บทความวิชาการ (Academic Article)

โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ
MULTILEVEL MIXTURE ITEM RESPONSE THEORY MODEL

Received: July 7, 2020 **Revised:** August 4, 2020 **Accepted:** August 11, 2020

ธีรุตม์ สุขสกุลวัฒน์[1*] ณัฏฐภรณ์ หลาวทอง[2] และสิวะโชติ ศรีสุทธิยากร[3]
Teerut Suksakulwat[1*] Nuttaporn Lawthong[2] and Siwachoat Srisuttiyakorn[3]

[1,2,3]คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
[1,2,3]Faculty of Education, Chulalongkorn University, Bangkok 10330, Thailand

*Corresponding Author, E-mail: teerut1234@gmail.com

บทคัดย่อ

โมเดลทฤษฎีการตอบสนองข้อสอบ (IRT) มีสมมติฐานข้อหนึ่งว่าผู้สอบต้องมาจากประชากรที่มีภาวะเอกพันธุ์กันเพียงกลุ่มเดียว เพื่อแสดงนัยถึงความไม่แปรเปลี่ยนของการวัด อย่างไรก็ตามสมมติฐานดังกล่าวอาจเป็นการฝ่าฝืนธรรมชาติของข้อมูล เนื่องจากโดยทั่วไปผู้สอบในประชากรเดียวกันอาจมีคุณลักษณะที่แตกต่างกันได้ นอกจากนี้โมเดล IRT ยังมีการละเลยต่อโครงสร้างพหุระดับของข้อมูล ซึ่งอาจก่อให้เกิดข้อผิดพลาดในการสรุปผลอันเนื่องมาจากการรวมระดับ ด้วยเหตุนี้จึงมีการรวมโมเดล IRT เข้ากับโมเดลกลุ่มแฝงและโมเดลพหุระดับ เรียกว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (multilevel mixture IRT) เพื่อผ่อนคลายข้อจำกัดดังกล่าว

คำสำคัญ: โมเดลทฤษฎีการตอบสนองข้อสอบ กลุ่มแฝง ข้อมูลพหุระดับ

Abstract

Item response theory (IRT) model has an assumption that test takers must come from a single homogeneous population to imply the invariance of measurement. However, this assumption may violate the nature of the data, as in general, test takers in the same population may have different characteristics. In addition, the IRT model still neglects the multilevel structure of the data, which may cause less accurate results due to aggregation. For this reason, IRT model is integrated with latent class model and multilevel model, multilevel mixture IRT model, to relax the limitations.

Keywords: Item Response Theory Model, Latent Class, Multilevel Data

บทนำ

โมเดลทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นตัวแบบที่อธิบายถึงความสัมพันธ์ระหว่างผลการตอบซึ่งเป็นค่าที่สังเกตได้กับคุณลักษณะแฝงของผู้สอบและข้อสอบ โดยในช่วงเริ่มต้นโมเดล IRT ถูกนำมาใช้เกี่ยวกับการวัดผลทางการศึกษา (Lord & Novick, 1968, p. 13) ต่อมาจึงมีการนำไปประยุกต์ใช้กับศาสตร์แขนงอื่น เช่น จิตวิทยา การแพทย์ การประเมินบุคลิกภาพ จนปัจจุบันโมเดล IRT ได้รับการยอมรับโดยทั่วไปว่ามีข้อได้เปรียบเหนือกว่าโมเดลทฤษฎีการทดสอบแบบดั้งเดิม เนื่องจากโมเดล IRT มีข้อตกลงเบื้องต้นหลายประการ เช่น ความเป็นเอกมิติ (Unidimensionality) ความเป็นอิสระ (Local Independence) ความไม่แปรเปลี่ยน

(Invariance) และการเป็นฟังก์ชันต่อเนื่อง (Continuous Function) ที่สามารถอธิบายถึงความสัมพันธ์ระหว่างความน่าจะเป็นของการตอบถูกกับคุณลักษณะแฝง (Kanjanawasee, 2012, pp. 75-78; Reckase, 2009, pp. 12-14)

จากข้อตกลงเกี่ยวกับความไม่แปรเปลี่ยน โมเดล IRT มีสมมติฐานว่าผู้สอบทุกคนต้องมาจากประชากรที่มีภาวะเอกพันธุ์กันเพียงกลุ่มเดียว (Single Homogeneous Population) เพื่อแสดงนัยว่าโค้งลักษณะข้อสอบ (ICCs: Item Characteristic Curves) เพียงชุดเดียวสามารถอธิบายความสัมพันธ์ระหว่างผลการตอบกับคุณลักษณะแฝงที่เกี่ยวข้องได้ กล่าวคือ ICCs จะมีความคงที่ข้ามกลุ่มผู้สอบและข้ามชุดข้อสอบ ส่งผลให้พารามิเตอร์ของข้อสอบไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ และความสามารถของผู้สอบไม่แปรเปลี่ยนไปตามชุดข้อสอบ (Kanjanawasee, 2012, pp. 62-63) อย่างไรก็ตาม สมมติฐานดังกล่าวอาจเป็นการฝ่าฝืนธรรมชาติของข้อมูล เนื่องจากโดยทั่วไปผู้สอบในประชากรเดียวกันอาจจะมีคุณลักษณะที่แตกต่างกันได้ เช่น นักเรียนจากโรงเรียนเดียวกันอาจมีแบบแผนการตอบ (Response Pattern) ข้อสอบที่แตกต่างมาก จนอาจเชื่อได้ว่าผู้สอบน่าจะมาจากประชากรกลุ่มย่อย (Subpopulation) คนละกลุ่มมากกว่า (Tay et al., 2011, pp. 184-185)

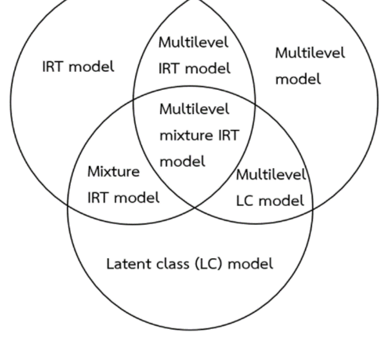
จากข้อจำกัดดังกล่าวจึงมีการบูรณาการรวมโมเดล IRT เข้ากับโมเดลกลุ่มแฝง (LC: Latent Class) เรียกว่า โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT: Mixture IRT) เพื่อผ่อนคลายปัญหาจากการฝ่าฝืนข้อตกลงความไม่แปรเปลี่ยน กล่าวคือ โมเดล MixIRT จะยอมให้พารามิเตอร์ในโมเดลมีค่าแตกต่างกันได้ตามกลุ่มแฝงหรือกลุ่มย่อยของประชากร (Rost, 1990, pp. 271-272)

นอกจากนี้ธรรมชาติของข้อมูลทางการศึกษา ยังพบว่าข้อมูลในระดับล่างมักซ้อน (Nested) อยู่ภายใต้ข้อมูลในระดับบน เช่น ข้อมูลนักเรียนซ้อนอยู่ภายใต้ห้องเรียน ห้องเรียนซ้อนอยู่ภายใต้โรงเรียน เป็นต้น เรียกว่าข้อมูลพหุระดับ (Multilevel Data) ทั้งนี้หากโมเดล MixIRT มีการละเลยโครงสร้างพหุระดับของข้อมูล ก็อาจส่งผลทำให้การตรวจสอบจำนวนกลุ่มแฝงของประชากรขาดความถูกต้อง รวมถึงการประมาณค่าพารามิเตอร์ยังอาจมีแนวโน้มคลาดเคลื่อนไปจากค่าจริง (Lee et al., 2018, pp. 151-152; Sen et al., 2019, pp. 273) ด้วยเหตุนี้จึงมีการบูรณาการรวมโมเดล MixIRT เข้ากับโมเดลพหุระดับ เรียกว่า โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT: Multilevel Mixture IRT) เพื่อผ่อนคลายข้อจำกัดที่เกิดขึ้น

จากการบูรณาการระหว่างโมเดล IRT โมเดล LC และโมเดลพหุระดับ ทำให้โมเดล MMixIRT มีคุณสมบัติในการให้สารสนเทศเชิงปริมาณเกี่ยวกับคุณลักษณะแฝงตามแนวคิดโมเดล IRT และให้สารสนเทศเชิงคุณภาพเกี่ยวกับกลุ่มแฝงตามแนวคิดโมเดล LC รวมถึงมีการคำนึงถึงโครงสร้างของข้อมูลตามแนวคิดโมเดลพหุระดับ โมเดล MMixIRT จึงสามารถให้สารสนเทศที่ครอบคลุมและสอดคล้องกับธรรมชาติของข้อมูลมากกว่าโมเดลทั่วไป โดยเฉพาะข้อมูลการทดสอบขนาดใหญ่หรือการทดสอบระดับชาติ ซึ่งมีผู้เข้าทดสอบเป็นนักเรียนที่มีคุณลักษณะหลากหลายจากโรงเรียนต่างๆ ทั่วประเทศ หรือจากประเทศต่างๆ ทั่วโลก

องค์ประกอบของโมเดล MMixIRT

องค์ประกอบของโมเดล MMixIRT สามารถแสดงได้ดังภาพ 1 และอาจนำไปเปรียบเทียบกับการวิเคราะห์ตัวแปรแฝงรูปแบบอื่นๆ ได้ดังตาราง 1



ภาพ 1 องค์ประกอบของโมเดล MMixIRT

ที่มา: Cho (2007, p. 18)

ตาราง 1 โมเดลการวิเคราะห์ตัวแปรแฝง

โครงสร้างข้อมูล	ตัวแปรแฝง	ตัวแปรสังเกตได้	
		ต่อเนื่อง (continuous)	จัดกลุ่ม (categorical)
ระดับเดียว (single-level)	ต่อเนื่อง (continuous)	การวิเคราะห์องค์ประกอบ (FA: factor analysis)	ทฤษฎีการตอบสนองข้อสอบ (IRT: item response theory)
	จัดกลุ่ม (categorical)	การวิเคราะห์โปรไฟล์แฝง (LP: latent profile analysis)	การวิเคราะห์กลุ่มแฝง (LC: latent class analysis)
	ต่อเนื่อง/จัดกลุ่ม (mixed)	การวิเคราะห์องค์ประกอบแบบผสม (MixFA: mixture FA)	ทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT: mixture IRT)
พหุระดับ (multilevel)	ต่อเนื่อง (continuous)	การวิเคราะห์องค์ประกอบพหุระดับ (multilevel FA)	ทฤษฎีการตอบสนองข้อสอบพหุระดับ (multilevel IRT)
	จัดกลุ่ม (categorical)	การวิเคราะห์โปรไฟล์แฝงพหุระดับ (multilevel LP)	การวิเคราะห์กลุ่มแฝงพหุระดับ (multilevel LC)
	ต่อเนื่อง/จัดกลุ่ม (mixed)	การวิเคราะห์องค์ประกอบแบบผสมพหุระดับ (MMixFA: multilevel mixture FA)	ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT: multilevel mixture IRT)

ที่มา: Lubke and Muthén (2005, p. 23); Tay et al. (2011, p. 182)

การบูรณาการองค์ความรู้ระหว่างโมเดล

1. โมเดลทฤษฎีการตอบสนองข้อสอบ (IRT)

โมเดล IRT เป็นโมเดลที่อธิบายถึงความสัมพันธ์ระหว่างโอกาสในการตอบถูกกับคุณลักษณะของข้อสอบและผู้สอบในรูปโค้งลักษณะข้อสอบ (ICCs) หรือฟังก์ชันการตอบสนองข้อสอบ ซึ่งนิยมเขียนอยู่ในรูปฟังก์ชันโลจิส (Kanjanawasee, 2012, pp. 53-54)

ข้อตกลงสำคัญของโมเดล IRT มี 3 ประการ คือ 1) ความเป็นเอกมิติ หมายถึงข้อสอบทุกข้อมุ่งวัดเพียงความสามารถเดียว 2) ความเป็นอิสระ กล่าวคือผลการตอบข้อสอบข้อเดียวกันของผู้สอบแต่ละคนจะเป็นอิสระจากกัน (ความเป็นอิสระระหว่างผู้สอบ) และผลการตอบข้อสอบแต่ละข้อของผู้สอบคนเดียวกันจะเป็นอิสระจากกัน (ความเป็นอิสระระหว่างข้อสอบ) ด้วยคุณสมบัติดังกล่าว พารามิเตอร์ของผู้สอบและข้อสอบจึงไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบหรือชุดข้อสอบ และ 3) การเป็นฟังก์ชันต่อเนื่องที่สามารถอธิบายความสัมพันธ์ระหว่างคุณลักษณะแฝงกับความน่าจะเป็นในการตอบข้อสอบได้ถูก (Kanjanawasee, 2012, pp. 75-78)

ในบทความนี้จะขอนำเสนอโมเดล IRT ในรูปฟังก์ชันโลจิสที่มีการตรวจให้คะแนนได้ 2 ค่า (Dichotomous Scoring) เท่านั้น เนื่องจากแบบสอบมาตรฐานที่นิยมใช้ทั่วไปมักเป็นข้อสอบปรนัยแบบหลายตัวเลือกที่มีการตรวจให้คะแนนเป็นถูกกับผิด รายละเอียดของโมเดลมีดังนี้

1.1 โมเดลโลจิส 1 พารามิเตอร์ (1PL) หรือโมเดล Rasch

โมเดล Rasch มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (1) ซึ่งแสดงถึงความน่าจะเป็นในการตอบข้อสอบ i ได้ถูกต้องของผู้สอบ j ($y_{ij}=1$ คือ ตอบถูก; $y_{ij}=0$ คือ ตอบผิด)

$$P(y_{ij}=1\,|\,\theta_j) = \frac{1}{1+\exp\left[-(\theta_j-b_i)\right]} \qquad (1)$$

เมื่อ θ_j คือ พารามิเตอร์ความสามารถของผู้สอบ j ภายใต้สมมติฐานว่าผู้สอบทุกคนมาจากประชากรเพียงกลุ่มเดียว โดย $\theta_j \sim N(0,1)$ และ b_i คือพารามิเตอร์ความยากของข้อสอบ i

1.2 โมเดลโลจิส 2 พารามิเตอร์ (2PL)

โมเดล 2PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (2)

$$P(y_{ij}=1\,|\,\theta_j) = \frac{1}{1+\exp\left[-a_i(\theta_j-b_i)\right]} \quad (2)$$

เมื่อ a_i คือ พารามิเตอร์อำนาจจำแนกของข้อสอบ i

1.3 โมเดลโลจิส 3 พารามิเตอร์ (3PL)

โมเดล 3PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (3)

$$P(y_{ij}=1\,|\,\theta_j) = c_i + (1-c_i)\frac{1}{1+\exp\left[-a_i(\theta_j-b_i)\right]} \quad (3)$$

เมื่อ c_i คือ พารามิเตอร์การเดาของข้อสอบ i

2. โมเดลกลุ่มแฝง (LC)

โมเดล LC มีแนวคิดคล้ายกับการวิเคราะห์องค์ประกอบ (FA: Factor Analysis) แต่มีจุดต่างกันตรงที่ตัวแปรแฝงในโมเดล FA เป็นตัวแปรต่อเนื่อง ส่วนตัวแปรแฝงในโมเดล LC เป็นตัวแปรจัดกลุ่ม จึงนิยมเรียกว่ากลุ่มแฝง หรืออาจมองความแตกต่างในแง่ของการจัดกลุ่มได้ว่าโมเดล FA เป็นการจัดกลุ่มให้กับตัวแปร (Variables) ขณะที่โมเดล LC เป็นการจัดกลุ่มให้กับบุคคล (Persons)

เป้าหมายของโมเดล LC คือการจัดกลุ่มให้กับบุคคลที่มีแบบแผนการตอบคล้ายคลึงกัน เพื่อจำแนกบุคคลออกเป็นกลุ่มย่อย ผลการวิเคราะห์จะให้สารสนเทศเกี่ยวกับจำนวนกลุ่มแฝงในประชากร สัดส่วนของสมาชิกในกลุ่มแฝง และอาจรวมถึงปัจจัยที่ใช้ในการจำแนกกลุ่ม โดยมีข้อตกลงสำคัญ 2 ประการ คือ 1) ความเป็นอิสระ เช่นเดียวกับโมเดล IRT และ 2) สมาชิกในกลุ่มแฝงต้องมาจากประชากรกลุ่มเดียวกัน กล่าวคือผลรวมของสัดส่วนของสมาชิกในแต่ละกลุ่มแฝงต้องมีค่าเป็น 1 หรือ $\sum_{k=1}^{K}\pi_k = 1$ เมื่อ π_k คือ สัดส่วนของสมาชิกในกลุ่มแฝง k และ $0 < \pi_k < 1$

รูปแบบทั่วไปของโมเดล LC (Tay et al., 2011, pp. 180) สามารถแสดงได้ดังสมการ (4)

$$P(\underset{\sim}{y}_j) = \sum_{k=1}^{K}\pi_k P(\underset{\sim}{y}_j\,|\,k) \quad (4)$$

เมื่อ $\underset{\sim}{y}_j$ คือ เวกเตอร์แบบแผนการตอบข้อสอบ I ข้อ ของผู้สอบ j ; $\underset{\sim}{y}_j = \left(y_{1j}\ \ y_{2j}\ \ldots\ y_{Ij}\right)$ และ π_k คือ สัดส่วนของสมาชิกในกลุ่มแฝง k และ $P(\underset{\sim}{y}_j\,|\,k)$ คือ ความน่าจะเป็นแบบมีเงื่อนไขของแบบแผนการตอบของผู้สอบ j ซึ่งอยู่ในกลุ่มแฝง k เมื่อ $P(\underset{\sim}{y}_j\,|\,k) = \prod_{i=1}^{I}P(y_{ij}\,|\,k)$

3. โมเดลพหุระดับ (Multilevel)

โมเดลพหุระดับหรือโมเดลเชิงเส้นแบบลดหลั่น (Hierarchical Linear Model) เป็นโมเดลการวิเคราะห์ที่มีการแบ่งความแปรปรวนของความคลาดเคลื่อนในระดับต่างๆ ออกจากกัน ถูกพัฒนาขึ้นเพื่อใช้แทนโมเดลถดถอยทั่วไป ซึ่งเป็นการวิเคราะห์ในระดับเดียว (Single-Level Analysis) ที่มักฝ่าฝืนข้อตกลงเกี่ยวกับความเป็นอิสระของตัวอย่างเมื่อนำไปวิเคราะห์กับข้อมูลทางการศึกษา (Raudenbush & Bryk, 2002, pp. 3-5) จึงอาจเกิดข้อจำกัด คือ 1) ผลการประมาณค่าพารามิเตอร์อาจคลาดเคลื่อนไปจากค่าจริง เนื่องจากการวิเคราะห์แบบรวมหน่วยตัวอย่างทุกกลุ่มเข้าด้วยกัน มักก่อให้เกิดข้อผิดพลาดในการสรุปผลระหว่างระดับ (Aggregation Bias) และ 2) ผลการวิเคราะห์อาจถูกรบกวนจากตัวแปรในระดับกลุ่ม ทำให้การประมาณค่าความคลาดเคลื่อนมาตรฐานของค่าพารามิเตอร์ในโมเดลมีแนวโน้มต่ำกว่าค่าจริง (Wongnam, 2012, pp. 27-28)

ในบทความนี้ขอนำเสนอเฉพาะโมเดลพหุระดับที่ไม่มีตัวแปรทำนายใดๆ หรือโมเดลปราศจากเงื่อนไขอย่างสมบูรณ์ (Fully Unconditional Model) ที่มี 2 ระดับคือ ระดับหน่วย (Micro-Level) เช่น ระดับนักเรียน และระดับกลุ่ม (Macro-Level) เช่น ระดับโรงเรียน เพื่อจะนำแนวคิดดังกล่าวไปบูรณาการเข้ากับโมเดล IRT และโมเดล LC ต่อไป

โมเดลระดับที่ 1 หรือโมเดลระดับนักเรียน (student-level model) เป็นโมเดลแสดงความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรต้นในระดับนักเรียน ดังสมการ (5)

$$y_{js} = \beta_{0s} + r_{js} \quad (5)$$

เมื่อ y_{js} คือ ตัวแปรตามซึ่งเป็นค่าสังเกตของนักเรียน j ในโรงเรียน s และ r_{js} คือ ค่าความคลาดเคลื่อนในระดับนักเรียน เมื่อ $r_{js} \sim N(0,\sigma^2)$; σ^2 คือ ความแปรปรวนภายในกลุ่ม และ β_{0s} คือค่าเฉลี่ยของค่าสังเกตของนักเรียนในโรงเรียน s ซึ่งถือว่าเป็นพารามิเตอร์สุ่ม (Random Parameter) ที่ต้องมีการกำหนดโมเดลเพิ่มเติมเพื่ออธิบายความแปรปรวนที่เกิดขึ้นต่อไป เรียกโมเดลดังกล่าวว่าโมเดลระดับที่ 2 หรือโมเดลระดับโรงเรียน (School-Level Model) ดังสมการ (6)

$$\beta_{0s} = \gamma_{00} + u_{0s} \quad (6)$$

เมื่อ γ_{00} คือ ค่าเฉลี่ยของค่าสังเกตของนักเรียนทุกโรงเรียน และ u_{0s} คือ ค่าความคลาดเคลื่อนในระดับโรงเรียน เมื่อ $u_{0s} \sim N(0,\tau)$; τ คือ ความแปรปรวนระหว่างกลุ่ม ทั้งนี้เมื่อนำสมการ (5) และ (6) มาเขียนรวมกันจะเรียกว่าโมเดลอิทธิพลผสม (Mixed-Effect Model) ดังสมการ (7)

$$y_{js} = \gamma_{00} + r_{js} + u_{0s} \quad (7)$$

โดยอาจเรียก γ_{00} ว่าอิทธิพลคงที่ และเรียก r_{js} และ u_{0s} ว่าอิทธิพลสุ่ม

การวิเคราะห์ด้วยโมเดลพหุระดับข้างต้นจะให้สารสนเทศว่า ข้อมูลมีความเหมาะสมสำหรับการนำมาวิเคราะห์แบบพหุระดับหรือไม่ ด้วยการพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC : intraclass correlation coefficient) ซึ่งมีสูตรคำนวณดังสมการ (8)

$$ICC = \frac{\tau}{\tau + \sigma^2} \quad (8)$$

เมื่อ ICC คือ ค่าสถิติที่บอกว่าความแปรปรวนทั้งหมดของข้อมูลสามารถถูกอธิบายด้วยข้อมูลระดับโรงเรียนเพียงใด หาก ICC > .05 จะถือว่าข้อมูลมีความเหมาะสม (Hox & Maas, 2001, p. 157)

4. โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT)

เนื่องจากโมเดล IRT ทั่วไปมีสมมติฐานว่าผู้สอบมาจากประชากรที่มีภาวะเอกพันธุ์กันเพียงกลุ่มเดียว อย่างไรก็ตาม สมมติฐานดังกล่าวอาจไม่สอดคล้องกับธรรมชาติของข้อมูลเสมอไป เพราะแม้ผู้สอบจะมาจากประชากรกลุ่มเดียวกัน แต่ก็อาจมีคุณลักษณะแตกต่างกันได้ (Tay et al., 2011, pp. 184-185) จากปัญหาดังกล่าวจึงมีการบูรณาการรวมโมเดล IRT เข้ากับโมเดล LC กลายเป็นโมเดล MixIRT เพื่อผ่อนคลายข้อตกลงเกี่ยวกับความไม่แปรเปลี่ยนลง พารามิเตอร์ในโมเดล MixIRT จึงมีค่าแตกต่างกันได้ตามกลุ่มแฝงหรือกลุ่มย่อยของประชากร (Rost, 1990, pp. 271-272)

เป้าหมายของโมเดล MixIRT จึงเป็นการจัดกลุ่มแฝงให้กับบุคคลที่มีแบบแผนการตอบคล้ายคลึงกันเช่นเดียวกับโมเดล LC แต่มีจุดแตกต่างตรงที่โมเดล MixIRT มีการให้สารสนเทศเชิงปริมาณเกี่ยวกับคุณลักษณะแฝงของผู้สอบและข้อสอบตามแนวคิดของโมเดล IRT ด้วย ทั้งนี้ การแบ่งรูปแบบของโมเดล MixIRT สามารถพิจารณาได้ในทำนองเดียวกับโมเดล IRT ดังนี้

4.1 โมเดล mixture Rasch (MixRasch)

โมเดล MixRasch เป็นการบูรณาการระหว่างโมเดล Rasch กับโมเดล LC โดยมีข้อตกลงว่า 1) สมาชิกในแต่ละกลุ่มแฝงต้องมาจากประชากรกลุ่มเดียวกัน เช่นเดียวกับโมเดล LC และ 2) โมเดล Rasch มีความสอดคล้องกับข้อมูลในกลุ่มแฝง โดยในกลุ่มแฝงแต่ละกลุ่มจะมีพารามิเตอร์ความยากของข้อสอบเพียงชุดเดียว เพื่อแสดงนัยว่าสมาชิกในกลุ่มแฝงเดียวกันมีแบบแผนการตอบที่คล้ายคลึงกัน (Cohen et al., 2002, pp. 4-6)

โมเดล MixRasch (Rost, 1990, p. 272) สามารถแสดงได้ดังสมการ (9)

$$P(y_{ij}=1 \mid \theta_{jk}) = \sum_{k=1}^{K} \pi_k \frac{1}{1 + \exp\left[-(\theta_{jk} - b_{ik})\right]} \quad (9)$$

เมื่อ π_k คือ สัดส่วนของสมาชิกในกลุ่มแฝง k และ θ_{jk} คือ พารามิเตอร์ความสามารถของผู้สอบ j ในกลุ่มแฝง k เมื่อ $\theta_{jk} \sim N(\mu_k, \sigma_k^2)$; μ_k และ σ_k^2 คือ ค่าเฉลี่ยและความแปรปรวนของความสามารถในกลุ่มแฝง k และ b_{ik} คือ พารามิเตอร์ความยากของข้อสอบ i ในกลุ่มแฝง k

4.2 โมเดล mixture 2PL (Mix2PL)

โมเดล Mix2PL ได้รับการขยายแนวคิดมาจากโมเดล MixRasch ด้วยการผ่อนคลายให้พารามิเตอร์อำนาจจำแนกของข้อสอบมีค่าแตกต่างกันได้ตามกลุ่มแฝง ภายใต้ข้อตกลงว่าโมเดล 2PL มีความสอดคล้องกับข้อมูลในกลุ่มแฝง โมเดล Mix2PL จึงมีรูปแบบดังสมการ (10)

$$P(y_{ij} = 1 \big| \theta_{jk}) = \sum_{k=1}^{K} \pi_k \frac{1}{1 + \exp\left[-a_{ik}(\theta_{jk} - b_{ik})\right]} \qquad (10)$$

เมื่อ a_{ik} คือ พารามิเตอร์อำนาจจำแนกของข้อสอบ i ในกลุ่มแฝง k

4.3 โมเดล mixture 3PL (Mix3PL)

โมเดล Mix3PL ได้รับการขยายแนวคิดมาจากโมเดล Mix2PL ด้วยการเพิ่มพารามิเตอร์การเดาให้มีค่าแตกต่างกันได้ตามกลุ่มแฝง ภายใต้ข้อตกลงว่าโมเดล 3PL มีความสอดคล้องกับข้อมูลในกลุ่มแฝง โมเดล Mix3PL จึงมีรูปแบบดังสมการ (11)

$$P(y_{ij} = 1 \big| \theta_{jk}) = \sum_{k=1}^{K} \pi_k \left(c_{ik} + (1 - c_{ik}) \frac{1}{1 + \exp\left[-a_{ik}(\theta_{jk} - b_{ik})\right]} \right) \qquad (11)$$

เมื่อ c_{ik} คือ พารามิเตอร์การเดาของข้อสอบ i ในกลุ่มแฝง k

5. โมเดลทฤษฎีการตอบสนองข้อสอบพหุระดับ (multilevel IRT)

จากที่กล่าวไปในตอนต้นแล้วว่า ธรรมชาติของข้อมูลทางการศึกษามักเป็นข้อมูลพหุระดับ แต่เนื่องจากโมเดล IRT ทั่วไปไม่มีการคำนึงถึงโครงสร้างของข้อมูล ผลการวัดในระดับหน่วย (ระดับนักเรียน) จึงอาจถูกรบกวนจากตัวแปรในระดับกลุ่ม (ระดับโรงเรียน) ทำให้เกิดความคลาดเคลื่อนขึ้นในการประมาณค่า ด้วยเหตุนี้โมเดล multilevel IRT จึงถูกพัฒนาเพื่อผ่อนคลายปัญหาดังกล่าว

แนวคิดการพัฒนาโมเดล multilevel IRT มีผู้เสนอไว้หลายวิธี เช่น Fox and Glas (2001); Kamata (2001) และ Maier (2001) อย่างไรก็ตาม ในบทความนี้ขอยกตัวอย่างแนวคิดการพัฒนาโมเดล multilevel IRT แบบ 3 ระดับ ของ Kamata (2001, pp. 85-86) เท่านั้น รายละเอียดดังนี้

5.1 โมเดล multilevel Rasch

โมเดล multilevel Rasch จะมีการเพิ่มโมเดล Rasch เข้ามาในโมเดลพหุระดับอีกหนึ่งระดับ เรียกโมเดลนี้ว่าโมเดลระดับข้อสอบ (Item-Level Model) ดังสมการ (12)

$$P(y_{ijs} = 1 \big| \theta_{js}) = \frac{1}{1 + \exp\left[-(\theta_{js} - b_{ijs})\right]} \qquad (12)$$

เมื่อ θ_{js} คือ พารามิเตอร์ความสามารถของนักเรียน j ในโรงเรียน s และ b_{ijs} คือ พารามิเตอร์ความยากของข้อสอบ i ซึ่งมีค่าขึ้นอยู่กับหน่วยตัวอย่างในระดับนักเรียนและโรงเรียน ทั้งนี้ θ_{js} และ b_{ijs} จะถือว่าเป็นพารามิเตอร์สุ่ม ซึ่งต้องมีการกำหนดโมเดลเพื่ออธิบายความแปรปรวนที่เกิดขึ้นต่อไป เรียกว่าโมเดลระดับนักเรียน ดังสมการ (13)

$$\theta_{js} = \gamma_{0s} + r_{js} \text{ และ } b_{ijs} = b_{is} \quad (13)$$

เมื่อ γ_{0s} คือ ความสามารถเฉลี่ยของนักเรียนในโรงเรียน s และ r_{js} คือ ค่าความคลาดเคลื่อนในระดับนักเรียน เมื่อ $r_{js} \sim N(0, \tau)$; τ คือ ความแปรปรวนของความสามารถในระดับนักเรียน ส่วน b_{ijs} จะกลายเป็น b_{is} เนื่องจากพารามิเตอร์ความยากเป็นคุณลักษณะของข้อสอบ จึงเป็นค่าคงที่สำหรับนักเรียนทุกคน ทั้งนี้ γ_{0s} และ b_{is} จะถือว่าเป็นพารามิเตอร์สุ่ม ซึ่งต้องมีการกำหนดโมเดลเพื่ออธิบายความแปรปรวนที่เกิดขึ้นต่อไป เรียกว่าโมเดลระดับโรงเรียน ดังสมการ (14)

$$\gamma_{0s} = \gamma_{00} + u_{0s} \text{ และ } b_{is} = b_i \quad (14)$$

เมื่อ γ_{00} คือ ความสามารถเฉลี่ยของนักเรียนทุกโรงเรียน และ u_{0s} คือ ค่าความคลาดเคลื่อนในระดับโรงเรียน เมื่อ $u_{0s} \sim N(0, \zeta)$; ζ คือ ความแปรปรวนของความสามารถในระดับโรงเรียน ส่วน b_{is} จะกลายเป็น b_i เนื่องจากพารามิเตอร์

ความยากเป็นคุณลักษณะของข้อสอบ จึงเป็นค่าคงที่สำหรับทุกโรงเรียน ทั้งนี้ อาจพิจารณาให้ $\gamma_{0s} = u_{0s}$ เมื่อ $u_{0s} \sim N(\gamma_{00}, \zeta)$ และเมื่อนำสมการ (12) - (14) มาเขียนเป็นโมเดลอิทธิพลผสมโมเดล multilevel Rasch จึงมีรูปแบบดังสมการ (15)

$$P(y_{ijs} = 1 \big| \theta_{js}) = \frac{1}{1 + \exp\left[-\left((u_{0s} + r_{js}) - b_i\right)\right]} \quad (15)$$

เมื่อ u_{0s} คือ ค่าความคลาดเคลื่อนในระดับโรงเรียน เมื่อ $u_{0s} \sim N(\gamma_{00}, \zeta)$ และ r_{js} คือ ค่าความคลาดเคลื่อนในระดับนักเรียน เมื่อ $r_{js} \sim N(0, \tau)$

5.2 โมเดล multilevel 2PL

โมเดล multilevel 2PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (16)

$$P(y_{ijs} = 1 \big| \theta_{js}) = \frac{1}{1 + \exp\left[-a_i\left((u_{0s} + r_{js}) - b_i\right)\right]} \quad (16)$$

5.3 โมเดล multilevel 3PL

โมเดล multilevel 3PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (17)

$$P(y_{ijs} = 1 \big| \theta_{js}) = c_i + (1 - c_i)\frac{1}{1 + \exp\left[-a_i\left((u_{0s} + r_{js}) - b_i\right)\right]} \quad (17)$$

ส่วน ICC สำหรับโมเดล multilevel IRT มีสูตรคำนวณดังสมการ (18)

$$ICC = \frac{\zeta}{\zeta + \tau} \quad (18)$$

เมื่อ τ และ ζ คือ ความแปรปรวนของความสามารถในระดับนักเรียนและระดับโรงเรียนตามลำดับ ส่วนความแปรปรวนของความคลาดเคลื่อนในระดับข้อสอบจะไม่ถูกนำมาพิจารณา เนื่องจากค่าดังกล่าวได้รวมอยู่ในโมเดลการวัดคุณลักษณะแฝงแล้ว (Fox & Glas, 2001, pp. 271-272)

6. โมเดลกลุ่มแฝงพหุระดับ (multilevel LC)

ข้อจำกัดหนึ่งของโมเดล LC คือ มีการละเลยโครงสร้างพหุระดับของข้อมูล ด้วยเหตุนี้โมเดล multilevel LC (Vermunt, 2003, pp. 216-217; Cho, 2007, p. 29) จึงถูกพัฒนาขึ้นเพื่อผ่อนคลายข้อจำกัดดังกล่าว โดยสามารถเขียนเป็นรูปแบบทั่วไปได้ดังสมการ (19)

$$P(\underset{\sim}{y}_s) = \sum_{g=1}^{G}\pi_g\left[\prod_{j=1}^{n_s}\left\{\sum_{k=1}^{K}\pi_{k|g}\left(\prod_{i=1}^{I}P(y_{ijs}\big|k, g)\right)\right\}\right] \quad (19)$$

เมื่อ $\underset{\sim}{y}_s$ คือ เวกเตอร์แบบแผนการตอบข้อสอบ I ข้อ ของนักเรียน n_s คน ในโรงเรียน s และ π_g คือ สัดส่วนของสมาชิกในกลุ่มแฝงระดับโรงเรียน g และ $\pi_{k|g}$ คือ สัดส่วนของสมาชิกในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g และ $P(y_{ijs}\big|k, g)$ คือ ความน่าจะเป็นแบบมีเงื่อนไขของผลการตอบข้อสอบ i ของนักเรียน j ในโรงเรียน s ซึ่งอยู่ในกลุ่มแฝงต่าง ๆ

7. โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT)

จากการบูรณาการองค์ความรู้ระหว่างโมเดลที่มีความแตกต่างกันจำนวนสองโมเดล ทำให้เกิดโมเดลใหม่ ได้แก่ 1) โมเดล MixIRT 2) โมเดล multilevel IRT และ 3) โมเดล multilevel LC อย่างไรก็ตามโมเดลเหล่านี้ยังอาจมีข้อจำกัดเกี่ยวกับการวิเคราะห์บางประการ ดังนี้

ข้อจำกัดของโมเดล MixIRT คือ โมเดลมีการละเลยต่อโครงสร้างพหุระดับของข้อมูล การตรวจสอบจำนวนกลุ่มแฝงของข้อมูลจึงอาจขาดความถูกต้อง และการประมาณค่าพารามิเตอร์อาจคลาดเคลื่อนไปจากค่าจริง (Lee et al., 2018, pp. 151-152; Sen et al., 2019, p. 273)

ข้อจำกัดของโมเดล multilevel IRT คือ โมเดลไม่มีการคำนึงถึงภาวะวิวิธพันธุ์ (Heterogeneity) ที่อาจเกิดขึ้นกับข้อมูล ผลการวิเคราะห์จึงอาจขาดสารสนเทศเกี่ยวกับสมาชิกในกลุ่มประชากรย่อยหรือกลุ่มแฝงไป (Cho & Cohen, 2010, pp. 337-338)

ข้อจำกัดของโมเดล multilevel LC คือ โมเดลมีการวัดความสามารถของผู้สอบเป็นตัวแปรจัดกลุ่มหรือกลุ่มแฝง โดยในกลุ่มแฝงนี้ผู้สอบทุกคนจะถูกพิจารณาให้มีความสามารถเท่าเทียมกัน ผลการวิเคราะห์จึงไม่มีการให้สารสนเทศเกี่ยวกับความผันแปรของความสามารถระหว่างผู้สอบแต่ละคน (Cho, 2007, p. 30)

ด้วยเหตุนี้โมเดล MMixIRT จึงถูกพัฒนาขึ้นเพื่อผ่อนคลายข้อจำกัดข้างต้น โดยมีเป้าหมายคือการจัดกลุ่มแฝงให้กับนักเรียนในแต่ละโรงเรียนที่มีแบบแผนการตอบคล้ายคลึงกัน และมีการให้สารสนเทศเชิงปริมาณเกี่ยวกับคุณลักษณะแฝงเช่นเดียวกับโมเดล MixIRT แต่มีจุดแตกต่างตรงที่โมเดล MMixIRT มีการพิจารณาให้กลุ่มแฝงระดับหน่วย (นักเรียน) ซ้อนอยู่ภายใต้กลุ่มแฝงระดับกลุ่ม (โรงเรียน) เพื่อแสดงนัยว่าโรงเรียนอาจมีผลต่อคุณลักษณะของนักเรียนตามแนวคิดโมเดลพหุระดับ ทั้งนี้การแบ่งรูปแบบของโมเดล MMixIRT สามารถพิจารณาได้ในทำนองเดียวกับโมเดล IRT ดังนี้

7.1 โมเดล multilevel mixture Rasch (MMixRasch)

โมเดล MMixRasch (Cho, 2007, p. 33; Cho & Cohen, 2010; p. 338) ได้รับการบูรณาการมาจากโมเดล Rasch โมเดล LC และโมเดลพหุระดับ จึงมีข้อตกลงเบื้องต้นคล้ายกับโมเดล MixIRT แต่มีข้อตกลงเพิ่มเติมว่าข้อมูลในระดับโรงเรียนอาจมีผลต่อความน่าจะเป็นของการเป็นสมาชิกในกลุ่มแฝงระดับนักเรียนด้วย (Tay et al., 2011, pp. 181-183) ทั้งนี้ โมเดล MMixRasch มีรูปแบบทั่วไปดังสมการ (20)

$$P(y_{ijs} = 1 \mid \theta_{jskg}) = \sum_{g=1}^{G} \sum_{k=1}^{K} \pi_g \cdot \pi_{k|g} \left(\frac{1}{1 + \exp\left[-(\theta_{jskg} - b_{ikg}) \right]} \right) \qquad (20)$$

เมื่อ π_g คือ สัดส่วนของสมาชิกในกลุ่มแฝงระดับโรงเรียน g และ $\pi_{k|g}$ คือ สัดส่วนของสมาชิกในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g และ θ_{jskg} คือ พารามิเตอร์ความสามารถของนักเรียน j ในโรงเรียน s ซึ่งอยู่ในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g เมื่อ $\theta_{jskg} \sim N(\mu_{kg}, \sigma_{kg}^2)$; μ_{kg} และ σ_{kg}^2 คือ ค่าเฉลี่ยและความแปรปรวนของความสามารถในกลุ่มแฝงระดับนักเรียนและระดับโรงเรียน และ b_{ikg} คือ พารามิเตอร์ความยากของข้อสอบ i ในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g

7.2 โมเดล multilevel mixture 2PL (MMix2PL)

โมเดล MMix2PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (21)

$$P(y_{ijs} = 1 \mid \theta_{jskg}) = \sum_{g=1}^{G} \sum_{k=1}^{K} \pi_g \cdot \pi_{k|g} \left(\frac{1}{1 + \exp\left[-a_{ikg} (\theta_{jskg} - b_{ikg}) \right]} \right) \qquad (21)$$

เมื่อ a_{ikg} คือพารามิเตอร์อำนาจจำแนกของข้อสอบ i ในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g

7.3 โมเดล multilevel mixture 3PL (MMix3PL)

โมเดล MMix3PL มีฟังก์ชันการตอบสนองข้อสอบดังสมการ (22)

$$P(y_{ijs} = 1 \mid \theta_{jskg}) = \sum_{g=1}^{G} \sum_{k=1}^{K} \pi_g \cdot \pi_{k|g} \left(c_{ikg} + (1 - c_{ikg}) \frac{1}{1 + \exp\left[-a_{ikg} (\theta_{jskg} - b_{ikg}) \right]} \right) \qquad (22)$$

เมื่อ c_{ikg} คือ พารามิเตอร์การเดาของข้อสอบ i ในกลุ่มแฝงระดับนักเรียน k ในกลุ่มแฝงระดับโรงเรียน g

การเปรียบเทียบพารามิเตอร์ของโมเดลและโปรแกรมที่รองรับการวิเคราะห์

เมื่อเปรียบเทียบโมเดลทั้ง 7 รูปแบบ อาจแบ่งตามลักษณะของการวัดความสามารถแฝงได้เป็น 4 ประเภท ได้แก่ 1) โมเดลที่ไม่มีการวัดความสามารถแฝง ได้แก่ โมเดลพหุระดับ 2) โมเดลที่วัดความสามารถแฝงเป็นค่าไม่ต่อเนื่อง ได้แก่ โมเดล LC และ

multilevel LC 3) โมเดลที่วัดความสามารถแฝงเป็นค่าต่อเนื่อง ได้แก่ โมเดล IRT และ multilevel IRT และ 4) โมเดลที่วัดความสามารถแฝงเป็นทั้งค่าต่อเนื่องและไม่ต่อเนื่อง (แบบผสม) ได้แก่ โมเดล MixIRT และ MMixIRT นอกจากนี้อาจสังเกตได้ว่าการวัดพารามิเตอร์ของข้อสอบจะมีเฉพาะในโมเดลที่มีการบูรณาการมาจากโมเดล IRT ในขณะที่การวัดพารามิเตอร์สัดส่วนของสมาชิกในกลุ่มแฝงก็จะมีเฉพาะในโมเดลที่มีการบูรณาการมาจากโมเดล LC เช่นกัน ส่วนโปรแกรมคอมพิวเตอร์ที่รองรับการวิเคราะห์สำหรับทุกโมเดล เช่น โปรแกรม Mplus (Muthén & Muthén, 2015) และโปรแกรม MultiBUGS (Goudie et al., 2017) รายละเอียดดังตาราง 2

ตาราง 2 การเปรียบเทียบพารามิเตอร์ของโมเดล และโปรแกรมที่รองรับการวิเคราะห์

โมเดล	**พารามิเตอร์**			**โปรแกรมที่รองรับการวิเคราะห์**
	ความสามารถ	**ข้อสอบ**	**กลุ่มแฝง**	
1. IRT	$\theta_j \sim N(0,1)$ * Continuous	a_i, b_i, c_i	-	- IRTPRO - Mplus - R package e.g. *ltm*, *mirt*, *irtoys* - WinBUGS/OpenBUGS/MultiBUGS
2. LC	$\theta_j = \sum_{k=1}^{K} \pi_k \cdot \theta_k$ * Categorical	-	π_k	- Latent Gold - Mplus - R package e.g. *randomLCA*, *poLCA* - WinBUGS/OpenBUGS/MultiBUGS
3. Multilevel	-	-	-	- HLM - Mplus - R package e.g. *lme4*, *nlme*, *rstan* - WinBUGS/OpenBUGS/MultiBUGS
4. MixIRT	$\theta_{jk} \sim N(\mu_k, \sigma_k^2)$ * Mixed	a_{ik}, b_{ik}, c_{ik}	π_k	- Latent Gold - Mplus - R package e.g. *psychomix*, *mixRasch* - WinBUGS/OpenBUGS/MultiBUGS - WINMIRA
5. Multilevel IRT	$\theta_{js} = u_{0s} + r_{js}$; $u_{0s} \sim N(\gamma_{00}, \zeta)$, $r_{js} \sim N(0, \tau)$ * Continuous	a_i, b_i, c_i		- Mplus - R package e.g. *lme4*, *nlme*, *rstan* - WinBUGS/OpenBUGS/MultiBUGS
6. Multilevel LC	$\theta_{js} = \sum_{g=1}^{G} \sum_{k=1}^{K} \pi_g\ \pi_{k\|g} \cdot \theta_{kg}$ * Categorical	-	$\pi_g, \pi_{k\|g}$	- Latent Gold - Mplus - WinBUGS/OpenBUGS/MultiBUGS
7. MMixIRT	$\theta_{jskg} \sim N(\mu_{kg}, \sigma_{kg}^2)$ * Mixed	$a_{ikg}, b_{ikg}, c_{ikg}$	$\pi_g, \pi_{k\|g}$	- Latent Gold - Mplus - WinBUGS/OpenBUGS/MultiBUGS

บทสรุป

โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) เป็นการบูรณาการองค์ความรู้จาก 3 โมเดล คือ โมเดลทฤษฎีการตอบสนองข้อสอบ โมเดลกลุ่มแฝง และโมเดลพหุระดับ โมเดลจึงสามารถให้สารสนเทศเชิงปริมาณและเชิงคุณภาพเกี่ยวกับคุณลักษณะแฝงและกลุ่มแฝงได้พร้อมกัน รวมถึงยังมีการคำนึงถึงโครงสร้างพหุระดับอันเป็นธรรมชาติของข้อมูลทางการศึกษา การวิเคราะห์ด้วยโมเดลดังกล่าวจึงมีการให้สารสนเทศที่ครอบคลุม สอดคล้องกับธรรมชาติของข้อมูล และมีความถูกต้องใน

การวิเคราะห์กลุ่มแฝงมากกว่าโมเดลกลุ่มแฝงทั่วไป โดยเฉพาะเมื่อนำไปวิเคราะห์กับข้อมูลการทดสอบขนาดใหญ่หรือการทดสอบระดับชาติ เช่น O-NET PISA และ TIMSS ซึ่งผู้เข้าทดสอบเป็นนักเรียนที่มีคุณลักษณะหลากหลายจากโรงเรียนต่างๆ ทั่วประเทศ หรือจากประเทศต่างๆ ทั่วโลก

การนำเสนอในบทความนี้เป็นเพียงการแนะนำแนวคิดในการบูรณาการโมเดลที่มีความเกี่ยวข้องกับโมเดล MMixIRT ซึ่งผู้เขียนต้องการแสดงให้เห็นถึงจุดเชื่อมโยงของแต่ละโมเดล ความแตกต่างของพารามิเตอร์ รวมถึงโปรแกรมที่รองรับการวิเคราะห์ ซึ่งน่าจะช่วยให้ผู้อ่านที่สนใจสามารถจำแนกรูปแบบการวิเคราะห์ของแต่ละโมเดลได้ชัดเจนยิ่งขึ้น อันจะเป็นประโยชน์ต่อการตัดสินใจเลือกใช้โมเดลให้เหมาะสมกับบริบทที่สนใจศึกษาต่อไป

References

Cho, S. J. (2007). *A multilevel mixture IRT model for DIF analysis* (Unpublished doctoral dissertation). Athens, GA: University of Georgia.

Cho, S. J., & Cohen, A. S. (2010). A multilevel mixture model with applications to DIF. *Journal of Educational and Behavioral Statistics,* 35(3), 336-370.

Cohen, A., Wollack, J., Bolt, D., & Mroch, A. (2002). A mixture Rasch model analysis of test speededness. *the annual meeting of the American Educational Research Association.* New Orleans: LA.

Fox, J. P., & Glas, C. (2001). Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika*, 66(2), 271-288.

Goudie, R. J., Turner, R. M., De Angelis, D., & Thomas, A. (2017). *Massively parallel MCMC for Bayesian hierarchical models* (arXiv Preprint No.1704.03216). Retrieved from https://arxiv.org/abs/1704.03216

Hox, J. J., & Maas, C. J. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling,* 8(2), 157-174.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. *Journal of Educational Measurement*, 38(1), 79-93.

Kanjanawasee, S. (2012). *Modern test theories* (4th ed.). Bangkok: Chulalongkorn University Prss. [in Thai]

Lee, W. Y., Cho, S. J., & Sterba, S. K. (2018). Ignoring a multilevel structure in mixture item response models: impact on parameter recovery and model selection. *Applied Psychological Measurement,* 42(2), 136-154.

Lord, F. M., & Novick, M. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Lubke, G., & Muthén, B. O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods,* 10(1), 21-39.

Maier, K. S. (2001). A Rasch hierarchical measurement model. *Journal of Educational and Behavioral Statistics,* 26(3), 307-330.

Muthén, L., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE.

Reckase, M. D. (2009). *Multidimensional item response theory*. New York, NY: Springer.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement,* 14(3), 271-282.

Sen, S., Cohen, A., & Kim, S. (2019). Model selection for multilevel mixture Rasch models. *Applied Psychological Measurement,* 43(4), 272-289.

Tay, L., Diener, E., Drasgow, F., & Vermunt, J. (2011). Multilevel mixed-measurement IRT analysis: An explication and application to self-reported emotions across the world. *Organizational Research Methods,* 14(1), 177-207.

Vermunt, J. (2003). Multilevel latent class models. *Sociological Methodology*, 33(1), 213-239.

Wongnam, P. (2012). Multilevel regression analysis with HLM 6. *Journal of Education*, 23(3), 27-42. [in Thai]