Simultaneous Estimation of Multinomial Logistic Regression Models: Factor Analysis of Polytomous Item Response Data

Carolyn J. Anderson

Department of Educational Psychology

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Introduction

- Flexible latent variable models for multivariate, polytomous nominal and/or ordinal variables with collateral information.
Introduction

- **Flexible latent variable models** for multivariate, polytomous nominal and/or ordinal variables with collateral information.

- **Observed discrete (multicategory) variables**, e.g.,
  - Multiple choice test items.
  - Surveys items.
  - “Likert”.
  - Coded qualitative data.
Flexible **latent variable models** for multivariate, polytomous nominal and/or ordinal variables with collateral information.

**Observed discrete (multicategory) variables**, e.g.,
- Multiple choice test items.
- Surveys items.
- “Likert”.
- Coded qualitative data.

**Collateral Information** includes observed covariates (discrete or continuous) that are measures of individuals, items, response options, latent variables, as well as specific hypotheses about items and response options.
Introduction

- **Flexible latent variable models** for multivariate, polytomous nominal and/or ordinal variables with collateral information.

- Observed discrete (multicategory) variables, e.g.,
  - Multiple choice test items.
  - Surveys items.
  - “Likert”.
  - Coded qualitative data.

- **Collateral Information** includes observed covariates (discrete or continuous) that are measures of individuals, items, response options, latent variables, as well as specific hypotheses about items and response options.

- **(Multiple) Unobserved continuous variables.**
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%)

Challanges for Analysis of ANES Data

- Ordering of response the options unknown.
Challenges for Analysis of ANES Data

- Ordering of response the options unknown.
- Dealing with “Don’t Know”.
Challanges 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.
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.
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.”
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)
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?
Major Existing Approaches

to Latent Variable Modeling of Discrete Response Data:
- Quantify the data and then use factor analysis (or SEM) for continuous data.
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.
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 (e.g., Bock’s multinomial response model)
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 (e.g., Bock’s multinomial response model)
  - Multiple latent variables is a problem for standard estimation algorithms (i.e., numerical integration).
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 (e.g., Bock’s multinomial response model)
  - Multiple latent variables is a problem for standard estimation algorithms (i.e., numerical integration).

- Factor analysis of discrete data (Bartholomew, Steele, Moustaki & Galbraith, 2008)
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 (e.g., Bock’s multinomial response model)
  - 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 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)
Conditional Specification Approach

or the Item Response Theory Approach

1. Bock’s nominal response model, e.g.,

\[ P(Y_i = j_i | \theta) = \frac{\exp[\lambda_{ij_i} + \nu_{ij_i} \theta]}{\sum_{h_i} \exp[\lambda_{ih_i} + \nu_{ih_i} \theta]} . \]
Conditional Specification Approach

or the Item Response Theory Approach

1. Bock’s nominal response model, e.g.,

\[ P(Y_i = j_i | \theta) = \frac{\exp[\lambda_{ij_i} + \nu_{ij_i} \theta]}{\sum_h \exp[\lambda_{ih_i} + \nu_{ih_i} \theta]} . \]

2. If we knew or had estimates of the \( \nu_{ij_i} \)'s for all the items, then \( \sum_i \nu_{ij_i} \) would be sufficient for \( \theta \) (Andersen, 1995).
Conditional Specification Approach

or the Item Response Theory Approach

1. Bock’s nominal response model, e.g.,

\[
P(Y_i = j_i | \theta) = \frac{\exp[\lambda_{ij} + \nu_{ij} \theta]}{\sum_h \exp[\lambda_{ih} + \nu_{ih} \theta]}.
\]

2. If we knew or had estimates of the \( \nu_{ij} \)'s for all the items, then \( \sum_i \nu_{ij} \) would be sufficient for \( \theta \) (Andersen, 1995).

3. Suppose we know all the \( \nu \)'s except for item \( i \)'s, then we could replace \( \theta \) with a “rest-score”,

\[
\tilde{\theta}_{-i} = \phi \sum_{k \neq i} \nu_{kj}.
\]
Conditional Specification Approach

or the Item Response Theory Approach

1. Bock’s nominal response model, e.g.,

\[ P(Y_i = j_i | \theta) = \frac{\exp[\lambda_{ij_i} + \nu_{ij_i} \theta]}{\sum_{h_i} \exp[\lambda_{ih_i} + \nu_{ih_i} \theta]} . \]

2. If we knew or had estimates of the \( \nu_{ij_i} \)'s for all the items, then \( \sum_i \nu_{ij_i} \) would be sufficient for \( \theta \) (Andersen, 1995).

3. Suppose we know all the \( \nu \)'s except for item \( i \)'s, then we could replace \( \theta \) with a “rest-score”,

\[ \tilde{\theta}_{-i} = \phi \sum_{k \neq i} \nu_{kj_k} . \]

\[ \text{This gives us a multinomial logistic regression model where the predictor variable is a rest-score.} \]
Conditional Specification Approach

or the **Item Response Theory Approach**

1. Bock’s nominal response model, e.g.,

\[ P(Y_i = j_i | \theta) = \frac{\exp[\lambda_{ij_i} + \nu_{ij_i} \theta]}{\sum_h \exp[\lambda_{ih_i} + \nu_{ih_i} \theta]} . \]

2. If we knew or had estimates of the \( \nu_{ij_i} \)'s for all the items, then \( \sum_i \nu_{ij_i} \) would be sufficient for \( \theta \) (Andersen, 1995).

3. Suppose we know all the \( \nu \)'s except for item \( i \)'s, then we could replace \( \theta \) with a “rest-score”,

\[ \tilde{\theta}_{-i} = \phi \sum_{k \neq i} \nu_{kj_k} . \]

- This gives us a multinomial logistic regression model where the predictor variable is a rest-score.
- Justification and precedence comes from item response theory and classical test theory (Junker, 1993; Junker & Sijtsma, 2000).
Conditional Specification Approach: ANES

Laws: \[ P(A = j_1 | \tilde{\theta}_A) = \exp[\lambda^A_{j_1} + \nu^A_{j_1} \phi(\nu^B_{j_2} + \nu^C_{j_3} + \nu^D_{j_4})] \kappa_A \]

Judges: \[ P(B = j_2 | \tilde{\theta}_B) = \exp[\lambda^B_{j_2} + \nu^B_{j_2} \phi(\nu^A_{j_1} + \nu^C_{j_3} + \nu^D_{j_4})] \kappa_B \]

House: \[ P(C = j_3 | \tilde{\theta}_C) = \exp[\lambda^C_{j_3} + \nu^C_{j_3} \phi(\nu^A_{j_1} + \nu^B_{j_2} + \nu^D_{j_4})] \kappa_C \]

Senate: \[ P(D = j_4 | \tilde{\theta}_D) = \exp[\lambda^D_{j_4} + \nu^D_{j_4} \phi^2(\nu^A_{j_1} + \nu^B_{j_2} + \nu^C_{j_3})] \kappa_D \]

The \( \kappa \)'s ensure that probabilities over response options sum to 1 for each item.
Conditional Specification Approach: ANES

Laws: \[ P(A = j_1|\tilde{\theta}_A) = \exp[\lambda_{j_1}^A + \nu_{j_1}^A \phi(\nu_{j_2}^B + \nu_{j_3}^C + \nu_{j_4}^D)] \kappa_A \]

Judges: \[ P(B = j_2|\tilde{\theta}_B) = \exp[\lambda_{j_2}^B + \nu_{j_2}^B \phi(\nu_{j_1}^A + \nu_{j_3}^C + \nu_{j_4}^D)] \kappa_B \]

House: \[ P(C = j_3|\tilde{\theta}_C) = \exp[\lambda_{j_3}^C + \nu_{j_3}^C \phi(\nu_{j_1}^A + \nu_{j_2}^B + \nu_{j_4}^D)] \kappa_C \]

Senate: \[ P(D = j_4|\tilde{\theta}_D) = \exp[\lambda_{j_4}^D + \nu_{j_4}^D \phi^2(\nu_{j_1}^A + \nu_{j_2}^B + \nu_{j_3}^C)] \kappa_D \]

The \( \kappa \)'s ensure that probabilities over response options sum to 1 for each item.

Compatibility Conditions are needed to ensure this set is consistent with some joint distribution:
Conditional Specification Approach: ANES

Laws: $P(A = j_1 | \tilde{\theta}_A) = \exp[\lambda_{j_1}^A + \nu_{j_1}^A \phi(\nu_{j_2}^B + \nu_{j_3}^C + \nu_{j_4}^D)] \kappa_A$

Judges: $P(B = j_2 | \tilde{\theta}_B) = \exp[\lambda_{j_2}^B + \nu_{j_2}^B \phi(\nu_{j_1}^A + \nu_{j_3}^C + \nu_{j_4}^D)] \kappa_B$

House: $P(C = j_3 | \tilde{\theta}_C) = \exp[\lambda_{j_3}^C + \nu_{j_3}^C \phi(\nu_{j_1}^A + \nu_{j_2}^B + \nu_{j_4}^D)] \kappa_C$

Senate: $P(D = j_4 | \tilde{\theta}_D) = \exp[\lambda_{j_4}^D + \nu_{j_4}^D \phi^2(\nu_{j_1}^A + \nu_{j_2}^B + \nu_{j_3}^C)] \kappa_D$

The $\kappa$'s ensure that probabilities over response options sum to 1 for each item.

Compatibility Conditions are needed to ensure this set is consistent with some joint distribution:

When variable $i$ is the response variable, the term relating variable $i'$ to $i$ must be the same as the term relating $i$ to $i'$ when $i'$ is the response.

The proof is in paper for all models discussed in this talk.
The Joint Distribution

A set of multinomial logistic regression models that meet the compatibility conditions uniquely implies a log-multiplicative association model for the joint distribution.
The Joint Distribution

A set of multinomial logistic regression models that meet the compatibility conditions uniquely implies a log-multiplicative association model for the joint distribution.

For the ANES data:

\[
\log(P(y)) = \lambda + \lambda_{j_1}^A + \lambda_{j_2}^B + \lambda_{j_3}^C + \lambda_{j_4}^D
\]
\[
+ \phi \left( \nu_{j_1}^A \nu_{j_2}^B + \nu_{j_1}^A \nu_{j_3}^C + \nu_{j_1}^A \nu_{j_4}^D + \nu_{j_2}^B \nu_{j_3}^C + \nu_{j_2}^B \nu_{j_4}^D + \nu_{j_3}^C \nu_{j_4}^D \right).
\]

A log-multiplicative association model.
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
- Are generalizations of Goodman’s (1979, 1986) multidimensional row-column or “\(RC(M)\)” association models for 2-way tables.
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
- Are generalizations of Goodman’s (1979, 1986) multidimensional row-column or “\(RC(M)\)” association models for 2-way tables.
- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
- Are generalizations of Goodman’s (1979, 1986) multidimensional row-column or “$RC(M)$” association models for 2-way tables.
- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).
- Can be derived from a distance based model (de Rooij & Heiser, 2005), a generalization of Newton’s Law of Gravity.
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
- Are generalizations of Goodman’s (1979, 1986) multidimensional row-column or “RC(M)” association models for 2-way tables.
- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).
- Can be derived from a distance based model (de Rooij & Heiser, 2005), a generalization of Newton’s Law of Gravity.
- Can be derived from a Formative latent variable model using statistical graphical models (Anderson & Böckenholt, 2000; Anderson & Vermunt, 2000; Anderson 2002; Anderson & Tettagah, 2007).
Log-multiplicative Association Models

- Are special cases of Poisson regression models (i.e., log-linear models).
- Are generalizations of Goodman's (1979, 1986) multidimensional row-column or “RC(M)” association models for 2-way tables.
- Are implied by underlying multivariate normal distribution (Goodman, 1979; Becker, 1989; others).
- Can be derived from a distance based model (de Rooij & Heiser, 2005), a generalization of Newton's Law of Gravity.
- Can be derived from a **Formative** latent variable model using statistical graphical models (Anderson & Böckenholt, 2000; Anderson & Vermunt, 2000; Anderson 2002; Anderson & Tettagah, 2007).
- Can be derived from a **Reflective** latent variable model using standard item response theory methodology (Anderson & Yu, 2007; Anderson, Li, & Vermunt, 2007; Anderson, Verkuilen & Peyton, accepted).
Extensions

To be discussed here...

1. $M$ Multiple (correlated) latent variables.
Extensions

To be discussed here...

1. $\mathcal{M}$ Multiple (correlated) latent variables.

2. Restrictions on parameters
   - Equality
   - Scaling
   - Other
Extensions

To be discussed here...

1. $M$ Multiple (correlated) latent variables.

2. Restrictions on parameters
   - Equality
   - Scaling
   - Other

3. Adding covariates
   - Information about response options.
   - Information about instructions (treatment condition).
Multiple (correlated) latent variables

\[ \tilde{\theta}_{m,-i} = \phi_{mm} \sum_{k \neq i} \nu_{k,j} m + \sum_{m' \neq m} \phi_{m'm} \left( \sum_{k} \nu_{k,j} m' \right) \]
Multiple (correlated) latent variables

\[
\tilde{\theta}_{m,-i} = \phi_{mm} \sum_{k \neq i} \nu_{kjk} m + \sum_{m' \neq m} \phi_{m'm} m \left( \sum_{k} \nu_{kjk} m' \right)
\]

and the conditional multinomial model is

\[
P(Y_i = j_i | \tilde{\theta}_{m,-i}, \forall m) = \frac{\exp[\lambda_{ij_i} + \sum_m \nu_{ijim} \tilde{\theta}_{m,-i}]}{\sum_{h_i} \exp[\lambda_{ih_i} + \sum_m \nu_{ihim} \tilde{\theta}_{m,-i}]}.
\]
Multiple (correlated) latent variables

\[ \tilde{\theta}_{m,-i} = \phi_{mm} \sum_{k \neq i} \nu_{kj} \sigma_{m} + \sum_{m' \neq m} \phi_{m'm} \left( \sum_{k} \nu_{kj} \sigma_{m'} \right) \]

and the conditional multinomial model is

\[ P(Y_i = j_i | \tilde{\theta}_{m,-i}, \forall m) = \frac{\exp[\lambda_{ij} + \sum m \nu_{ij} \sigma_{m} \tilde{\theta}_{m,-i}]}{\sum h_i \exp[\lambda_{ih} + \sum m \nu_{ih} \sigma_{m} \tilde{\theta}_{m,-i}]} . \]

- If there is no direct relationship between an item and a latent variable, then the corresponding \( \nu_{ij} \sigma_{m} = 0 \).
Multiple (correlated) latent variables

\[ \tilde{\theta}_{m,-i} = \phi_{mm} \sum_{k \neq i} \nu_{kj}m + \sum_{m' \neq m} \phi_{m'm} \left( \sum_{k} \nu_{kj}m' \right) \]

and the conditional multinomial model is

\[ P(Y_i = j_i | \tilde{\theta}_{m,-i}, \forall m) = \frac{\exp[\lambda_{ij_i} + \sum_m \nu_{ij_i}m \tilde{\theta}_{m,-i}]}{\sum_{h_i} \exp[\lambda_{ih_i} + \sum_m \nu_{ih_i}m \tilde{\theta}_{m,-i}]} . \]

- If there is no direct relationship between an item and a latent variable, then the corresponding \( \nu_{ij_i}m = 0 \).

- The compatibility conditions are

\[ \phi_{mm'} \nu_{ij_i}m \nu_{kj_k}m' = \phi_{m'm} \nu_{kj_k}m' \nu_{ij_i}m \]

for all items, categories and latent variables.

i.e., \( \phi_{mm'} = \phi_{m'm} \).
The terms in the model include

1. A $\lambda$ to ensure that probabilities summed over all possible response patterns equal 1.
The terms in the model include

1. A $\lambda$ to ensure that probabilities summed over all possible response patterns equal 1.

2. All marginal effects terms $\lambda_{ij}$ (i.e., the intercepts in the multinomials).
Joint Distribution For M latent variables

The terms in the model include

1. A $\lambda$ to ensure that probabilities summed over all possible response patterns equal 1.

2. All marginal effects terms $\lambda_{ij}$ (i.e., the intercepts in the multinomials).

3. All multiplicative terms $\phi_{mm'}\nu_{ij}m\nu_{kj}k$ representing the relationship between pairs of items in the set of models defined by the multinomials.
Joint Distribution For M latent variables

The terms in the model include

1. A $\lambda$ to ensure that probabilities summed over all possible response patterns equal 1.

2. All marginal effects terms $\lambda_{ij}$ (i.e., the intercepts in the multinomials).

3. All multiplicative terms $\phi_{mm'}\nu_{ijm}\nu_{kk'm'}$ representing the relationship between pairs of items in the set of models defined by the multinomials.

Applying these rules yields

$$P(y) = \exp \left[ \lambda + \sum_i \lambda_{ij} + \sum_i \sum_{k>i} \sum_m \sum_{m'} \phi_{mm'}\nu_{ijm}\nu_{kk'm'} \right]$$
Restrictions on Parameters

- Equality and ordinal restrictions on category scores are straight forward.
Restrictions on Parameters

- Equality and ordinal restrictions on category scores are straightforward.

- For identification, need to set the
  - Location the $\lambda_{ij}$ and category scores $\nu_{ij}$, e.g.,

\[
\sum_{j_i} \lambda_{ij} = \sum_{j_i} \nu_{ij} = 0
\]
Restrictions on Parameters

- Equality and ordinal restrictions on category scores are straight forward.

- For identification, need to set the
  - Location the $\lambda_{ij}$ and category scores $\nu_{ij}$, e.g.,
    \[
    \sum_{j_i} \lambda_{ij} = \sum_{j_i} \nu_{ij} = 0
    \]
  - Scale of one set of $\nu_{ij}$ per latent variable, e.g.,
    \[
    \sum_{j_i} \nu_{ij}^2 = 1
    \]
Restrictions on Parameters

- Equality and ordinal restrictions on category scores are straightforward.

- For identification, need to set the
  - Location the $\lambda_{ij}$ and category scores $\nu_{ij}$, e.g.,
    \[ \sum_{j_i} \lambda_{ij} = \sum_{j_i} \nu_{ij} = 0 \]
  - Scale of one set of $\nu_{ij}$ per latent variable, e.g.,
    \[ \sum_{j_i} \nu_{ij}^2 = 1 \]
  - The category scale can be re-expressed as
    \[ \nu_{ij} = \omega_i \nu_{ij}^* \]
    where $\omega_i = \sqrt{\sum_{j_i} \nu_{ij}^2}$. 

- Extensions
  - Multiple (correlated) latent variables
  - Joint Distribution For M latent variables
  - Hybrid Item Response Models
  - Covariates

Analysis of the ANES Data

Conclusion & Beyond
Hybrid Item Response Models

- The category scale can be re-expressed as

\[ \nu_{ij} = \omega_i \nu_{ij}^* \]

where \( \omega_i = \sqrt{\sum_{j_i} \nu_{ij}^2} \).
Hybrid Item Response Models

- The category scale can be re-expressed as

\[ \nu_{ij} = \omega_i \nu_{ij}^* \]

where \( \omega_i = \sqrt{\sum_{j} \nu_{ij}^2} \).

- Set some (or all) \( \omega_i = 1 \) yields a “hybrid” item response model where
  - Each item is equally strongly related to the latent variable (Rasch-like).
Hybrid Item Response Models

- The category scale can be re-expressed as
  \[ \nu_{ij} = \omega_i \nu_{ij}^* \]

  where \( \omega_i = \sqrt{\sum_j \nu_{ij}^2} \).

- Set some (or all) \( \omega_i = 1 \) yields a “hybrid” item response model where
  - Each item is equally strongly related to the latent variable (Rasch-like).
  - The category scores \( \nu_{ij}^* \) carry category specific information (Bock NRM-like).
Hybrid Item Response Models

- The category scale can be re-expressed as
  \[ \nu_{ij} = \omega_i \nu^*_{ij} \]
  where \( \omega_i = \sqrt{\sum_j \nu^2_{ij}}. \)

- Set some (or all) \( \omega_i = 1 \) yields a “hybrid” item response model where
  - Each item is equally strongly related to the latent variable (Rasch-like).
  - The category scores \( \nu^*_{ij} \) carry category specific information (Bock NRM-like).

- The restriction of equal \( \omega_i(=1) \) for items can be imposed by placing scaling restriction on the category scores; that is,
  \[ \sum_{ji} \nu^2_{ij} = 1 \]

  on categories scores for more items than what is needed for identification.
Covariates can be added to the model by creating or modifying sub-models for the intercepts and/or latent variable(s).
Covariates can be added to the model by creating or modifying sub-models for the intercepts and/or latent variable(s).

Possible models for the intercept:

\[
\lambda_{ij} = \lambda^*_{ij} + \sum_{q} \beta_{ij} q x_q
\]

\[
\lambda_{ij} = \lambda^*_{ij} + \sum_{q} \beta_{ij} q x_{iq}
\]

\[
\lambda_{ij} = \lambda^*_{ij} + \sum_{q} \beta_q x_{ij} q
\]
Covariates

Covariates can be added to the model by creating or modifying sub-models for the intercepts and/or latent variable(s).

- Possible models for the intercept:

  \[ \lambda_{ij} = \lambda^{*}_{ij} + \sum_{q} \beta_{ijq} x_{iq} \]

- A possible model for a latent variable (i.e., covariate adds information about a person’s value on the latent variable):

  \[ \tilde{\theta}_{-i} = \phi \left( \sum_{k \neq i} \nu_{kjk} + \sum_{q} \eta_{q} x_{iq} \right) \]
Analysis of the ANES Data

1. Hypotheses regarding “Don’t know”:
   - Mondak (2001): “Don’t know” responders might have intermediate states of knowledge but be unwilling to venture a substantive response and thus be scored wrong even though they knew more than responders who answer incorrectly.
Analysis of the ANES Data

1. Hypotheses regarding “Don’t know”:
   - Mondak (2001): “Don’t know” responders might have intermediate states of knowledge but be unwilling to venture a substantive response and thus be scored wrong even though they knew more than responders who answer incorrectly.
   - “Don’t know” responders have less knowledge than those who respond with a substantive but incorrect answer.
1. Hypotheses regarding “Don’t know”:

- Mondak (2001): “Don’t know” responders might have intermediate states of knowledge but be unwilling to venture a substantive response and thus be scored wrong even though they knew more than responders who answer incorrectly.

- “Don’t know” responders have less knowledge than those who respond with a substantive but incorrect answer.

- “Don’t know” responses are simply not comparable to substantive responses, be they correct or incorrect.
Analysis of the ANES Data

1. Hypotheses regarding “Don’t know”:
   - Mondak (2001): “Don’t know” responders might have intermediate states of knowledge but be unwilling to venture a substantive response and thus be scored wrong even though they knew more than responders who answer incorrectly.
   - “Don’t know” responders have less knowledge than those who respond with a substantive but incorrect answer.
   - “Don’t know” responses are simply not comparable to substantive responses, be they correct or incorrect.

2. Instructions given to respondents would effect the way they responded. Treatment condition:
   - Standard or control group \( n = 606 \)
   - Instructions that encouraged respondents to guess if they were not sure of the correct answer \( n = 570 \).
Analysis of the ANES Data

1. Hypotheses regarding “Don’t know”:
   ■ Mondak (2001): “Don’t know” responders might have intermediate states of knowledge but be unwilling to venture a substantive response and thus be scored wrong even though they knew more than responders who answer incorrectly.
   ■ “Don’t know” responders have less knowledge than those who respond with a substantive but incorrect answer.
   ■ “Don’t know” responses are simply not comparable to substantive responses, be they correct or incorrect.

2. Instructions given to respondents would effect the way they responded. Treatment condition:
   ■ Standard or control group \((n = 606)\)
   ■ Instructions that encouraged respondents to guess if they were not sure of the correct answer \((n = 570)\).

3. Goal: Measure political knowledge.
Possible Graphs for the ANES Data
Possible Graphs for the ANES Data

Instructions

A

B

C

D

\[ \Theta_1 \]

Instructions

A

B

C

D

\[ \Theta_1 \]

\[ \Theta_2 \]
Possible Graphs for the ANES Data

Instructions

A

B

C

D

Θ₁

Θ₂

Θ₃

Instructions

A

B

C

D

Θ₁

Θ₂

Θ₃
## Analysis of the ANES Data

Fit statistics for models fit to the cross-classification of instructions and responses to the four ANES items.

<table>
<thead>
<tr>
<th>Model</th>
<th># par</th>
<th>$G^2$</th>
<th>BIC</th>
<th>Percent account</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log-linear Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independence</td>
<td>11</td>
<td>1533.02</td>
<td>−418.27</td>
<td>—</td>
</tr>
<tr>
<td>All 2-way</td>
<td>58</td>
<td>189.66</td>
<td>−1429.34</td>
<td>88%</td>
</tr>
<tr>
<td><strong>Log-multiplicative association models (LMA)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One dimension</td>
<td>31</td>
<td>418.22</td>
<td>−1391.67</td>
<td>73%</td>
</tr>
<tr>
<td>Two dimensions</td>
<td>32</td>
<td>312.88</td>
<td>−1489.94</td>
<td>80%</td>
</tr>
<tr>
<td>2D Hybrid</td>
<td>30</td>
<td>313.12</td>
<td>−1503.84</td>
<td>80%</td>
</tr>
<tr>
<td>2D Hybrid with $\nu_{11}^A = \nu_{21}^A$</td>
<td>29</td>
<td>316.79</td>
<td>−1507.23</td>
<td>79%</td>
</tr>
<tr>
<td>Three dimensional LMA</td>
<td>38</td>
<td>227.54</td>
<td>−1532.86</td>
<td>85%</td>
</tr>
<tr>
<td>3D Hybrid</td>
<td>36</td>
<td>229.69</td>
<td>−1544.84</td>
<td>85%</td>
</tr>
<tr>
<td>3D Hybrid w/ $\nu_{11}^A = \nu_{21}^A$</td>
<td>35</td>
<td>230.33</td>
<td>−1551.28</td>
<td>85%</td>
</tr>
<tr>
<td>3D Hybrid w/ $\nu_{11}^A = \nu_{21}^A$, $\nu_{j32}^C = \nu_{j32}^D$</td>
<td>34</td>
<td>230.34</td>
<td>−1558.33</td>
<td>85%</td>
</tr>
</tbody>
</table>
## Scoring and “Don’t Know”

### The estimated scale values/scores:

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Determines</td>
<td>President</td>
<td>$\nu_{11}^A$</td>
<td>$-0.036$</td>
</tr>
<tr>
<td></td>
<td>constitutionality</td>
<td>$\nu_{21}^A$</td>
<td>$-0.036$</td>
</tr>
<tr>
<td></td>
<td>Congress</td>
<td>$\nu_{21}^A$</td>
<td>$0.741$</td>
</tr>
<tr>
<td></td>
<td>Supreme Ct.</td>
<td>$\nu_{31}^A$</td>
<td>$-0.700$</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
<td>$\nu_{41}^A$</td>
<td>$-0.700$</td>
</tr>
<tr>
<td>B: Responsible for</td>
<td>President</td>
<td>$\nu_{11}^B$</td>
<td>$0.775$</td>
</tr>
<tr>
<td></td>
<td>nominating federal</td>
<td>Congress</td>
<td>$\nu_{21}^B$</td>
</tr>
<tr>
<td></td>
<td>judges</td>
<td>Supreme Ct.</td>
<td>$\nu_{31}^B$</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
<td>$\nu_{41}^B$</td>
<td>$-0.511$</td>
</tr>
<tr>
<td>C: Holds majority</td>
<td>Republican</td>
<td>$\nu_{12}^C$</td>
<td>$0.549$</td>
</tr>
<tr>
<td></td>
<td>in the House</td>
<td>Democrat</td>
<td>$\nu_{322}^C$</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
<td>$\nu_{32}^C$</td>
<td>$-0.798$</td>
</tr>
<tr>
<td>D: Holds majority</td>
<td>Republican</td>
<td>$\nu_{12}^D$</td>
<td>$0.625$</td>
</tr>
<tr>
<td></td>
<td>in the Senate</td>
<td>Democrat</td>
<td>$\nu_{22}^D$</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
<td>$\nu_{32}^D$</td>
<td>$-0.768$</td>
</tr>
</tbody>
</table>
Effect of Instructions: Constitutionality of Laws

Standard Instructions

Encourage Guessing

Don’t Know

Supreme Court

Congress

President

Don’t Know

Congress

Supreme Court

President

Don’t Know

Fitted Probability

Fitted Probability

Responsible, $\theta_1$

Responsible, $\theta_1$
Single Measure of Political Knowledge

and another look at the effect of Instructions

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>Control Group</th>
<th>Guessing Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary</td>
<td>.68 (.64, .72)</td>
<td>.58 (.53, .63)</td>
</tr>
<tr>
<td>LMA category scores</td>
<td>.75 (.68, .82)</td>
<td>.68 (.60, .75)</td>
</tr>
</tbody>
</table>
Single Measure of Political Knowledge

and another look at the effect of Instructions

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>Cronbach’s alpha (95% CI’s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Group</td>
<td>Guessing Group</td>
</tr>
<tr>
<td>Binary</td>
<td>.68 (.64, .72)</td>
</tr>
<tr>
<td>LMA category scores</td>
<td>.75 (.68, .82)</td>
</tr>
</tbody>
</table>

Comments:

- Scores from all the LMA models are all very similar.
Single Measure of Political Knowledge

and another look at the effect of Instructions

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>Control Group</th>
<th>Guessing Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary</td>
<td>.68 (.64, .72)</td>
<td>.58 (.53, .63)</td>
</tr>
<tr>
<td>LMA category scores</td>
<td>.75 (.68, .82)</td>
<td>.68 (.60, .75)</td>
</tr>
</tbody>
</table>

Comments:

- Scores from all the LMA models are all very similar.
- Model based scores lead to larger reliabilities.
Single Measure of Political Knowledge

and another look at the effect of Instructions

<table>
<thead>
<tr>
<th>Scoring Method</th>
<th>Cronbach’s alpha (95% CI’s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Group</td>
</tr>
<tr>
<td>Binary</td>
<td>.68 (.64, .72)</td>
</tr>
<tr>
<td>LMA category scores</td>
<td>.75 (.68, .82)</td>
</tr>
</tbody>
</table>

Comments:

- Scores from all the LMA models are all very similar.
- Model based scores lead to larger reliabilities.
- Less reliable measures when guessing is encouraged.
Conclusion & Beyond

- LMA models are effective tools for analysis and measurement of polytomous response data where latent variables may underlie responses.
Conclusion & Beyond

- LMA models are effective tools for analysis and measurement of polytomous response data where latent variables may underlie responses.

- **Estimation** that’s more flexible and capable of larger problems.
  - $\ell_{EM}$ can handle about $2^{12}$.
  - SAS PROC NLP can handle about $2^{20}$.
  - Pseudolikelihood of models in the Rasch family can handle at least $5^{100}$.
  - Experimental algorithm can handle at least $2^{100}$.
Conclusion & Beyond

- LMA models are effective tools for analysis and measurement of polytomous response data where latent variables may underlie responses.

- **Estimation** that’s more flexible and capable of larger problems.
  - $\ell_{\text{EM}}$ can handle about $2^{12}$.
  - SAS PROC NLP can handle about $2^{20}$.
  - Pseudolikelihood of models in the Rasch family can handle at least $5^{100}$.
  - Experimental algorithm can handle at least $2^{100}$.

- **Some Additional Possibilities:**
  - Other transformations of scale values, including ordinal and partial ordinal restrictions, spline, linear.
  - Higher level factors (i.e., lower rank decomposition of matrix of $\phi$’s).
This model did **not** fit the ANES data

\[ \phi_{11} = \ell_1^2, \quad \phi_{12} = \ell_1 \ell_2, \text{ and } \phi_{22} = \ell_2^2. \]