of responses of students from a midwestern middle school to items of the Illinois Victimization Scale (Espelage & Holt, 2001).


The data consist of responses of students from a midwestern middle school to items of the Illinois Victimization Scale (Espelage & Holt, 2001).

**Bully Sub-scale Items:**
- I upset other students for the fun of it.
- I helped harass other students.
- I teased other students.


The data consist of responses of students from a midwestern middle school to items of the Illinois Victimization Scale (Espelage & Holt, 2001).

Bully Sub-scale Items:
- I upset other students for the fun of it.
- I helped harass other students.
- I teased other students.

Fight Sub-scale Items:
- I got in a physical fight.
- I threatened to hurt or hit another student.
- I hit back when someone hits me first.


The data consist of responses of students from a midwestern middle school to items of the Illinois Victimization Scale (Espelage & Holt, 2001).

**Bully Sub-scale Items:**
- I upset other students for the fun of it.
- I helped harass other students.
- I teased other students.

**Fight Sub-scale Items:**
- I got in a physical fight.
- I threatened to hurt or hit another student.
- I hit back when someone hits me first.

**Response scale:**
Never, 1 or 2 times, 3 or 4 times, 5 or 6 times, and 7 or more times.


The data consist of responses of students from a midwestern middle school to items of the Illinois Victimization Scale (Espelage & Holt, 2001).

**Bully Sub-scale Items:**
- I upset other students for the fun of it.
- I helped harass other students.
- I teased other students.

**Fight Sub-scale Items:**
- I got in a physical fight.
- I threatened to hurt or hit another student.
- I hit back when someone hits me first.

**Response scale:**
Never, 1 or 2 times, 3 or 4 times, 5 or 6 times, and 7 or more times.

**The role Gender?** Do girls tend to be more verbal bullies and boys more physical? … *The findings are mixed*…
Review of Major Existing Approaches to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
Review of Major Existing Approaches to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
  - Need to know the order of the response options.
  - Does not allow for alternative scoring for different latent variables.
Review of Major Existing Approaches
to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
  - Need to know the order of the response options.
  - Does not allow for alternative scoring for different latent variables.

- Item response theory models for polytomous items
Review of Major Existing Approaches to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
  - Need to know the order of the response options.
  - Does not allow for alternative scoring for different latent variables.

- Item response theory models for polytomous items
  - Multiple latent variables is a problem for standard estimation algorithms (i.e., numerical integration).
Review of Major Existing Approaches to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
  - Need to know the order of the response options.
  - Does not allow for alternative scoring for different latent variables.

- Item response theory models for polytomous items
  - Multiple latent variables is a problem for standard estimation algorithms (i.e., numerical integration).

- Factor analysis of discrete data (Bartholomew, Steele, Moustaki & Galbraith, 2008)
Review of Major Existing Approaches

to Latent Variable Modeling of Discrete Response Data:

- Quantify the data and then use factor analysis (or SEM) for continuous data.
  - Need to know the order of the response options.
  - Does not allow for alternative scoring for different latent variables.

- Item response theory models for polytomous items
  - Multiple latent variables is a problem for standard estimation algorithms (i.e., numerical integration).

- Factor analysis of discrete data (Bartholomew, Steele, Moustaki & Galbraith, 2008)
  - Lack of available of software and flexibility of implementation.
  - Methods and programs for nominal data are sorely lacking and “...work on ordinal categorical variables is nearer the research frontier and is consequently more incomplete, and in some sense, more difficult than other methods.” (p. 243)
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
- Generalization of Goodman’s (1979, 1986) $RC(M)$ association model for two-way tables to multi-way tables, i.e.,

$$
\log(P(y_i = j, y_k = \ell_k)) = \lambda + \lambda_{ij}^R + \lambda_{k\ell}^C + \sum_m \phi_m \nu_{ij}^R \nu_{k\ell}^C m
$$
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
- Generalization of Goodman’s (1979, 1986) $RC(M)$ association model for two-way tables to multi-way tables, i.e.,

$$\log(P(y_i = j_i, y_k = \ell_k)) = \lambda + \lambda_{ij}^R + \lambda_{k\ell}^C + \sum_m \phi_m \nu_{ijm}^R \nu_{k\ell m}^C$$

- When equally spaced scores are input for the $\nu$’s (and $M = 1$), then the model is known as a the uniform association model.
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
- Generalization of Goodman’s (1979, 1986) $RC(M)$ association model for two-way tables to multi-way tables, i.e.,

$$
\log(P(y_i = j, y_k = \ell)) = \lambda + \lambda_{ij}^R + \lambda_{k\ell}^C + \sum_m \phi_m \nu_{ij}^R \nu_{k\ell}^C
$$

- When equally spaced scores are input for the $\nu$’s (and $M = 1$), then the model is known as the uniform association model.
- Takane (1987): Ideal point discriminant analysis without a centroid restriction on the columns (criterion groups) is equivalent to the $RC$ association model.
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
- Generalization of Goodman’s (1979, 1986) $RC(M)$ association model for two-way tables to multi-way tables, i.e.,

$$\log(P(y_i = j, y_k = \ell)) = \lambda + \lambda_{ij}^R + \lambda_{k\ell}^C + \sum_m \phi_m \nu_{ij}^R \nu_{k\ell}^C$$

- When equally spaced scores are input for the $\nu$’s (and $M = 1$), then the model is known as the uniform association model.

- Takane (1987): Ideal point discriminant analysis without a centroid restriction on the columns (criterion groups) is equivalent to the $RC$ association model.

- Andersen (1995): Rasch model for polytomous items where an item’s response options are the rows and the columns are (categorical) estimates of ability/latent trait.
Log-multiplicative Association Models

- Structured Poisson regression model with 2-way interactions.
- Generalization of Goodman’s (1979, 1986) $RC(M)$ association model for two-way tables to multi-way tables, i.e.,

$$\log(P(y_i = j_i, y_k = \ell_k)) = \lambda + \lambda_{ij}^R + \lambda_{k\ell_k}^C + \sum_m \phi_m \nu_{ijm}^R \nu_{k\ell_k}^C$$

- When equally spaced scores are input for the $\nu$’s (and $M = 1$), then the model is known as a the uniform association model.

- Takane (1987): Ideal point discriminant analysis without a centroid restriction on the columns (criterion groups) is equivalent to the $RC$ association model.

- Andersen (1995): Rasch model for polytomous items where an item’s response options are the rows and the columns are (categorical) estimates of ability/latent trait.

- The general log-multiplicative association (LMA) model

$$\log(P(y)) = \lambda + \sum_i \lambda_{ij} + \sum_i \sum_{k>i} \sum_m \sum_{m'\geq m} \sigma_{mm'} \nu_{ijm} \nu_{k\ell m'}$$
Underlying Models that Imply a LMA Model

that I know of...

- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).
Underlying Models that Imply a LMA Model

- Are implied by underlying **multivariate normal distribution** (Goodman, 1979; Becker, 1989; others).

- Can be derived from a **distance based model** (de Rooij & Heiser, 2005), and a generalization of Newton’s Law of Gravity (de Rooij, 2008).
Underlying Models that Imply a LMA Model

that I know of...

- Are implied by underlying **multivariate normal distribution** (Goodman, 1979; Becker, 1989; others).

- Can be derived from a **distance based model** (de Rooij & Heiser, 2005), and a generalization of Newton’s Law of Gravity (de Rooij, 2008).

- For dichotomous items, derived via the **Dutch Identity** (Holland, 1993).
Underlying Models that Imply a LMA Model

that I know of...

- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).

- Can be derived from a distance based model (de Rooij & Heiser, 2005), and a generalization of Newton’s Law of Gravity (de Rooij, 2008).

- For dichotomous items, derived via the Dutch Identity (Holland, 1993).

- A generalization of the Dutch Identity to Rasch models for polytomous items, multidimensional traits and covariates. (Li, 2010).
Underlying Models that Imply a LMA Model

that I know of...

- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).

- Can be derived from a distance based model (de Rooij & Heiser, 2005), and a generalization of Newton’s Law of Gravity (de Rooij, 2008).

- For dichotomous items, derived via the Dutch Identity (Holland, 1993).

- A generalization of the Dutch Identity to Rasch models for polytomous items, multidimensional traits and covariates. (Li, 2010).

- Can be derived from a formative latent variable model using statistical graphical models (Anderson & Böckenholt, 2000; Anderson & Vermunt, 2000; Anderson, 2002; Anderson & Tettagah, 2007).
Underlying Models that Imply a LMA Model

that I know of...

- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).

- Can be derived from a distance based model (de Rooij & Heiser, 2005), and a generalization of Newton’s Law of Gravity (de Rooij, 2008).

- For dichotomous items, derived via the Dutch Identity (Holland, 1993).

- A generalization of the Dutch Identity to Rasch models for polytomous items, multidimensional traits and covariates. (Li, 2010).

- Can be derived from a formative latent variable model using statistical graphical models (Anderson & Böckenholt, 2000; Anderson & Vermunt, 2000; Anderson, 2002; Anderson & Tettagah, 2007).

A Possible Graph & Model for Our Data

Got in fight
 Threatened to hurt
 Hit back
 Upset others for fun
 Help harass
 Tease others
A Possible Graph & Model for Our Data

- Got in fight
- Threatened to hurt
- Hit back
- Upset others for fun
- Help harass
- Tease others

Fight

\( \Theta_1 \)

Bully

\( \Theta_2 \)

Overview

Example Data Set

Existing Approaches

Log Multiplicative Association Models

Graphical Approach

○ A Possible Graph & Model for Our Data
○ Assumptions & Implications

Conditional Approach

Fighters, bullies and gender

Conclusions

A Nominal Example
A Possible Graph & Model for Our Data

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
● A Possible Graph & Model for Our Data
● Assumptions & Implications
Conditional Approach
Fighters, bullies and gender
Conclusions
A Nominal Example

Got in fight

Threatened to hurt

Hit back

Upset others for fun

Help harass

Tease others

Fight

Bully

$\Theta_1$

$\Theta_2$

$\nu_{1j1}$

$\nu_{2j1}$

$\nu_{3j1}$

$\nu_{4j2}$

$\nu_{5j2}$

$\nu_{6j2}$

$\sigma_{11}$

$\sigma_{12}$

$\sigma_{22}$
A Possible Graph & Model for Our Data

- Got in fight
- Threatened to hurt
- Hit back
- Upset others for fun
- Help harass
- Tease others

Fight

\[ \Theta_1 \]

Bully

\[ \Theta_2 \]

Gender

\[ \nu_{1j1}, \nu_{2j1}, \nu_{3j1}, \nu_{4j2}, \nu_{5j2}, \nu_{6j2}, \nu_{7j1}, \nu_{7j2}, \sigma_{11}, \sigma_{12}, \sigma_{22} \]
A Possible Graph & Model for Our Data

\[
\begin{align*}
\log(P(y)) &= \lambda + \sum_{i=1}^{7} \lambda_{ij} + \sum_{i=1}^{7} \sum_{k>i}^{7} \sum_{m=1}^{2} \sum_{m'>m} \sigma_{mm'} \nu_{ijm} \nu_{k\ell m'} \\
\Theta_1 &\text{ Fight} \\
\Theta_2 &\text{ Bully} \\
\nu_{1j1} &\text{ Got in fight} \\
\nu_{2j1} &\text{ Threatened to hurt} \\
\nu_{3j1} &\text{ Hit back} \\
\nu_{4j2} &\text{ Upset others for fun} \\
\nu_{5j2} &\text{ Help harass} \\
\nu_{6j2} &\text{ Tease others} \\
\nu_{7j1} &\text{ Gender} \\
\sigma_{11} &
\end{align*}
\]
The response pattern $y$ follows a multinomial distribution.
Assumptions & Implications

- The response pattern $y$ follows a multinomial distribution.
- Absence of a line connecting variables indicates conditional independence.
Assumptions & Implications

- The response pattern $y$ follows a multinomial distribution.
- Absence of a line connecting variables indicates conditional independence.
- For all possible response patterns $y$,

$$\Theta | y \sim MVN \left( \mu_y, \Sigma \right) \ i.i.d.$$ 

where

$$\mu_y = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1M} \\ \sigma_{12} & \sigma_{22} & \ldots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1M} & \sigma_{2M} & \ldots & \sigma_{MM} \end{pmatrix} \begin{pmatrix} \sum_i \nu_{ij1} \\ \sum_i \nu_{ij2} \\ \vdots \\ \sum_i \nu_{ijM} \end{pmatrix}$$
Assumptions & Implications

- The response pattern $y$ follows a multinomial distribution.
- Absence of a line connecting variables indicates conditional independence.
- For all possible response patterns $y$,

$$\Theta \mid y \sim MVN \left( \mu_y, \Sigma \right) \text{ i.i.d.}$$

where

$$\mu_y = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1M} \\ \sigma_{12} & \sigma_{22} & \ldots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1M} & \sigma_{2M} & \ldots & \sigma_{MM} \end{pmatrix} \begin{pmatrix} \sum_i \nu_{ij1} \\ \sum_i \nu_{ij2} \\ \vdots \\ \sum_i \nu_{ijM} \end{pmatrix}$$

Our example,

Fight:  \[ \mu_1 \mid y = \sigma_{11} \left( \sum_i \nu_{ij1} \right) + \sigma_{12} \left( \sum_i \nu_{ij2} \right) \]

Bully:  \[ \mu_2 \mid y = \sigma_{22} \left( \sum_i \nu_{ij2} \right) + \sigma_{12} \left( \sum_i \nu_{ij1} \right) \]
Conditional Approach

Consider the following conditional logistic regression model,

\[ P(Y_i = j | y_{k\ell}, k \neq i, x) = \frac{\exp(\lambda_{ij} + \sum_p \beta_{ijp} x_p + \sum_{k \neq i} \psi_{ij | k\ell})}{\sum_h \exp(\lambda_{ih} + \sum_p \beta_{ihp} x_p + \sum_{k \neq i} \psi_{ih | k\ell})} \]
Conditional Approach

Consider the following conditional logistic regression model,

\[
P(Y_i = j | y_{k\ell}, k \neq i, x) = \frac{\exp(\lambda_{ij} + \sum_p \beta_{ijp} x_p + \sum_{k \neq i} \psi_{ij|k\ell})}{\sum_h \exp(\lambda_{ih} + \sum_p \beta_{ihp} x_p + \sum_{k \neq i} \psi_{ih|k\ell})}
\]

where
- \(\lambda_{ij}\) is an intercept or location parameter.
- \(\psi_{ij|k\ell}\) is the parameter for variable \(y_k\) when predicting variable \(y_i\).
Conditional Approach

Consider the following conditional logistic regression model,

\[
P(Y_i = j|y_{k\ell}, k \neq i, \mathbf{x}) = \frac{\exp(\lambda_{ij} + \sum_p \beta_{ijp}x_p + \sum_{k \neq i} \psi_{ij|k\ell})}{\sum_h \exp(\lambda_{ih} + \sum_p \beta_{ihp}x_p + \sum_{k \neq i} \psi_{ih|k\ell})}
\]

where

- \(\lambda_{ij}\) is an intercept or location parameter.
- \(\psi_{ij|k\ell}\) is the parameter for variable \(y_k\) when predicting variable \(y_i\).

If we have this model for each item \(i (i = 1, \ldots, I)\) and \(\psi_{ij|k\ell} = \psi_{k\ell|ij}\), then model for the joint distribution of all items is

\[
\log(P(y_{1j}, \ldots, y_{IJ}|\mathbf{x})) = \lambda + \sum_i \lambda_{ij} + \sum_i \sum_p \beta_{ijp}x_p + \sum_i \sum_{k > i} \psi_{ij|k\ell}
\]

This is basically a log-linear model with all 2-way interactions.
Conditional Approach

Consider the following conditional logistic regression model,

\[
P(Y_i = j | y_{k\ell}, k \neq i, \mathbf{x}) = \frac{\exp(\lambda_{ij} + \sum_p \beta_{ijp} x_p + \sum_{k \neq i} \psi_{ij|k\ell})}{\sum_h \exp(\lambda_{ih} + \sum_p \beta_{ihp} x_p + \sum_{k \neq i} \psi_{ih|k\ell})}
\]

where

- \(\lambda_{ij}\) is an intercept or location parameter.
- \(\psi_{ij|k\ell}\) is the parameter for variable \(y_k\) when predicting variable \(y_i\).

If we have this model for each item \(i (i = 1, \ldots, I)\) and \(\psi_{ij|k\ell} = \psi_{k\ell|i,j}\), then model for the joint distribution of all items is

\[
\log(P(y_{1j}, \ldots, y_{Ij} | \mathbf{x})) = \lambda + \sum_i \lambda_{ij} + \sum_i \sum_p \beta_{ijp} x_p + \sum_i \sum_{k > i} \psi_{ij|k\ell}
\]

This is basically a log-linear model with all 2-way interactions.

Proof for the dichotomous, Joe & Liu (1996); for other cases, Anderson, Li & Vermunt (2007), and Anderson, Verkuilen & Peyton (in press).
Simplifying the Model

For each pair of items, there is a \((J \times L)\) matrix of \(\psi\)'s,

\[
\Psi_{i|k} = \begin{pmatrix}
\psi_{i1|k1} & \psi_{i1|k2} & \cdots & \psi_{i1|kL} \\
\psi_{i2|k1} & \psi_{i2|k2} & \cdots & \psi_{i2|kL} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{iJ|k1} & \psi_{iJ|k2} & \cdots & \psi_{iJ|kL}
\end{pmatrix}
\]
Simplifying the Model

For each pair of items, there is a \((J \times L)\) matrix of \(\psi\)'s,

\[
\Psi_{i|k} = \begin{pmatrix}
\psi_{i1|k1} & \psi_{i1|k2} & \cdots & \psi_{i1|kL} \\
\psi_{i2|k1} & \psi_{i2|k2} & \cdots & \psi_{i2|kL} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{iJ|k1} & \psi_{iJ|k2} & \cdots & \psi_{iJ|kL}
\end{pmatrix}
\]

The model can be simplified by considering lower rank decompositions:

\[
\Psi_{i|k} = \mathbf{N}_{i}^{[ik]} \Sigma^{[ik]} \mathbf{N}_{k}^{[ik]}' \quad \text{where} \quad \Sigma^{[ik]} \text{ is diagonal.}
\]
Simplifying the Model

For each pair of items, there is a \((J \times L)\) matrix of \(\psi\)'s,

\[
\Psi_{i|k} = \begin{pmatrix}
\psi_{i1|k1} & \psi_{i1|k2} & \cdots & \psi_{i1|kL} \\
\psi_{i2|k1} & \psi_{i2|k2} & \cdots & \psi_{i2|kL} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{iJ|k1} & \psi_{iJ|k2} & \cdots & \psi_{iJ|kL}
\end{pmatrix}
\]

The model can be simplified by considering lower rank decompositions:

\[
\Psi_{i|k} = N_i^{[ik]} \Sigma^{[ik]} N_k^{[ik]}'
\]

where \(\Sigma^{[ik]}\) is diagonal.

However, here we’ll mostly consider those of the form

\[
\Psi_{i|k} = N_i \Sigma N_k'
\]

where \(\Sigma\) is not necessarily diagonal and

\[
N_i = \begin{pmatrix}
\nu_{i11} & \nu_{i12} & \cdots & \nu_{i1M} \\
\nu_{i21} & \nu_{i22} & \cdots & \nu_{i2M} \\
\vdots & \vdots & \ddots & \vdots \\
\nu_{iJ1} & \nu_{iJ2} & \cdots & \nu_{iJM}
\end{pmatrix}
\]
**Overview**

**Example Data Set**

**Existing Approaches**

- Log Multiplicative Association Models
- Graphical Approach
- Conditional Approach
  - Simplifying the Model
  - Special Case #1: $M = 1$
  - Special Case #2: $M$
  - Recent Developments: LMA as IRT Models
  - Common and Novel IRT Models as LMAs

**Conditional Approach**

- Fighters, bullies and gender

**Conclusions**

**A Nominal Example**

---

**Special Case #1: $M = 1$**

$$
\Psi_i|k = \nu_i \sigma_{11} \nu'_k = \{\sigma_{11} \nu_{ij} \nu_{k\ell}\}
$$
Special Case \#1: \( M = 1 \)

\[ \Psi_{i|k} = \nu_{i1} \sigma_{11}^\prime \nu_{k1} = \{ \sigma_{11} \nu_{ij1} \nu_{k\ell1} \} \]

The conditional logistic regression model for each item \( i \) is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))}
\]
Special Case \#1: $M = 1$

\[ \Psi_{i|k} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k\ell1}\} \]

The conditional logistic regression model for each item $i$ is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11}\sum_{k \neq i} \nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11}\sum_{k \neq i} \nu_{k\ell1}))} = \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}
\]
Special Case #1: $M = 1$

$$\Psi_{ik} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k1}\}$$

The conditional logistic regression model for each item $i$ is

$$P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}\sigma_{11}\sum_{k\neq i}\nu_{k1})}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}\sigma_{11}\sum_{k\neq i}\nu_{k1})}$$

$$= \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}$$

$\lambda_{ij} = b_{ij}$ is an intercept or "difficulty" parameter.
Special Case #1: $M = 1$

\[
\Psi_{i|k} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k1}\}
\]

The conditional logistic regression model for each item $i$ is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11}\sum_{k \neq i} \nu_{k1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11}\sum_{k \neq i} \nu_{k1}))} = \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}
\]

- $\lambda_{ij} = b_{ij}$ is an intercept or “difficulty” parameter.
- $\nu_{ij1} = a_{ij}$ is a slope or “discrimination” parameter.
Special Case  #1:  $M = 1$

$$\Psi_{i|k} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k\ell1}\}$$

The conditional logistic regression model for each item $i$ is

$$P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))} = \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}$$

- $\lambda_{ij} = b_{ij}$ is an intercept or “difficulty” parameter.
- $\nu_{ij1} = a_{ij}$ is a slope or “discrimination” parameter.
- The predictor variable is a (weighted) rest-score: $\tilde{\theta} = \sigma_{11} \sum_{k \neq i} \nu_{k\ell1}$. 
Special Case \#1: \( M = 1 \)

\[
\Psi_{i|k} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k\ell1}\}
\]

The conditional logistic regression model for each item \( i \) is

\[
P(Y_i = j|y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11}\sum_{k\neq i}\nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11}\sum_{k\neq i}\nu_{k\ell1}))}
\]

\[
= \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}
\]

- \( \lambda_{ij} = b_{ij} \) is an intercept or “difficulty” parameter.
- \( \nu_{ij1} = a_{ij} \) is a slope or “discrimination” parameter.
- The predictor variable is a (weighted) rest-score: \( \tilde{\theta} = \sigma_{11}\sum_{k\neq i}\nu_{k\ell1} \).

Justification, see Junker (1993), and Junker & Sijtsma (2000)
Special Case \#1: \( M = 1 \)

\[
\Psi_{i|k} = \nu_{i1} \sigma_{11} \nu'_{k1} = \{\sigma_{11} \nu_{ij1} \nu_{k\ell1}\}
\]

The conditional logistic regression model for each item \( i \) is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1} (\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1} (\sigma_{11} \sum_{k \neq i} \nu_{k\ell1}))} = \frac{\exp(b_{ij} + a_{ij} \tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih} \tilde{\theta})}
\]

- \( \lambda_{ij} = b_{ij} \) is an intercept or "difficulty" parameter.
- \( \nu_{ij1} = a_{ij} \) is a slope or "discrimination" parameter.
- The predictor variable is a (weighted) rest-score: \( \tilde{\theta} = \sigma_{11} \sum_{k \neq i} \nu_{k\ell1} \).

Justification, see Junker (1993), and Junker & Sijtsma (2000)

- Bock’s nominal response model and all it’s special cases.
Special Case #1: $M = 1$

$$\Psi_{i|k} = \nu_{i1}\sigma_{11}\nu'_{k1} = \{\sigma_{11}\nu_{ij1}\nu_{k\ell1}\}$$

The conditional logistic regression model for each item $i$ is

$$P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \nu_{ij1}(\sigma_{11}\sum_{k \neq i}\nu_{k\ell1}))}{\sum_h \exp(\lambda_{ih} + \nu_{ih1}(\sigma_{11}\sum_{k \neq i}\nu_{k\ell1}))} = \frac{\exp(b_{ij} + a_{ij}\tilde{\theta})}{\sum_h \exp(b_{ih} + a_{ih}\tilde{\theta})}$$

- $\lambda_{ij} = b_{ij}$ is an intercept or “difficulty” parameter.
- $\nu_{ij1} = a_{ij}$ is a slope or “discrimination” parameter.
- The predictor variable is a (weighted) rest-score: $\tilde{\theta} = \sigma_{11}\sum_{k \neq i}\nu_{k\ell1}$.
- Justification, see Junker (1993), and Junker & Sijtsma (2000)
- Bock’s nominal response model and all it’s special cases.
- The LMA

$$P(y) = \lambda + \sum_i \lambda_{ij} + \sigma_{11} \sum_i \sum_{k > i} \nu_{ij1}\nu_{k\ell1}$$
Special Case #2: $M$

$$\Psi_{i|k} = N_i \sum N_k' = \left\{ \sum_{m} \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{klm'} \right\}$$
**Special Case #2: \( M \)**

\[
\Psi_{i|k} = N_i \sum N'_k = \left\{ \sum_m \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{k\ell m'} \right\}
\]

The conditional logistic regression model for each item \( i \) is

\[
P(Y_i = j \mid y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \sum_m \nu_{ijm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}{\sum_h \exp(\lambda_{ih} + \sum_m \nu_{ihm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}
\]
Special Case #2: $M$

\[ \Psi_{i|k} = N_i \sum N'_k = \left\{ \sum_m \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{k\ell m'} \right\} \]

The conditional logistic regression model for each item $i$ is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \sum_m \nu_{ijm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}{\sum_h \exp(\lambda_{ih} + \sum_m \nu_{ihm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))} \]

\[
= \frac{\exp(b_{ij} + \sum_m a_{ijm} \tilde{\theta}_m)}{\sum_h \exp(b_{ih} + \sum_m a_{ihm} \tilde{\theta}_m)}
\]
Special Case #2: $M$

\[
\Psi_i | k = N_i \Sigma N'_k = \left\{ \sum_m \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{k\ell m'} \right\}
\]

The conditional logistic regression model for each item $i$ is

\[
P(Y_i = j | y_k\ell, k \neq i) = \frac{\exp(\lambda_{ij} + \sum_m \nu_{ijm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}{\sum_h \exp(\lambda_{ih} + \sum_m \nu_{ihm} (\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))} = \frac{\exp(b_{ij} + \sum_m a_{ijm} \tilde{\theta}_m)}{\sum_h \exp(b_{ih} + \sum_m a_{ihm} \tilde{\theta}_m)}
\]

$\nu_{ijm} = a_{ijm}$ is the slope or discrimination parameter for variable $m$. 

---

**Overview**

- Example Data Set
- Existing Approaches
- Log Multiplicative Association Models
- Graphical Approach

**Conditional Approach**
- Simplifying the Model
- Special Case #1: $M = 1$
- Special Case #2: $M$
- Recent Developments: LMA as IRT Models
- Common and Novel IRT Models as LMAs

**Fighters, bullies and gender**

**Conclusions**

**A Nominal Example**
Special Case #2: $M$

\[
\Psi_{i|k} = N_i \Sigma N'_k = \left\{ \sum_m \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{k\ell m'} \right\}
\]

The conditional logistic regression model for each item $i$ is

\[
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \sum_m \nu_{ijm}(\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}{\sum_h \exp(\lambda_{ih} + \sum_m \nu_{ihm}(\sum_{m'} \sigma_{mm'} \sum_{k \neq i} \nu_{k\ell m'}))}
= \frac{\exp(b_{ij} + \sum_m a_{ijm} \tilde{\theta}_m)}{\sum_h \exp(b_{ih} + \sum_m a_{ihm} \tilde{\theta}_m)}
\]

- $\nu_{ijm} = a_{ijm}$ is the slope or discrimination parameter for variable $m$.
- $\tilde{\theta}_m$ is predictor variable $m$ or weighted sum of rest-scores or test-totals:

\[
\tilde{\theta}_m = \sigma_{mm'} \left( \sum_{k \neq i} \nu_{k\ell m} \right) + \sum_{m' \neq m} \sigma_{mm'} \left( \sum_{k \neq i} \nu_{k\ell m'} \right)
\]

rest-score $m$  \quad rest-score or test-total $m'$
Special Case \#2: $M$

$$
\Psi_{i|k} = N_i \Sigma N'_{k} = \left\{ \sum_{m} \sum_{m'} \nu_{ijm} \sigma_{mm'} \nu_{k\ell m'} \right\}
$$

The conditional logistic regression model for each item $i$ is

$$
P(Y_i = j | y_{k\ell}, k \neq i) = \frac{\exp(\lambda_{ij} + \sum_{m} \nu_{ijm} (\sum_{m'} \sigma_{mm'} \sum_{k\neq i} \nu_{k\ell m'}))}{\sum_{h} \exp(\lambda_{ih} + \sum_{m} \nu_{ihm} (\sum_{m'} \sigma_{mm'} \sum_{k\neq i} \nu_{k\ell m'}))}
\quad = \frac{\exp(b_{ij} + \sum_{m} a_{ijm} \tilde{\theta}_m)}{\sum_{h} \exp(b_{ih} + \sum_{m} a_{ihm} \tilde{\theta}_m)}
$$

- $\nu_{ijm} = a_{ijm}$ is the slope or discrimination parameter for variable $m$.
- $\tilde{\theta}_m$ is predictor variable $m$ or weighted sum of rest-scores or test-totals:

$$
\tilde{\theta}_m = \sigma_{mm} \left( \sum_{k\neq i} \nu_{k\ell m} \right) + \sum_{m' \neq m} ^{\sigma_{mm'}} \left( \sum_{k\neq i} ^{\nu_{k\ell m'}} \right).
$$

- Multidimensional compensatory IRT model for polytomous items.
Recent Developments: LMA as IRT Models

- Anderson, Li & Vermunt (2007):
  - Models in the Rasch family—dichotomous and polytomous items, uni- and multi-dimensional latent variables.
  - Pseudo-likelihood estimation.

Recent Developments: LMA as IRT Models

- Anderson, Li & Vermunt (2007):
  - Models in the Rasch family—dichotomous and polytomous items, uni- and multi-dimensional latent variables.
  - Pseudo-likelihood estimation.
Recent Developments: LMA as IRT Models

- Anderson, Li & Vermunt (2007):
  - Models in the Rasch family—dichotomous and polytomous items, uni- and multi-dimensional latent variables.
  - Pseudo-likelihood estimation.

- Anderson & Yu (2007):
  - Dichotomous items, 1 underlying latent variable.
  - Different underlying alternative marginal distributions of the latent variable.

- Recent Developments: LMA as IRT Models
  - Models in the Rasch family—dichotomous and polytomous items, uni- and multi-dimensional latent variables.
  - Pseudo-likelihood estimation.

- Anderson & Y u (2007):
  - Dichotomous items, 1 underlying latent variable.
  - Different underlying alternative marginal distributions of the latent variable.
Recent Developments: LMA as IRT Models

- Anderson, Li & Vermunt (2007):
  - Models in the Rasch family—dichotomous and polytomous items, uni– and multi–dimensional latent variables.
  - Pseudo-likelihood estimation.

- Anderson & Yu (2007):
  - Dichotomous items, 1 underlying latent variable.
  - Different underlying alternative marginal distributions of the latent variable.

- Anderson, Verkuilen & Peyton (in press):
  - Multicategory items and two latent variables (also 3 latent variable and higher order models, but these aren’t in the paper).
  - Covariate that influenced the choice of the “don’t know” response option (i.e., instructions). The covariate came in from the “left”.
  - Hybrid model.
  - Equality restrictions on some parameters.
  - Models fit using SAS/NLP.
## Common and Novel IRT Models as LMAs

<table>
<thead>
<tr>
<th>Common models</th>
<th>Restrictions on LMA Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>(\lambda_{ij} (b_{ij})) none (\nu_{ijm} (a_{ijm})) none</td>
</tr>
<tr>
<td>Nominal response model</td>
<td>none</td>
</tr>
<tr>
<td>Multidimensional compensatory</td>
<td>none</td>
</tr>
<tr>
<td>Rasch family</td>
<td>none</td>
</tr>
<tr>
<td>Graded × Nominal response</td>
<td>none</td>
</tr>
</tbody>
</table>

### Restrictions on LMA Parameters
- \(\lambda_{ij} (b_{ij})\) none
- \(\nu_{ijm} (a_{ijm})\) none
- input/fixed (ordered)
- ordinal

### Conditional Approach
- Special Case #1: \(M = 1\)
- Special Case #2: \(M\)
- Recent Developments: LMA as IRT Models
- Common and Novel IRT Models as LMAs

### Restrictions
- Common models
- Nominal response model
- Multidimensional compensatory
- Rasch family
- Graded × Nominal response

### Conditions
- \(M = 1\)
- \(M\)
- Input/fixed (ordered)
- Ordinal
# Common and Novel IRT Models as LMAs

<table>
<thead>
<tr>
<th>Common models</th>
<th>Restrictions on LMA Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>( \lambda_{ij} (\beta_{ij}) )</td>
</tr>
<tr>
<td>Nominal response model</td>
<td>none</td>
</tr>
<tr>
<td>Multidimensional compensatory</td>
<td>none</td>
</tr>
<tr>
<td>Rasch family</td>
<td>none</td>
</tr>
<tr>
<td>Graded × Nominal response</td>
<td>none</td>
</tr>
</tbody>
</table>

**Novel ones** can be created by modeling or placing restrictions on location parameters (i.e., \( \lambda_{ij} \)), category scores (i.e., \( \nu_{ijm} \)), and/or \( \sigma_{mm} \)s:
Common and Novel IRT Models as LMAs

<table>
<thead>
<tr>
<th>Common models</th>
<th>Restrictions on LMA Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>$\lambda_{ij}$ ($b_{ij}$)</td>
</tr>
<tr>
<td>Nominal response model</td>
<td>none</td>
</tr>
<tr>
<td>Multidimensional compensatory</td>
<td>none</td>
</tr>
<tr>
<td>Rasch family</td>
<td>none</td>
</tr>
<tr>
<td>Graded × Nominal response</td>
<td>none</td>
</tr>
</tbody>
</table>

Novel ones can be created by modeling or placing restrictions on location parameters (i.e., $\lambda_{ij}$), category scores (i.e., $\nu_{ijm}$), and/or $\sigma_{mm}$s:

- Input/fixed.
- Equality.
- Ordinal.
- Linear functions (e.g.,
  \[ \nu_{ij} = \omega_i x_{ij} \text{ where} \]
  \[ x_{ij} = 0, 1, \ldots J \text{ or any values}. \]
## Common and Novel IRT Models as LMAs

<table>
<thead>
<tr>
<th>Common models</th>
<th>Restrictions on LMA Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>λ_{ij} \ (b_{ij}) \ ν_{ijm} \ (a_{ijm})</td>
</tr>
<tr>
<td>Nominal response model</td>
<td>none</td>
</tr>
<tr>
<td>Multidimensional compensatory</td>
<td>none</td>
</tr>
<tr>
<td>Rasch family</td>
<td>none</td>
</tr>
<tr>
<td>Graded × Nominal response</td>
<td>none</td>
</tr>
</tbody>
</table>

### Novel ones

Can be created by modeling or placing restrictions on location parameters (i.e., λ_{ij}), category scores (i.e., ν_{ijm}), and/or σ_{mm}:

- Input/fixed.
- Equality.
- Ordinal.
- Linear functions (e.g., ν_{ij} = \omega_i x_{ij} where x_{ij} = 0, 1, \ldots J or any values).

### Restrictions on LMA Parameters

- Set minimum and maximum (e.g., 0 and 1) and estimate scores in between.
- Model σ_{mm} (e.g.,
  \[ σ_{mm} = σ_{mm}^* + β_m x \] or
  \[ σ_{mm} = σ_{mm}^* β_m x \].)
Common and Novel IRT Models as LMAs

<table>
<thead>
<tr>
<th>Common models</th>
<th>Restrictions on LMA Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>( \lambda_{ij} ) (( b_{ij} ))</td>
</tr>
<tr>
<td>Nominal response model</td>
<td>none</td>
</tr>
<tr>
<td>Multidimensional compensatory</td>
<td>none</td>
</tr>
<tr>
<td>Rasch family</td>
<td>none</td>
</tr>
<tr>
<td>Graded ( \times ) Nominal response</td>
<td>none</td>
</tr>
</tbody>
</table>

Novel ones can be created by modeling or placing restrictions on location parameters (i.e., \( \lambda_{ij} \)), category scores (i.e., \( \nu_{ijm} \)), and/or \( \sigma_{mm} \)s:

- Input/fixed.
- Equality.
- Ordinal.
- Linear functions (e.g., \( \nu_{ij} = \omega_i x_{ij} \) where \( x_{ij} = 0, 1, \ldots J \) or any values).
- Other: e.g., \( \nu_{ij} = \omega_i \nu_{ij}^* \) where \( \sum_j \nu_{ij}^* = 1 \) or set min and max \( \nu_{ij}^* \).
-2lnlike versus Number Parameters

- Independence
- Rasch 1D
- Rasch 2D

Color → Restrictions on parameters

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
- -2lnlike versus Number Parameters
- Estimated Item Scale Values
- Item Characteristic Curves
- Item Cumulative Probability Curves
- Estimated Scale Values for Gender
Conclusions
A Nominal Example
Overview

Example Data Set

Existing Approaches

Log Multiplicative Association Models

Graphical Approach

Conditional Approach

Fighters, bullies and gender

-2lnlike versus Number Parameters

-2lnlike versus Number Parameters w/ Gender

Estimated Item Scale Values

Item Characteristic Curves

Item Cumulative Probability Curves

Estimated Scale Values for Gender

Conclusions

A Nominal Example

-2lnlike versus Number Parameters
-2lnlike versus Number Parameters

- Independence
- Rasch 1D
- Rasch 2D
- log-linear X linear
- Ima1D
- Ima2D
-2lnlike versus Number Parameters

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
- Independence
- Rasch 1D
- Rasch 2D

Color → Restrictions on parameters

Advances in Models for Multivariate Nominal or Ordinal Variables
-2lnlike versus Number Parameters

- Independence
- Rasch 1D
- Rasch 2D

Color → Restrictions on parameters

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
- Independence
- Rasch 1D
- Rasch 2D

Conclusions
A Nominal Example
-2lnlike versus Number Parameters

- Independence
- Rasch 1D
- Rasch 2D

Color → Restrictions on parameters

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
- Independence
- Rasch 1D
- Rasch 2D

Features:
- Estimated Item Scale Values
- Item Characteristic Curves
- Item Cumulative Probability Curves
- Estimated Scale Values for Gender

Conclusions
A Nominal Example

Advances in Models for Multivariate Nominal or Ordinal Variables
-2lnlike versus Number Parameters

- Independence
- Rasch 1D
- Rasch 2D
- Ima2D-LinearScores
- Ima2D-MixedScores
- Ima2D-OrdinalScores
- log-linearXlinear
- Min/MaxScore
- Ima1D
- Ima2D

Color → Restrictions on parameters

- Overview
- Example Data Set
- Existing Approaches
  - Log Multiplicative Association Models
  - Graphical Approach
  - Conditional Approach
- Fighters, bullies and gender
  - Independence
  - Rasch 1D
  - Rasch 2D
  - Ima2D-LinearScores
- Conclusions
- A Nominal Example
-2lnlike versus Number Parameters

- Independence

Color → Restrictions on parameters

- Rasch 1D

- Rasch 2D

Ima2D-LinearScores

log-linearXlinear

Min/MaxScore

Ima1D

Ima2D

Ima2D-MixedScores

MixedScores/OrdinalIntercepts

Ima2D-OrdinalScores

Overview

Example Data Set

Existing Approaches

Log Multiplicative Association Models

Graphical Approach

Conditional Approach

Fighters, bullies and gender

- Independence

- Rasch 1D

- Rasch 2D

Ima2D-LinearScores

log-linearXlinear

Min/MaxScore

Ima1D

Ima2D

Ima2D-MixedScores

MixedScores/OrdinalIntercepts

Ima2D-OrdinalScores
-2lnlike versus Number Parameters w/ Gender

Red ——> Gender included
-2lnlike versus Number Parameters w/ Gender

Independence

Red → Gender included

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
-2lnlike versus Number Parameters
-2lnlike versus Number Parameters w/ Gender
Estimated Item Scale Values
Item Characteristic Curves
Item Cumulative Probability Curves
Estimated Scale Values for Gender
Conclusions
A Nominal Example
Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach

Fighters, bullies and gender
-2lnlike versus Number Parameters
-2lnlike versus Number Parameters w/ Gender
Estimated Item Scale Values
Item Characteristic Curves
Item Cumulative Probability Curves
Estimated Scale Values for Gender

Conclusions
A Nominal Example

Advances in Models for Multivariate Nominal or Ordinal Variables
-2lnlike versus Number Parameters w/ Gender

Independence

Red → Gender included

Advances in Models for Multivariate Nominal or Ordinal Variables
-2lnlike versus Number Parameters w/ Gender

Independence

Red → Gender included

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
-2lnlike versus Number Parameters
-2lnlike versus Number Parameters w/ Gender
Estimated Item Scale Values
Item Characteristic Curves
Item Cumulative Probability Curves
Estimated Scale Values for Gender
Conclusions
A Nominal Example

Advances in Models for Multivariate Nominal or Ordinal Variables
Estimated Item Scale Values

Got in Fight

Threatened to hit/hurt

Hit back

Upset others for fun

Help harass

Tease Other Students

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options

Estimated $v$

Response Options
Item Characteristic Curves

- Got in Fight
- Threatened to Hurt
- Hit Back
- Upset Others for Fun
- Help to Harass Students
- Tease Other Students

- Overview
- Example Data Set
- Existing Approaches
  - Log Multiplicative Association Models
  - Graphical Approach
  - Conditional Approach
- Fighters, bullies and gender
  - $-2 \ln \text{like}$ versus Number Parameters
  - $-2 \ln \text{like}$ versus Number Parameters with Gender
    - Estimated Item Scale Values
    - Item Characteristic Curves
    - Item Cumulative Probability Curves
    - Estimated Scale Values for Gender
- Conclusions
  - A Nominal Example
Item Cumulative Probability Curves

- Got in Fight
- Threatened to Hurt
- Hit Back
- Upset Others for Fun
- Help to Harass Students
- Tease Other Students
Estimated Scale Values for Gender

Gender on Fight

Gender on Bully

Gender

Male Female

Estimated v

Male Female

Estimated v

-0.5 -0.3 -0.1 0.0

-0.5 -0.3 -0.1 0.0

-2lnlike versus Number Parameters
-2lnlike versus Number Parameters w/ Gender
Estimated Item Scale Values
Item Characteristic Curves
Item Cumulative Probability Curves
Estimated Scale Values for Gender

Conclusions

A Nominal Example
Example Illustrated...

Noteworthy in today’s example for multycategory items:

- Marginal distribution of traits are very skewed.
Example Illustrated...

Noteworthy in today’s example for multcategory items:

- Marginal distribution of traits are very skewed.
- Ordinal restrictions on responses options (and linear transformations).
Example Illustrated...

Noteworthy in today’s example for multcategory items:

- Marginal distribution of traits are very skewed.
- Ordinal restrictions on responses options (and linear transformations).
- Ordinal restrictions on intercepts (“difficulties” or location parameters).
Example Illustrated... 

Noteworthy in today’s example for multcategory items:

- Marginal distribution of traits are very skewed.
- Ordinal restrictions on responses options (and linear transformations).
- Ordinal restrictions on intercepts (“difficulties” or location parameters).
- Covariate came in from the “right” (i.e., helps model underlying latent variable).
Example Illustrated... 

Noteworthy in today’s example for multcategory items:

- Marginal distribution of traits are very skewed.
- Ordinal restrictions on responses options (and linear transformations).
- Ordinal restrictions on intercepts (“difficulties” or location parameters).
- Covariate came in from the “right” (i.e., helps model underlying latent variable).
- The scale values for categories of the 3 three bully items are nearly identical from a uni-dimensional LMA model fit to all 9 bully items (and different estimation methods). Correlations > .99.
Example Illustrated...

Noteworthy in today’s example for multcategory items:

- Marginal distribution of traits are very skewed.
- Ordinal restrictions on responses options (and linear transformations).
- Ordinal restrictions on intercepts (“difficulties” or location parameters).
- Covariate came in from the “right” (i.e., helps model underlying latent variable).

- The scale values for categories of the 3 three bully items are nearly identical from a uni-dimensional LMA model fit to all 9 bully items (and different estimation methods). Correlations > .99.

The Importance of this: Illustrates that major criticisms of conditional models do not hold up for LMA models as latent variable models:

- Models parameters are essentially the same.
- No interpretational difficulty.
The major problem is the size of a table (i.e., number of items/categories), but not the number of latent variables.

- The largest problem that I’ve successfully fit using $\ell_{EM}$ (Vermunt, 1997) is $2^{12} = 4096$ response patterns.
Estimation Developments

The major problem is the size of a table (i.e., number of items/categories), but not the number of latent variables.

- The largest problem that I’ve successfully fit using $\ell_{EM}$ (Vermunt, 1997) is $2^{12} = 4096$ response patterns.

- Bayesian methods for the $RC(M)$ association model for 2-way tables (Iliopoulos, Kateri, & Ntzoufras, 2007; Iliopoulos & Kateri, 2009)
The major problem is the size of a table (i.e., number of items/categories), but not the number of latent variables.

- The largest problem that I’ve successfully fit using $\ell_{EM}$ (Vermunt, 1997) is $2^{12} = 4096$ response patterns.

- **Bayesian methods** for the $RC(M)$ association model for 2-way tables (Iliopoulos, Kateri, & Ntzoufras, 2007; Iliopoulos & Kateri, 2009)

- Models can be fit to data using **SAS/NLP** (probably also R and MatLab using their optimization capabilities.). Although this approach can fit larger numbers of items/categories than $\ell_{EM}$, it is still somewhat limited.
Estimation Developments

The major problem is the size of a table (i.e., number of items/categories), but not the number of latent variables.

- The largest problem that I’ve successfully fit using $\ell EM$ (Vermunt, 1997) is $2^{12} = 4096$ response patterns.

- Bayesian methods for the $RC(M)$ association model for 2-way tables (Iliopoulos, Kateri, & Ntzoufras, 2007; Iliopoulos & Kateri, 2009)

- Models can be fit to data using SAS/NLP (probably also R and MatLab using their optimization capabilities.). Although this approach can fit larger numbers of items/categories than $\ell EM$, it is still somewhat limited.

- Models in the Rasch family can be fit by pseudo-likelihood estimation in any program that can fit conditional logistic regression models (Anderson, Li & Vermunt, 2007) and can include covariates.

No limit hit (yet) in terms of number of items/categories or number of latent variables.
The major problem is the size of a table (i.e., number of items/categories), but not the number of latent variables.

- The largest problem that I’ve successfully fit using $\ell EM$ (Vermunt, 1997) is $2^{12} = 4096$ response patterns.

- **Bayesian methods** for the $RC(M)$ association model for 2-way tables (Iliopoulos, Kateri, & Ntzoufras, 2007; Iliopoulos & Kateri, 2009)

- Models can be fit to data using SAS/NLP (probably also R and MatLab using their optimization capabilities.). Although this approach can fit larger numbers of items/categories than $\ell EM$, it is still somewhat limited.

- Models in the Rasch family can be fit by **pseudo-likelihood estimation** in any program that can fit conditional logistic regression models (Anderson, Li & Vermunt, 2007) and can include covariates. 

  No limit hit (yet) in terms of number of items/categories or number of latent variables.

- For models with estimated category scores, an **experimental algorithm**.
For models with estimated category scores, an experimental algorithm that iteratively fits conditional logistic regressions using MLE to estimate scale values and a pseudo-likelihood step to estimate the association parameters.

- Takes advantage of conditional specification of models.
Experimental Algorithm

For models with estimated category scores, an experimental algorithm that iteratively fits conditional logistic regressions using MLE to estimate scale values and a pseudo-likelihood step to estimate the association parameters.

- Takes advantage of conditional specification of models.
- Applications to data sets yield nearly identical estimates as MLE.
Experimental Algorithm

For models with estimated category scores, an experimental algorithm that iteratively fits conditional logistic regressions using MLE to estimate scale values and a pseudo-likelihood step to estimate the association parameters.

- Takes advantage of conditional specification of models.
- Applications to data sets yield nearly identical estimates as MLE.
- In simulation studies, the algorithm yields parameters estimates nearly identical (up to linear transformation) of the parameters used to simulate data (by some IRT model).
Experimental Algorithm

For models with estimated category scores, an experimental algorithm that iteratively fits conditional logistic regressions using MLE to estimate scale values and a pseudo-likelihood step to estimate the association parameters.

- Takes advantage of conditional specification of models.
- Applications to data sets yield nearly identical estimates as MLE.
- In simulation studies, the algorithm yields parameters estimates nearly identical (up to linear transformation) of the parameters used to simulate data (by some IRT model).
- Converges relatively quickly.
For models with estimated category scores, an experimental algorithm that iteratively fits conditional logistic regressions using MLE to estimate scale values and a pseudo-likelihood step to estimate the association parameters.

- Takes advantage of conditional specification of models.
- Applications to data sets yield nearly identical estimates as MLE.
- In simulation studies, the algorithm yields parameters estimates nearly identical (up to linear transformation) of the parameters used to simulate data (by some IRT model).
- Converges relatively quickly.
- I haven’t hit a limit in terms of number of items/categories.
Estimation.
Areas for Future Research

- Estimation.

- Applications, especially those with multiple latent variables (e.g., testlets, Q-matrix).
Areas for Future Research

- Estimation.

- Applications, especially those with multiple latent variables (e.g., testlets, Q-matrix).

- Further comparisons with traditional IRT methods (i.e., estimation of item parameters and individuals’ values on latent variables).
## Areas for Future Research

- Estimation.

- Applications, especially those with multiple latent variables (e.g., testlets, Q-matrix).

- Further comparisons with traditional IRT methods (i.e., estimation of item parameters and individuals’ values on latent variables).

- Handling missing data
Areas for Future Research

- Estimation.

- Applications, especially those with multiple latent variables (e.g., testlets, Q-matrix).

- Further comparisons with traditional IRT methods (i.e., estimation of item parameters and individuals’ values on latent variables).

- Handling missing data

- Addition of random effects
Areas for Future Research

- Estimation.
- Applications, especially those with multiple latent variables (e.g., testlets, Q-matrix).
- Further comparisons with traditional IRT methods (i.e., estimation of item parameters and individuals’ values on latent variables).
- Handling missing data
- Addition of random effects
- Multidimensional, partially-compensatory models leads to higher-way interactions in model for the data where the higher-way interactions have higher-way decompositions (Tucker 3-mode and other higher-way type decompositions).
Will be able to download SAS/NLP programs used in this talk and various papers from

http://faculty.ed.uiuc.edu/cja/homepage/software_index.html

and slides from

http://faculty.ed.uiuc.edu/cja/homepage
A: Who has the final responsibility to decide if a law is constitutional or not?
A: Who has the final responsibility to decide if a law is constitutional or not?

President (9.0%), Congress (27.6%), Supreme Court (57.9%)
A: Who has the final responsibility to decide if a law is constitutional or not?

President (9.0%), Congress (27.6%), Supreme Court (57.9%), Don’t know (5.5%)

A: Who has the final responsibility to decide if a law is constitutional or not?
   President (9.0%), Congress (27.6%), Supreme Court (57.9%), Don’t know (5.5%)

B: Whose responsibility is it to nominate judges to the Federal courts?
   President (50.8%), Congress (15.6%), Supreme Court (25.5%), Don’t know (8.2%)

A: Who has the final responsibility to decide if a law is constitutional or not?
   President (9.0%), Congress (27.6%), Supreme Court (57.9%), Don’t know (5.5%)

B: Whose responsibility is it to nominate judges to the Federal courts?
   President (50.8%), Congress (15.6%), Supreme Court (25.5%), Don’t know (8.2%)

C: Do you happen to know which party has the most members in the House of Representatives in Washington? Republicans (70.3%), Democrats (16.0%), Don’t know (13.7%)

A: Who has the final responsibility to decide if a law is constitutional or not?
   President (9.0%), Congress (27.6%), Supreme Court (57.9%), Don’t know (5.5%)

B: Whose responsibility is it to nominate judges to the Federal courts?
   President (50.8%), Congress (15.6%), Supreme Court (25.5%), Don’t know (8.2%)

C: Do you happen to know which party has the most members in the House of Representatives in Washington? Republicans (70.3%), Democrats (16.0%), Don’t know (13.7%)

D: Do you happen to know which party has the most members in the U.S. Senate? Republicans (62.8%), Democrats (19.0%), Don’t know (18.2%)

A: Who has the final responsibility to decide if a law is constitutional or not?
   President (9.0%), Congress (27.6%), Supreme Court (57.9%), Don’t know (5.5%)

B: Whose responsibility is it to nominate judges to the Federal courts?
   President (50.8%), Congress (15.6%), Supreme Court (25.5%), Don’t know (8.2%)

C: Do you happen to know which party has the most members in the House of Representatives in Washington? Republicans (70.3%), Democrats (16.0%), Don’t know (13.7%)

D: Do you happen to know which party has the most members in the U.S. Senate? Republicans (62.8%), Democrats (19.0%), Don’t know (18.2%)

Ordering of response the options unknown.
Challanges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
- Scoring of responses needed.
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
- Scoring of responses needed.
- Latent variable structure unknown.
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
- Scoring of responses needed.
- Latent variable structure unknown.
  - One dominant underlying dimension of “Knowledge.”
Challanges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
- Scoring of responses needed.
- Latent variable structure unknown.
  - One dominant underlying dimension of “Knowledge.”
  - Two correlated latent variables: Structure of political system (items A & B) and Party in Power (items C & D)
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
- Different number of response options.
- Scoring of responses needed.
- Latent variable structure unknown.
  - One dominant underlying dimension of “Knowledge.”
  - Two correlated latent variables: Structure of political system (items A & B) and Party in Power (items C & D)
- How do the instructions given to respondents affect all of this?
Overview

Example Data Set

Existing Approaches

Log Multiplicative Association Models

Graphical Approach

Conditional Approach

Fighters, bullies and gender

Conclusions

A Nominal Example

● Challenges for Analysis of ANES Data

Effect of Instructions:
Constitutionality of Laws

ANES Graphs
ANES Graphs

Instructions

A
B
C
D

\(\Theta_1\)

Instructions

A
B
C
D

\(\Theta_1\)

\(\Theta_2\)

A Nominal Example

- Challenges for Analysis of ANES Data
- ANES Graphs
- Effect of Instructions:
  Constitutionality of Laws
ANES Graphs

Overview
Example Data Set
Existing Approaches
Log Multiplicative Association Models
Graphical Approach
Conditional Approach
Fighters, bullies and gender
Conclusions

A Nominal Example
● Challenges for Analysis of ANES Data

ANES Graph
● Effect of Instructions:
  Constitutionality of Laws

Challenges for Analysis of ANES Data

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D

Instructions

A
B
C
D
Effect of Instructions: Constitutionality of Laws

**Standard Instructions**

- Don’t Know
- Supreme Court
- Congress
- President

**Encourage Guessing**

- Congress
- Supreme Court
- President
- Don’t Know

A Nominal Example

- Challenges for Analysis of ANES Data
- ANES Graphs
- Effect of Instructions: Constitutionality of Laws


Goodman, L.A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-


29-3