Polytomous Item Explanatory IRT Models with Random Item Effects: An Application to Carbon Cycle Assessment Data

Jinho Kim and Mark Wilson
University of California, Berkeley

Presented on April 11, 2018
at the 2018 IOMW, New York City, New York

Research Support and Disclaimer

The project described was supported by Grant Number 0815993 from the National Science Foundation (NSF). Any ideas, opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

OUTLINE

BEAR CENTER

1. Empirical Example & Motivation

2. Two Item Explanatory Extensions of the Partial Credit Model

3. Polytomous Item Explanatory Models with Random Item Errors

4. Empirical Study: An application to the Carbon Cycle assessment data

5. Conclusion and Discussion

Carbon TIME

NSF

RESEARCH GOALS

- How can we explain or predict the polytomous item difficulties by observed item properties?

- Considering the uncertainty in explanation or prediction, how can we add random item errors to account for residual variation in the polytomous item difficulties?

- How can we develop polytomous item explanatory models with random item errors and apply them to Carbon Cycle assessment data?

CONSTRUCTION OF CARBON CYCLE ITEMS

- **The item 01 (BODYTEMP) consists of cellular respiration process, energy progress variable, and multiple choice with explanation item format.**

Q1. [BODYTEMP] Your body produces **heat to maintain its normal temperature**. Where does **the heat mainly come from**? Please **choose the ONE answer** that you think is best **and explain why** you think that the answer you chose is better than the others. If you think some of the other answers are also partially right, please explain why you think so.

A. The heat mainly comes from sunlight.
B. The heat mainly comes from the clothes you are wearing.
C. The heat mainly comes from the foods you eat.
D. The heat mainly comes from your body when you are exercising.

CONSTRUCTION OF CARBON CYCLE ITEMS

- **The item 01 (BODYTEMP) consists of cellular respiration process, energy progress variable, and multiple choice with explanation item format.**

Q1. [BODYTEMP] Your body produces **heat to maintain its normal temperature**. Where does **the heat mainly come from**? Please **choose the ONE answer** that you think is best **and explain why** you think that the answer you chose is better than the others. If you think some of the other answers are also partially right, please explain why you think so.

A. The heat mainly comes from sunlight.
B. The heat mainly comes from the clothes you are wearing.
C. The heat mainly comes from the foods you eat.
D. The heat mainly comes from your body when you are exercising.

CONSTRUCTION OF CARBON CYCLE ITEMS

- **The item 01 (BODYTEMP) consists of cellular respiration process, energy progress variable, and multiple choice with explanation item format.**

Q1. [BODYTEMP] Your body produces **heat to maintain its normal temperature**. Where does **the heat mainly come from**? Please **choose the ONE answer** that you think is best **and explain why** you think that the answer you chose is better than the others. If you think some of the other answers are also partially right, please explain why you think so.

A. The heat mainly comes from sunlight.
B. The heat mainly comes from the clothes you are wearing.
C. The heat mainly comes from the foods you eat.
D. The heat mainly comes from your body when you are exercising.

CONSTRUCTION OF CARBON CYCLE ITEMS

- 13 items were polytomously scored with three achievement levels of ordered-categories based on the scoring rubric.
- The items can be reconstructed by combining the three item properties: (1) carbon cycling process, (2) progress variable, (3) item format

Item		Process			Progress				Format	
Num.	Name	Cellular Respiration (CR)	Photosyn -thesis (PS)	Digestion / Biosynthesis (DB)	Large Scale (LS)	Micro Scale (MS)	Energy (EN)	Mass (MA)	Multiple Choice (MC)	Yes.No Explain (YN)
01	BODYTEMP	1	0	0	0	0	1	0	1	0
02	CARBCYC2	1	0	0	1	0	0	0	0	1
03	CARBPLNT	0	1	0	0	1	0	0	0	1
~										
12	PLNTGRWTH	0	1	0	0	0	0	1	0	1
13	THINGTREE	0	1	0	0	0	0	1	0	1

MOTIVATION

- **LLTM+e approach as a random item explanatory IRT model**
  - The Linear Logistic Test Model (LLTM; Fischer, 1973) decomposes the difficulties of specific items into linear combinations of elementary components such as item properties.
  - The Linear Logistic Test Model with item error (LLTM+e; Janssen et al., 2004) can consider the uncertainty in explanation and enhance the precision of estimation of the item difficulties by accounting for residual variation.
- **LLTM+e approach to polytomous data**
  - Extensions of the LLTM+e approach to polytomous data is more complicated and demanding than to dichotomous data.
  - (1) item parameterization with item properties in the polytomous models is complicated,
  - (2) adding random item errors to the polytomous models is also complicated,
  - (3) allowing random effects on both item and person sides makes for crossed random effects, which is demanding for estimation.

Item Explanatory Extensions of the Partial Credit Model

1. **Partial Credit Model**
2. **Many-Facet Rasch Model (MFRM) with a Decomposition of the Item Location Parameters**
3. **Linear Partial Credit Model (LPCM) with a Decomposition of the Step Difficulty Parameters**

Partial Credit Model with two different item parameterizations

- **PCM with onefold item parameterization (original parameterization)**

$$ln\frac{Pr(y_{pi}=m|\theta_p)}{Pr(y_{pi}=m-1|\theta_p)}=\theta_p-\delta_{im}, \qquad m=1,\ldots,M_i, \delta_{i0}=0, \text{and } \theta_p \sim N(0,\sigma_\theta^2)$$

δ_{im} is a **step difficulty** parameter for scoring m rather than $m-1$ on item i

- **PCM with twofold item parameterization (ConQuest parameterization)**

$$ln\frac{Pr(y_{pi}=m|\theta_p)}{Pr(y_{pi}=m-1|\theta_p)}=\theta_p-(\beta_i+\tau_{im}), \qquad m=1,\ldots,M_i, \tau_{i0}=0, \text{and } \textstyle\sum_{m=1}^{M_i}\tau_{im}=0$$

β_i is an item location parameter for item *i*, interpreted as the **overall item difficulty**,
τ_{im} is a step deviation parameter for the *m*-th step within item *i*

Item Explanatory Extensions with two different item parameterizations

Polytomous Rasch Model	Partial Credit Model	
Item Parameterization	$\delta_{im} = \beta_i + \tau_{im}$ (two fold)	δ_{im} (one fold)
Target Item Difficulty	Overall Item Difficulty β_i	Step Difficulty δ_{im}
Relevant Item Explanatory Approach	Many-Facet Rasch Model (MFRM; Linacre, 1989) with a Decomposition of the Item Location Parameters	Linear Partial Credit Model (LPCM; Fischer and Ponocny, 1994) with a Decomposition of the Step Difficulty Parameters
Item Explanatory Model with Fixed Item Effects	$\delta'_{im} = \sum_{k=0}^{K} \gamma_k x_{ik} + \tau_{im}$ $\sum_{m=1}^{M_i} \tau_{im} = 0$	$\delta'_{im} = \sum_{k=0}^{K} \omega_{km} x_{ik}$
Item Property Effects	item specific property effects	item-by-step specific property effects

Polytomous Item Explanatory Models with Random Item Errors

1. Many-Facet Rasch Model (MFRM) with Overall Item Random Errors
2. Linear Partial Credit Model (LPCM) with Item-by-step Random Errors
3. Linear Partial Credit Model (LPCM) with Step Specific Random Errors

Polytomous Item Explanatory Models with Random Item Errors

Relevant Approach	Many-Facet Rasch Model (MFRM)	Linear Partial Credit Model (LPCM)
Target Item Difficulty	Overall Item Difficulty β_i	Step Difficulty δ_{im}
MFRM with Overall Item Random Errors	$\delta_{im} = \sum_{k=0}^{K} \gamma_k x_{ik} + \boldsymbol{\epsilon_i} + \tau_{im}$ $\epsilon_i \sim N(0, \sigma_\epsilon^2), \sum_{m=1}^{M_i} \tau_{im} = 0$	
LPCM with Item-by-step Random Errors		$\delta_{im} = \sum_{k=0}^{K} \omega_{km} x_{ik} + \boldsymbol{\varepsilon_{im}}$ $\varepsilon_{im} \sim N(0, \sigma_\varepsilon^2)$
LPCM with Step Specific Random Errors		$\delta_{im} = \sum_{k=0}^{K} \omega_{km} x_{ik} + \boldsymbol{\xi_{im}}$ $\boldsymbol{\xi_i}\ (\xi_{i1}, \dots, \xi_{im})' \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$

Empirical Study: An application to the Carbon Cycle assessment data

1. Many-Facet Rasch Model (MFRM) with Overall Item Random Errors
2. Linear Partial Credit Model (LPCM) with Item-by-step Random Errors
3. Linear Partial Credit Model (LPCM) with Step Specific Random Errors

Performance of the MFRM with Overall Item Random Errors

		PCM	MFRM (Fixed Item Effects)	MFRM with Overall Item Random Errors
Variance	σ_θ^2	1.18	0.80	1.15
	σ_ϵ^2			1.25
Goodness-of-fit	D	10933.10	11858.17	10947.72
	AIC	10987.10	11900.17	10991.72
	BIC	11123.55	12006.30	11102.90
	DIC	11747.50	12613.20	11757.00
Parameter	K	27	21	22

Residual variance for overall item difficulties

Correlation of δ_{im}	PCM	MFRM (Fixed Item Effects)	MFRM with Overall Item Random Errors
PCM	1	0.91	1

Item Property Effects in the MFRM with Overall Item Random Errors

- How does the item format property affect the overall item difficulties?

Item Property	Item Property Effects on the Overall Item Difficulties	Coefficient (Posterior Mean)	Posterior SD
	γ_0 (Intercept)	0.66	0.90
Process Property (Ref: Digestion/Biosynthesis)	γ_{CR} (Cellular Respiration)	0.60	1.04
	γ_{PS} (Photosynthesis)	-0.25	0.90
Progress Property (Ref: Mass)	γ_{LS} (Large Scale)	-1.74	1.45
	γ_{MS} (Micro Scale)	-0.79	0.91
	γ_{EN} (Energy)	-0.65	1.03
Item Format Property (Ref: Yes/No Explain)	γ_{MC} (Multiple Choice Explain)	0.28	1.09

Performance of the LPCM with Item-by-step Random Errors

		PCM	LPCM (Fixed Item Effects)	LPCM with Item-by-step Random Errors
Variance	σ_θ^2	1.18	0.79	1.17
	σ_ϵ^2			1.20
Goodness-of-fit	D	10933.10	12072.30	10937.86
	AIC	10987.10	12102.30	10969.86
	BIC	11123.55	12178.10	11050.72
	DIC	11747.50	12819.50	11749.30
Parameter	K	27	15	16

Residual variance for step difficulties

Correlation of δ_{im}	PCM	LPCM (Fixed Item Effects)	LPCM with Item-by-step Random Errors
PCM	1	0.91	1

Item Property Effects in the LPCM with Item-by-step Random Errors

- How does the item format property affect the step difficulties for each step?

Item Property	Item Property Effects on the Step Difficulties	Coefficient (Posterior Mean)	Posterior SD
	ω_{01}(Intercept for the 1st step)	-1.03	0.86
	ω_{02} (Intercept for the 2nd step)	2.22	0.84
Process Property (Ref: Digestion/Biosynthesis)	ω_{CR1} (Cellular Respiration for the 1st step)	0.87	1.04
	ω_{CR2} (Cellular Respiration for the 2nd step)	0.41	1.00
	ω_{PS1} (Photosynthesis for the 1st step)	0.38	0.90
	ω_{PS2} (Photosynthesis for the 2nd step)	-0.78	0.89
Progress Property (Ref: Mass)	ω_{LS1} (Large Scale for the 1st step)	-3.86	1.50
	ω_{LS2} (Large Scale for the 2nd step)	0.25	1.47
	ω_{MS1} (Micro Scale for the 1st step)	-0.89	0.84
	ω_{MS2} (Micro Scale for the 2nd step)	-0.58	0.86
	ω_{EN1} (Energy for the 1st step)	-0.58	1.01
	ω_{EN2} (Energy for the 2nd step)	-0.72	1.00
Item Format Property (Ref: Yes/No Explain)	ω_{MC1} (Multiple Choice for the 1st step)	0.66	1.09
	ω_{MC2} (Multiple Choice for the 2nd step)	-0.06	1.08

Performance of the LPCM with Step Specific Random Errors

		PCM	LPCM (Fixed Item Effects)	LPCM with Step Specific Random Errors
Variance	σ_θ^2	1.18	0.79	1.17
	$\sigma_{\delta 1}^2$			1.55
	$\sigma_{\delta 2}^2$			0.91
	$\sigma_{\delta 12}$			0.79
Goodness-of-fit	D	10933.10	12072.30	10937.64
	AIC	10987.10	12102.30	10973.64
	BIC	11123.55	12178.10	11064.60
	DIC	11747.50	12819.50	11748.90
Parameter	K	27	15	18

Residual variance for
- **the 1st step**
- **the 2nd step**
- **covariance**

Correlation of δ_{im}	PCM	LPCM (Fixed Item Effects)	LPCM with Step Specific Random Errors
PCM	1	0.91	1

Item Property Effects in the LPCM with Step Specific Random Errors

- How does the item format property affect the step difficulties for each step?

Item Property	Item Property Effects on the Step Difficulties	Coefficient (Posterior Mean)	Posterior SD
	ω_{01}(Intercept for the 1st step)	-0.97	0.95
	ω_{02} (Intercept for the 2nd step)	2.25	0.77
Process Property (Ref: Digestion/Biosynthesis)	ω_{CR1} (Cellular Respiration for the 1st step)	0.80	1.15
	ω_{CR2} (Cellular Respiration for the 2nd step)	0.35	0.89
	ω_{PS1} (Photosynthesis for the 1st step)	0.33	0.97
	ω_{PS2} (Photosynthesis for the 2nd step)	-0.80	0.78
Progress Property (Ref: Mass)	ω_{LS1} (Large Scale for the 1st step)	-3.81	1.69
	ω_{LS2} (Large Scale for the 2nd step)	0.40	1.27
	ω_{MS1} (Micro Scale for the 1st step)	-0.89	0.96
	ω_{MS2} (Micro Scale for the 2nd step)	-0.57	0.75
	ω_{EN1} (Energy for the 1st step)	-0.64	1.16
	ω_{EN2} (Energy for the 2nd step)	-0.75	0.88
Item Format Property (Ref: Yes/No Explain)	ω_{MC1} (Multiple Choice for the 1st step)	0.70	1.19
	ω_{MC2} (Multiple Choice for the 2nd step)	-0.05	0.92

CONCLUSION AND DISCUSSION

- **We found that**
  - Random item effect models performed better than fixed item effect models.
  - In the MFRM with overall item random errors, we could see how item properties have an effect on the overall item difficulties.
  - In the LPCM with item-by-step random errors and the LPCM with step specific random errors, we could see how item properties have an effect on the step difficulties for each step, but variance of the residuals was differently estimated between the two LPCM variations.

- **Potential value of the proposed models**
  - Can predict item locations or step difficulties of the newly developed items in a more scientific and systematic rather than intuitive manner.
  - Can enhance the precision of estimation of the polytomous item difficulties.
  - Can be a methodological foundation for advanced measurement models.

REFERENCE

Linacre, J. M. (1989). *Multi-facet Rasch measurement*. Chicago: MESA Press.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.

Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359-374.

Fischer, G. H., & Ponocny, I. (1994). An extension of the partial credit model with an application to the measurement of change. *Psychometrika*, 59(2), 177-192.

Janssen, R., Schepers, J., & Peres, D. (2004). Models with item and item group predictors. In De Boeck, P. & Wilson, M. (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 189-212). New York: Springer-Verlag.

Spiegelhalter, D., Thomas, A., & Best, N. (2003). *WinBUGS version 1.4* [Computer program]. UK: MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR.

Van den Noortgate, W., De Boeck, P., & Meulders, M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics*, 28(4), 369-386.

Thank you!

potatopaul@berkeley.edu