การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับสำหรับการวิเคราะห์กลุ่มความสามารถแฝง

นางภัทรภร เจนสุทธิเวชกุล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2558
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

DEVELOPMENT OF THE MULTILEVEL MIXTURE ITEM RESPONSE THEORY MODEL FOR IDENTIFYING LATENT CLASSES

Mrs. Pattaraporn Jensuttiwetchakul

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2015
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับสำหรับการวิเคราะห์กลุ่มความสามารถแฝง
โดย	นางภัทรภร เจนสุทธิเวชกุล
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

.....คณบดีคณะครุศาสตร์

(รองศาสตราจารย์ ดร.บัญชา ชลาภิรมย์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก)

.....กรรมการ

(รองศาสตราจารย์ ดร.วรรณี แกมเกตุ)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง)

.....กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร.อวยพร เรืองตระกูล)

ภัทรภร เจนสุทธิเวชกุล : การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับสำหรับการวิเคราะห์กลุ่มความสามารถแฝง (DEVELOPMENT OF THE MULTILEVEL MIXTURE ITEM RESPONSE THEORY MODEL FOR IDENTIFYING LATENT CLASSES) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ศ. ดร.ศิริชัย กาญจนวาสี, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ผศ. ดร.สังวรณ์ งัดกระโทก, 231 หน้า.

งานวิจัยนี้มีวัตถุประสงค์คือ (1) เพื่อพัฒนาโมเดล การประมาณค่ากลุ่มความสามารถแฝงของบุคคลตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ (2) เพื่อตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น จากข้อมูลจำลอง และ (3) เพื่อวิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น จากข้อมูลผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ข้อมูลที่ใช้ในการศึกษาครั้งนี้ มี 2 ส่วนคือ (1) ข้อมูลจำลองผลการตอบข้อสอบของนักเรียน ภายใต้เงื่อนไขคือ จำลองกลุ่มแฝงระดับโรงเรียน 6 กลุ่มแฝง แต่ละกลุ่มแฝงจำลองผลการตอบข้อสอบสองสถานการณ์คือ จำนวนข้อสอบ 6 ข้อ และ 15 ข้อ ของนักเรียนในโรงเรียน 4 ขนาด และ (2) ข้อมูลทุติยภูมิ ผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) การศึกษาครั้งนี้จำลองข้อมูลโดยใช้การเขียนคำสั่งการประมวลผลบนโปรแกรม R ด้วยแพ็คเกจ IRTOYS และวิเคราะห์ข้อมูลโดยใช้การเขียนคำสั่งการประมวลผลบนโปรแกรม R ด้วยแพ็คเกจ LTM และ MCLUST และตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มแฝงด้วยการวิเคราะห์ซ้ำแต่ละสถานการณ์จำนวน 25 ครั้ง

ผลการวิจัยที่สำคัญสรุปได้ดังนี้

1. โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น สำหรับจำแนกกลุ่มความสามารถแฝงของบุคคลเป็นโมเดลที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ ซึ่งมีขั้นตอนการวิเคราะห์ทั้งสิ้น 5 ขั้นตอนได้แก่ (1) ประมาณค่าพารามิเตอร์ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อของข้อมูลจำลอง (2) แบ่งกลุ่มความสามารถแฝงระดับบุคคล (3) คำนวณความสามารถของโรงเรียน (4) แบ่งกลุ่มความสามารถแฝงระดับโรงเรียน (5) คำนวณหาความน่าจะเป็นที่จะตอบคำถามในแต่ละข้อถูกของแต่ละโรงเรียน

2. การศึกษาด้วยข้อมูลจำลอง ผลการวิจัยให้ข้อค้นพบว่าโมเดลที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงของบุคคลได้ถูกต้องและแม่นยำ สำหรับข้อสอบ 6 ข้อ จำแนกได้ถูกต้องคิดเป็นร้อยละ 96 และสำหรับข้อสอบ 15 ข้อ จำแนกได้ถูกต้อง คิดเป็นร้อยละ 100

3. การศึกษาด้วยข้อมูลผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พบว่า สามารถจำแนกกลุ่มความสามารถแฝงของบุคคล ในระดับบุคคลและระดับโรงเรียนได้ 20 กลุ่ม และ 9 กลุ่มตามลำดับ

ภาควิชา	วิจัยและจิตวิทยาการศึกษา	ลายมือชื่อนิสิต .....
สาขาวิชา	การวัดและประเมินผลการศึกษา	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
ปีการศึกษา	2558	ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

5384240827 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORDS: MULTILEVEL MIXTURE ITEM RESPONSE THEORY MODEL, LATENT CLASS MODEL, THREE PARAMETER ITEM RESPONSE THEORY MODEL, MULTILEVEL MODEL

PATTARAPORN JENSUTTIWETCHAKUL: DEVELOPMENT OF THE MULTILEVEL MIXTURE ITEM RESPONSE THEORY MODEL FOR IDENTIFYING LATENT CLASSES. ADVISOR: PROF. SIRICHAI KANJANAWASEE, Ph.D., CO-ADVISOR: ASST. PROF. SUNGWORN NGUDGRATOKE, Ph.D., 231 pp.

This research aims to (1) develop a model for estimating individual's latent ability based on Item response theory by integrating three parameters IRT model, latent class model, and multilevel model, called 3PL-MMixIRT, (2) examine the correctness and accuracy of the developed model in classifying individual's latent class by using simulated data, (3) apply the model to classify individual's latent class using the empirical data from grade 12 students' O-NET test score in mathematics.

The data can be divided into (1) simulated data and (2) empirical data. In the first case, the simulated data consist of two situations. The first situation is 6 test items with 4 school sizes, while the other situation is 15 test items with 4 school sizes. In the other case, the model was applied for the empirical data retrieved from grade 12 students' O-NET test score in mathematics. Program R was adopted to develop the model which IRTOYS package was used to create simulation data, LTM package was utilized to estimate parameters, and MCLUST package was applied to classify latent class. Moreover, to ensure correctness and accuracy of the developed model, 25 iterations are applied.

This study provides three important results. First, the developed model for estimating individual's latent ability based on Item response theory by integrating three parameters IRT model, latent class model, and multilevel model consists of five steps, which are (1) parameter estimation based on three parameter IRT, (2) individual latent class analysis, (3) school ability calculation based on proportional probability, (4) school latent class analysis, and finally (5) probability to answer each item correctly. Second, the analysis with simulated data indicates the correctness and accuracy of the model in both 6-item test and 15-item test. The first situation reveals 96% correctness, while the other situation unveils 100% correctness. Finally, the analysis with empirical data from grade 12 students' O-NET test score in mathematics reveals that there are 20 school latent classes, and 9 individual latent classes, respectively.

CHULALONGKORN UNIVERSITY

Department:	Educational Research and Psychology	Student's Signature .....
		Advisor's Signature .....
Field of Study:	Educational Measurement and Evaluation	Co-Advisor's Signature .....
Academic Year:	2015	

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงด้วยความเมตตากรุณาและความเอาใจใส่อย่างดียิ่งจากท่านศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ให้ความรู้ด้านการวัดและประเมินผล ให้คำปรึกษาแนะนำ ชี้แนะประเด็นสำคัญในการทำวิจัย ให้โอกาสผู้วิจัยได้มีประสบการณ์การเรียนรู้ที่ลึกซึ้งมากยิ่งขึ้น รวมถึงดูแลเอาใจใส่ ห่วงใย ให้กำลังใจ ติดตามการทำงานของผู้วิจัยจนทำให้วิทยานิพนธ์ฉบับนี้มีคุณค่าอย่างสมบูรณ์ และกราบของพระคุณ ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ผู้ให้คำปรึกษา แนะนำ และปรับแก้ข้อบกพร่องอันเป็นประโยชน์อย่างสูงในการทำวิทยานิพนธ์ด้วยดีตลอดมา ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่งและขอกราบขอบพระคุณท่านอาจารย์เป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ ประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.อวยพร เรืองตระกูล รองศาสตราจารย์ ดร.วรรณี แกมเกตุ และผู้ช่วยศาสตราจารย์ณัฐภรณ์ หลาวทอง กรรมการสอบวิทยานิพนธ์ ที่ได้ให้คำแนะนำในการปรับปรุงวิทยานิพนธ์ฉบับนี้ ให้มีความสมบูรณ์ยิ่งขึ้น ตลอดจนให้ข้อเสนอแนะข้อคิดเห็นต่างๆ ที่เป็นประโยชน์ต่อการวิจัยในครั้งนี้

ขอกราบขอบพระคุณ คณาจารย์ภาควิชาวิจัยการศึกษาทุกท่านที่ได้ประสิทธิ์ประสาทความรู้แก่ผู้วิจัย ให้คำแนะนำและข้อคิดเห็นต่าง ๆ คอยห่วงใย ให้กำลังใจผู้วิจัยด้วยความเมตตากรุณาเสมอมา

ขอกราบขอบพระคุณ ศาสตราจารย์กิตติคุณ ดร.นงลักษณ์ วิรัชชัย รองศาสตราจารย์ศักดา บุญยไวโรจน์ ที่ให้คำแนะนำ ข้อเสนอแนะต่าง ๆ ในการทำงานวิจัย และการเขียนบทความวิจัยที่ดีแก่ผู้วิจัยเสมอ รวมทั้งคอยติดตามความก้าวหน้าในการวิจัยด้วยความห่วงใย และให้กำลังใจผู้วิจัย เป็นแรงใจให้กับผู้วิจัยอย่างดียิ่ง

ขอขอบคุณบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย เป็นอย่างสูงที่มอบทุนสนับสนุนการทำวิจัยหรือทุน “90 ปี จุฬาลงกรณ์มหาวิทยาลัย” ให้แก่ผู้วิจัยทำให้สามารถวิทยานิพนธ์ได้อย่างเต็มที่

ขอขอบคุณผู้อำนวยการ และคณาจารย์โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยมทุกท่าน โดยเฉพาะอย่างยิ่ง คณาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ที่ให้ความอนุเคราะห์ และสนับสนุนให้ผู้วิจัยได้มีโอกาสในการพัฒนาตนเอง รวมทั้งคอยให้ความช่วยเหลือ และเป็นกัลยาณมิตรที่ดีตลอดเวลา

ขอขอบคุณพี่ ๆ เพื่อน ๆ สาขาการวัดและการประเมินผลทางการศึกษาทุกท่าน ที่คอยให้กำลังใจและให้ความช่วยเหลือต่าง ๆ ตลอดเวลาที่ทำวิจัยนี้

สุดท้ายนี้ ผู้วิจัยขอกราบขอบพระคุณบุคคลผู้เป็นที่รักและเคารพยิ่งของผู้วิจัย คือ คุณพ่อไพศาล คุณแม่สมหมาย แสงไชย และน้องสาวกรวรรณ แสงตระกูล ผู้ที่รักและห่วงใยผู้วิจัย คอยให้กำลังใจ และสนับสนุนผู้วิจัยตลอดมาในทุก ๆ ด้าน และที่สำคัญที่สุดขอขอบคุณอาจารย์ธนพล เจนสุทธิเวชกุล คู่สมรสที่คอยให้ทั้งความรัก กำลังใจ ให้ความช่วยเหลือ คำแนะนำและเป็นเพื่อนคู่คิดในการทำวิจัยครั้งนี้อย่างดียิ่ง จนทำให้ผู้วิจัยประสบความสำเร็จตามที่มุ่งหวัง

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ญ
สารบัญภาพ ..... ฒ
บทที่ 1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 คำถามวิจัย 5
1.3 วัตถุประสงค์การวิจัย 6
1.4 ขอบเขตการวิจัย 6
1.5 นิยามศัพท์ที่ใช้ในการวิจัย 7
1.6 ประโยชน์ที่ได้รับ 8
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 10
ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ 11
1.1 ทฤษฎีการตอบสนองข้อสอบแบบผสม (Mixture IRT : MixIRT) 12
1.1.1 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) 13
1.1.2 โมเดลการวิเคราะห์กลุ่มแฝง (Latent class Model) หรือ โมเดลโครงสร้างแฝง 19
1.1.3 ทฤษฎีการตอบสนองข้อสอบแบบผสม (Mixture Item Response Theory: MixIRT) 28
1.2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT:MMixIRT) 32

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หน้า

1.2.1 การวิเคราะห์พหุระดับ 33
1.2.2 โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel mixture IRT) . 33
1.3 การประมาณค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบ 47
1.3.1 วิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood: ML)..... 49
1.3.2 การประมาณค่าพารามิเตอร์แบบเบส์ 57
1.4 วิธีการเลือกโมเดล (Model Selection Methods)..... 60
ตอนที่ 2 การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล..... 62
2.1 ความเป็นมาของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล 62
2.2 ขั้นตอนการจำลองข้อมูล..... 63
ตอนที่ 3 การทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O– NET)..... 83
3.1 ความสำคัญของการสอบ O-NET..... 84
3.2 ลักษณะทั่วไปของข้อสอบ O – NET..... 85
3.3 วิธีการคิดคะแนน..... 86
3.4 ลักษณะข้อสอบ O-NET วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6 86
ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง..... 88
4.1 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 88
4.2 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม..... 89
ตอนที่ 5 กรอบแนวคิดในการวิจัย..... 100
บทที่ 3 วิธีดำเนินการวิจัย 102
ตอนที่ 1 การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ..... 104
ตอนที่ 2 การศึกษากับข้อมูลจำลอง..... 106
ตอนที่ 3 การศึกษากับข้อมูลเชิงประจักษ์ 113
บทที่ 4 ผลการวิเคราะห์ข้อมูล 118

หน้า

ตอนที่ 1 ผลการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 118
ตอนที่ 2 ผลการตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล จากการใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับด้วยข้อมูลจำลอง 124
ตอนที่ 3 ผลการวิเคราะห์ข้อมูลเชิงประจักษ์ เพื่อจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น 158
บทที่ 5 สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ 186
5.1 สรุปผลการวิจัย 187
5.1.1. ผลการพัฒนาโมเดลการประมาณค่าความสามารถตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) 187
5.1.2 ผลการทดสอบความถูกต้องแม่นยำในการจำแนกกลุ่มแฝงโดยใช้โมเดลที่พัฒนาขึ้นกับข้อมูลจำลอง 188
5.1.3 ผลการจำแนกกลุ่มแฝงโดยใช้โมเดลที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ 191
5.2 อภิปรายผลการวิจัย 193
5.3 ข้อเสนอแนะในการนำผลการวิจัยไปใช้ 196
5.4 ข้อเสนอแนะในการทำวิจัยต่อไป 197
รายการอ้างอิง 199
ภาคผนวก 211
ภาคผนวก ก คำสั่งที่ใช้ในการศึกษาด้วยโปรแกรม R 212
ภาคผนวก ข รายชื่อโรงเรียนในแต่ละกลุ่มแฝง ที่จำแนกโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 223
ประวัติผู้เขียนวิทยานิพนธ์ 231

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตารางที่ 2.1 การจำแนกวิธีการใช้ในการจำแนกความแตกต่างของประชากร 20

ตารางที่ 2.2 ความแตกต่างของโมเดลที่มีตัวแปรแฝง 23

ตารางที่ 2.3 โครงสร้างสัดส่วนของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 38

ตารางที่ 2.4 โมเดลโครงสร้างที่แตกต่างกันภายในกรอบแนวคิดตัวแปรแฝงทั่วไป 46

ตารางที่ 2.5 จำนวนข้อสอบ O-NET วิชาคณิตศาสตร์ แยกตามรูปแบบของข้อสอบ ปีการศึกษา 2555 86

ตารางที่ 2.6 จำนวนข้อสอบ O-NET วิชาคณิตศาสตร์ แยกตามสาระ และมาตรฐานการเรียนรู้ 87

ตารางที่ 2.7 งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับและทฤษฎีการตอบสนองข้อสอบแบบผสม 96

ตารางที่ 3.1 กำหนดการจำลองกลุ่มแฝงของโรงเรียนจำนวน 24 โรงเรียน 108

ตารางที่ 3.2 การจำลองกลุ่มแฝงของโรงเรียน 4 กลุ่มจำแนกตามความสามารถในการตอบคำถาม 109

ตารางที่ 3.3 การกำหนดค่าพารามิเตอร์ในการจำลองรูปแบบการตอบ จำนวน 6 ข้อ 110

ตารางที่ 3.4 การกำหนดค่าพารามิเตอร์ในการจำลองรูปแบบการตอบ จำนวน 15 ข้อ 111

ตารางที่ 3.5 จำนวนนักเรียนที่จำลองขึ้นสำหรับแต่ละกลุ่มแฝงของแต่ละขนาดโรงเรียน 111

ตารางที่ 3.6 จำนวนประชากรจำแนกตามสังกัดและขนาดของโรงเรียน 113

ตารางที่ 3.7 จำนวนกลุ่มตัวอย่างจำแนกตามสังกัดและขนาดของโรงเรียน 115

ตารางที่ 3.8 ลำดับข้อสอบของข้อมูลเชิงประจักษ์ และและข้อสอบที่เลือกใช้ในการจำแนกกลุ่มความสามารถแฝง 116

ตารางที่ 4.1 ผลการประมาณค่าพารามิเตอร์ของข้อมูลจำลองจำนวน 6 ข้อ 125

ตารางที่ 4.2 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบสำหรับข้อมูลจำลองจำนวน 6 ข้อ 126

หน้า

ตารางที่ 4.3 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก สำหรับข้อมูลจำลองจำนวน 6 ข้อ 128

ตารางที่ 4.4 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 6 ข้อ 130

ตารางที่ 4.5 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มแฝงระดับบุคคล สำหรับข้อมูลจำลองจำนวน 6 ข้อ 130

ตารางที่ 4.6 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มี ความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 6 ข้อ 131

ตารางที่ 4.7 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ 132

ตารางที่ 4.8 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝงระดับบุคคล สำหรับข้อมูลจำลองจำนวน 6 ข้อ 134

ตารางที่ 4.9 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 6 ข้อ 135

ตารางที่ 4.10 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับโรงเรียน สำหรับข้อมูลจำลอง จำนวน 6 ข้อ 136

ตารางที่ 4.11 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 6 ข้อ 136

ตารางที่ 4.12 ผลการจำแนกกลุ่มแฝงระดับโรงเรียนจำแนกตามโรงเรียนสำหรับข้อมูลจำลอง จำนวน 6 ข้อ 137

ตารางที่ 4.13 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ 138

ตารางที่ 4.14 ผลการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากการวิเคราะห์ซ้ำ 25 ครั้ง สำหรับข้อมูลจำลอง 6 ข้อ 140

ตารางที่ 4.15 ผลการประมาณค่าพารามิเตอร์ สำหรับข้อมูลจำลองจำนวน 15 ข้อ 141

หน้า

ตารางที่ 4.16 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบ
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 143

ตารางที่ 4.17 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 144

ตารางที่ 4.18 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3 อันดับสูงสุด
สำหรับข้อมูลจำลอง 15 ข้อ 145

ตารางที่ 4.19 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับบุคคล g
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 145

ตารางที่ 4.20 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของ
นักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 146

ตารางที่ 4.21 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 148

ตารางที่ 4.22 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝง
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 149

ตารางที่ 4.23 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน 3 อันดับสูงสุด
สำหรับข้อมูลจำลอง 15 ข้อ 151

ตารางที่ 4.24 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับโรงเรียน
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 151

ตารางที่ 4.25 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของ
นักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม
สำหรับข้อมูลจำลองจำนวน 15 ข้อ 152

ตารางที่ 4.26 ผลการจำแนกกลุ่มแฝงระดับโรงเรียนจำแนกตามโรงเรียน
สำหรับข้อมูลจำลอง จำนวน 15 ข้อ 153

ตารางที่ 4.27 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับ
โรงเรียน สำหรับข้อมูลจำลองจำนวน 15 ข้อ 154

หน้า

ตารางที่ 4.28 ผลการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากการวิเคราะห์ซ้ำ 25 ครั้ง สำหรับข้อมูลจำลอง 15 ข้อ 156

ตารางที่ 4.29 ผลการประมาณค่าพารามิเตอร์ สำหรับข้อมูลเชิงประจักษ์จำนวน 15 ข้อ 158

ตารางที่ 4.30 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบสำหรับข้อมูลเชิงประจักษ์ 160

ตารางที่ 4.31 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูกสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 161

ตารางที่ 4.32 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3อันดับสูงสุดสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 162

ตารางที่ 4.33 จำนวนรูปแบบ สัดส่วนในแต่ละกลุ่มแฝงระดับบุคคลสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 163

ตารางที่ 4.34 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม 164

ตารางที่ 4.35 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียนสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 171

ตารางที่ 4.36 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝงสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 172

ตารางที่ 4.37 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน สูงสุด 3 อันดับแรกสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 173

ตารางที่ 4.38 จำนวนโรงเรียนและสัดส่วนของแต่ละกลุ่มความสามารถแฝงระดับโรงเรียนสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 173

ตารางที่ 4.39 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกในแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม
สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 174

ตารางที่ 4.40 รหัสโรงเรียนที่ได้จากการจำแนกกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝงสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 176

หน้า

ตารางที่ 4.41 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียนสำหรับข้อมูลเชิงประจักษ์ 15 ข้อ 180

ตารางที่ 4.42 เปรียบเทียบผลการวิเคราะห์ข้อมูลจำลอง 6 ข้อ 15 ข้อ และ ข้อมูลเชิงประจักษ์ 15 ข้อ ในการจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคลด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 185

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญภาพ

หน้า

ภาพที่ 2.1 โค้งคุณลักษณะข้อสอบสำหรับโมเดล 3 พารามิเตอร์ 17

ภาพที่ 2.2 ตัวแปรแฝง ที่มีตัวแปรสังเกตได้ 4 ตัว เป็นตัวชี้วัด 21

ภาพที่ 2.3 แผนภาพเวนน์แสดงการบูรณาการกันเป็นโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ 34

ภาพที่ 2.4 โครงสร้างของข้อมูลพหุระดับ 37

ภาพที่ 2.5 ความชัดเจนของความยากของข้อสอบในกลุ่ม 38

ภาพที่ 2.6 การแจกแจงแบบยูนิฟอร์ม 65

ภาพที่ 2.7 การแจกแจงแบบปกติ 65

ภาพที่ 2.8 การแจกแจงแบบแกมมา 67

ภาพที่ 2.9 การแจกแจงแบบเบต้า 68

ภาพที่ 2.10 การแจกแจงแบบไคว์สแควร์ 69

ภาพที่ 2.11 การแจกแจงแบบที 70

ภาพที่ 2.12 กรอบแนวคิดในการวิจัย 101

ภาพที่ 3.1 ขั้นตอนการวิเคราะห์ข้อมูล 112

ภาพที่ 4.1 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลจำลอง 6 ข้อ 125

ภาพที่ 4.2 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝง สำหรับข้อมูลจำลองจำนวน 6 ข้อ 132

ภาพที่ 4.3 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ 137

ภาพที่ 4.4 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน จากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 6 ข้อ 139

ภาพที่ 4.5 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง15 ข้อ (ข้อที่ 1 ถึง 5) 142

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

หน้า

ภาพที่ 4.6 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง 15 ข้อ (ข้อที่ 6 ถึง 10)..... 142

ภาพที่ 4.7 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง 15 ข้อ (ข้อที่ 11 ถึง 15)..... 142

ภาพที่ 4.8 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 15 ข้อ..... 147

ภาพที่ 4.9 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 15 ข้อ..... 153

ภาพที่ 4.10 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียนจากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 15 ข้อ..... 155

ภาพที่ 4.11 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 2, 4, 6, 8 และ 9)..... 159

ภาพที่ 4.12 โค้งคุณลักษณะข้อสอบจากการ ประมาณค่าพารามิเตอร์ สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 10, 13, 15, 16 และ 18)..... 159

ภาพที่ 4.13 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 20, 22, 27, 29 และ 31)..... 159

ภาพที่ 4.14 กราฟแท่งแสดงจำนวนข้อสอบจำแนกตามระดับความน่าจะเป็นในการตอบข้อสอบถูกของกลุ่มแฝงระดับบุคคล 20 กลุ่มแฝง..... 166

ภาพที่ 4.15 ความน่าจะเป็นที่จะตอบข้อสอบในแต่ละข้อถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝงของข้อมูลเชิงประจักษ์จำนวน 15 ข้อ..... 170

ภาพที่ 4.16 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ..... 175

ภาพที่ 4.17 กราฟแท่งแสดงจำนวนข้อสอบจำแนกตามระดับความน่าจะเป็นในการตอบข้อสอบถูกของกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง..... 181

ภาพที่ 4.18 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียนจากการจำแนกกลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ..... 182

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การตรวจสอบคุณภาพการจัดการศึกษาของประเทศจำเป็นต้องมีการประเมินผลเพื่อให้ได้ข้อมูลย้อนกลับที่สามารถสะท้อนถึงการบรรลุเป้าหมายตามมาตรฐานการศึกษาของสถานศึกษาทั้งประเทศ รวมทั้งทราบจุดอ่อน หรือปัญหาที่ต้องการแก้ไข ซึ่งในการทดสอบระดับชาติ หรือระดับนานาชาติได้ดำเนินการติดตามและประเมินผลการจัดการศึกษาของสถานศึกษามาอย่างต่อเนื่อง การทดสอบระดับนานาชาติ เช่น การทดสอบในโครงการประเมินผลนักเรียนระดับนานาชาติ (Program for International Student Assessment: PISA) การทดสอบในโครงการการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระดับนานาชาติ (The International Mathematics and Science Study: TIMSS) การทดสอบระดับชาติ เช่น การทดสอบทางศึกษาระดับชาติขั้นพื้นฐาน (Ordinary National Educational Test: O – NET) การทดสอบความถนัดทั่วไป (General Aptitude Test: GAT) การทดสอบความถนัดทางวิชาชีพและวิชาการ (Professional and Academic Aptitude Test: PAT) การประเมินคุณภาพการศึกษาขั้นพื้นฐานเพื่อประกันคุณภาพผู้เรียน (National test: NT) ที่ผ่านมาการประเมินต่าง ๆ เป็นการศึกษารูปแบบประสิทธิภาพของการจัดการศึกษา และเป็นการศึกษาเพื่อเปรียบเทียบผลสัมฤทธิ์ของแต่ละประเทศ แต่ละกลุ่มโรงเรียน มีการจัดอันดับ และให้ข้อมูลป้อนกลับทำให้ได้สารสนเทศอย่างหลากหลายที่มีประโยชน์ในการตรวจสอบการปฏิบัติงานได้ และโดยทั่วไปการประเมินผลจะรายงานในรูปคะแนนเฉลี่ย เช่น การรายงานผลของ PISA ได้รายงานคะแนนเฉลี่ยเปรียบเทียบระหว่างประเทศ ซึ่งการรายงานผลโดยใช้คะแนนเฉลี่ยทำให้ไม่ทราบข้อมูลว่านักเรียนมีความสามารถด้านใดสูง ด้านใดต่ำ และอยู่ในความสามารถระดับใด ทำให้ขาดข้อมูลย้อนกลับที่จำเพาะเจาะจงสำหรับการพัฒนานักเรียน และพัฒนาการศึกษาโดยรวม

จากการประเมินผล PISA ในปี 2012 ซึ่งมีการเน้นคณิตศาสตร์เป็นการประเมินหลัก และ ปี 2009 พบว่า ประเทศไทยยังอยู่ห่างไกลจากเป้าหมายของความเข้มแข็งทางการศึกษาและไม่สามารถเตรียมเยาวชนให้มีศักยภาพในการแข่งขันในอนาคต นักเรียนไทยมีผลการประเมินต่ำทุกวิชาเมื่อเทียบกับระดับนานาชาติ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554, 2557) นอกจากนั้นเมื่อเปรียบเทียบระหว่างกลุ่มโรงเรียนพบว่านักเรียนในเมืองใหญ่ เช่น โรงเรียนในกรุงเทพฯ และปริมณฑล มีผลการประเมินแตกต่างกับนักเรียนในชนบท เช่น ในภาคกลาง ภาคใต้ และภาคอีสานล่าง นักเรียนในกลุ่มโรงเรียนจุฬาภรณราชวิทยาลัย และในกลุ่มโรงเรียนสาธิตมีผลการเรียนรู้ต่างจากโรงเรียนกลุ่มอื่น ๆ ราวกับเป็นผลจากคนละประเทศกัน โดยทั้งสองกลุ่มนี้มีคะแนนเทียบเท่ากับ

ประเทศในกลุ่มสิบประเทศบน ในขณะที่คะแนนเฉลี่ยทั้งหมดของนักเรียนไทยรู้เรื่องคณิตศาสตร์อยู่ระดับปลายล่างของตารางนานาชาติ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554, 2557) เมื่อวิเคราะห์ผลการประเมินเป็นระดับการรู้เรื่องคณิตศาสตร์ 6 ระดับ โดยกำหนดให้ระดับ 2 เป็นระดับพื้นฐานซึ่งถือเป็นระดับต่ำสุดที่นักเรียนแสดงว่ามีศักยภาพสามารถใช้ประโยชน์จากคณิตศาสตร์ในชีวิตจริงได้ในอนาคต ผลการประเมินในภาพรวมทั้งประเทศชี้ว่านักเรียน จำนวนครึ่งหนึ่ง มีทักษะทางคณิตศาสตร์ไม่ถึงระดับพื้นฐาน และมีนักเรียนเพียงร้อยละ 2.5 ที่ทักษะทางคณิตศาสตร์อยู่ในระดับสูงที่ระดับ 5 และระดับ 6 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2557) เห็นได้ว่า นักเรียนไทยมีการรู้เรื่องคณิตศาสตร์ในระดับที่ไม่น่าพึงพอใจและไม่แสดงศักยภาพที่จะสามารถใช้คณิตศาสตร์ให้เป็นประโยชน์ในชีวิตได้ ทั้ง ๆ ที่คณิตศาสตร์มีบทบาทสำคัญในการพัฒนาความคิด ส่งผลต่อความสามารถในการคิดสร้างสรรค์ คิดอย่างมีเหตุผล และคิดเป็นระบบ มีแบบแผน และสามารถวิเคราะห์ปัญหาและสถานการณ์ได้ถี่ถ้วนรอบคอบ ทำให้สามารถคาดการณ์วางแผน ตัดสินใจ และแก้ปัญหาได้อย่างถูกต้องเหมาะสม (ศึกษาธิการ, 2552) ยิ่งไปกว่านั้นเมื่อติดตามแนวโน้มในเวลาที่ผ่านไปพบว่านักเรียนส่วนใหญ่มีแนวโน้มการรู้เรื่องคณิตศาสตร์ลดลง (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554, 2557) ดังนั้นระบบการศึกษาควรต้องเร่งการยกระดับคุณภาพการเรียนรู้ของนักเรียน

ในการเร่งยกระดับคุณภาพนี้มีความจำเป็นต้องวินิจฉัยคุณลักษณะแฝงของนักเรียนที่มีความเกี่ยวข้องกับความสามารถทางคณิตศาสตร์ อาทิเช่น ในวิชาคณิตศาสตร์นักเรียนกลุ่มที่มีความสามารถในการให้เหตุผลสูงจะเป็นกลุ่มที่มีคะแนนสูงสุด และนักเรียนที่ขาดความสามารถในด้านการแก้ปัญหาจะเป็นกลุ่มที่มีคะแนนต่ำสุด ซึ่งในการศึกษาที่ผ่านมาจากผลการประเมิน PISA (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2557) มีการเปรียบเทียบผลของนักเรียนแยกตามกลุ่มโรงเรียนพบว่ากลุ่มนักเรียนโรงเรียนสาธิตมีคะแนนเฉลี่ยสูงกว่ากลุ่มนักเรียนโรงเรียน สพฐ. แต่ในความเป็นจริงแล้วการแบ่งในลักษณะดังกล่าวไม่ถูกต้องเสมอไป เพราะความสามารถแฝงของแต่ละคนอาจเหมือนกันในขณะที่อยู่ต่างองค์กรกัน และอาจจะแตกต่างกันทั้ง ๆ ที่อยู่องค์กรเดียวกัน (Tay, Diener, Drasgow, & Vermunt, 2011) นักเรียนกลุ่มโรงเรียน สพฐ. บางคนมีความสามารถทางคณิตศาสตร์ในระดับสูง และนักเรียนกลุ่มโรงเรียนสาธิตบางคนมีความสามารถทางคณิตศาสตร์อยู่ในระดับต่ำ หรือ นักเรียนบางกลุ่มของประเทศไทยอาจอยู่กลุ่มที่มีความสามารถสูงเทียบเท่ากับประเทศที่มีคะแนนเฉลี่ยสูงกว่าประเทศไทย และนักเรียนบางกลุ่มในประเทศที่มีคะแนนเฉลี่ยสูงกว่าประเทศไทยอาจมีความสามารถเทียบเท่าประเทศไทย ด้วยเหตุผลนี้จึงทำให้นักการศึกษา และนักสถิติวิเคราะห์ คิดค้นวิธีการหรือโมเดลการวัดที่สามารถวัดสิ่งที่แฝงอยู่ภายใน (unobserved variable) มากกว่าสิ่งที่เห็นอยู่ภายนอก (observed variable) เช่น โมเดลการวิเคราะห์กลุ่มแฝง (Latent Class Model)

การวิเคราะห์กลุ่มแฝง เป็นวิธีในการจำแนกคุณลักษณะหรือความสามารถของกลุ่มบุคคล เมื่อให้ตัวแปรกลุ่มบุคคล เป็นตัวแปรแฝงที่สัมพันธ์กับตัวแปรสังเกตได้ ซึ่งเป็นตัวแปรจัดประเภท จากงานวิจัยของ Rost (1990) ได้ศึกษาเกี่ยวกับการวิเคราะห์กลุ่มแฝง และนำแนวคิดของทฤษฎีการตอบสนองข้อสอบไปใช้ในการวิเคราะห์กลุ่มแฝง โดยบูรณาการทั้งสองโมเดลเป็นโมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT Model) ซึ่งทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถที่มีอยู่ภายในบุคคลกับผลการตอบข้อสอบ หรือข้อคำถามโดยใช้โค้งลักษณะข้อสอบ ทฤษฎีการตอบสนองข้อสอบสามารถอธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในกับพฤติกรรมการตอบสนองข้อสอบของบุคคลว่ามีโอกาสตอบถูกมากน้อยเพียงใด ตามทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองข้อสอบ (ถูกหรือผิด) ของผู้สอบเป็นสิ่งที่สังเกตได้โดยตรง ซึ่งถูกกำหนดโดยคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในบุคคล ที่ไม่สามารถสังเกตได้โดยตรง ซึ่งเป็นองค์ความรู้ที่สามารถเชื่อมโยงโอกาสในการตอบข้อสอบได้ถูกต้องเข้ากับความสามารถของผู้ตอบ และคุณลักษณะของข้อสอบ (ศิริชัย กาญจนวาสี, 2555) จากการบูรณาการทฤษฎีการตอบสนองข้อสอบกับการวิเคราะห์กลุ่มแฝงนี้ ทำให้สามารถจำแนกกลุ่มคุณลักษณะ หรือความสามารถของกลุ่มบุคคล เมื่อตัวแปรกลุ่มบุคคลเป็นตัวแปรแฝงที่สัมพันธ์กับตัวแปรสังเกตได้แบบต่อเนื่อง ซึ่งการจำแนกกลุ่มนี้เป็นประโยชน์ทำให้เราทราบได้ว่าลักษณะภายในของแต่ละคนเป็นอย่างไร เพื่อทำการวิจัยได้ถูกต้องมากขึ้น (Mislevy & Huang, 2007)

โดยทั่วไปของการจัดการองค์กรทางการศึกษามีหลายระดับที่ลดหลั่นกัน และสอดแทรกกันอยู่ เช่น นักเรียนจะถูกสอนร่วมกันเป็นชั้นเรียนภายในโรงเรียน โรงเรียนอยู่ในแต่ละสังกัด แต่ละภูมิภาค แต่ละอำเภอแตกต่างกันไป ข้อมูลทางการศึกษาจึงมักเป็นข้อมูลที่เป็นหลายระดับ (multilevel data) สอดแทรกกันอยู่ (hierarchical nested data) ซึ่งเป็นผลมาจากที่ระบบทางการศึกษาเป็นระบบที่มีความซับซ้อน การบริหารงานมีโครงสร้างเป็นหน่วยที่ซ้อนกันเป็นลำดับชั้น โดยหน่วยที่อยู่ระดับล่างสุดได้รับอิทธิพลร่วมกันจากตัวแปรที่อยู่ในระดับที่สูงกว่า (Waller & Meehl, 1998) Clogg (1988) ได้ศึกษาอิทธิพลของตัวแปรภายนอกที่มีอิทธิพลต่อการเรียนรู้ของผู้เรียน และพบว่าความถนัดมีอิทธิพลปฏิสัมพันธ์กับตัวแปรความแตกต่างระหว่างห้องเรียนมากกว่าในระดับบุคคล ซึ่งแตกต่างจากงานวิจัยก่อนหน้าที่พบว่ามีความสัมพันธ์ของตัวแปรความถนัดกับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ว่ามีปฏิสัมพันธ์มากกับผู้เรียน ซึ่งในอดีตมีข้อจำกัดเกี่ยวกับสถิติวิเคราะห์ทำให้มีความคลาดเคลื่อนของการสรุปผลการวิจัยเกิดขึ้น ดังนั้นหากเกิดความผันแปรขึ้นไปตามบริบทของโรงเรียนหรือห้องเรียน การวิเคราะห์ในแบบเดิมที่มีการกำหนดให้ตัวแปรต่างระดับเสมือนอยู่ในระดับเดียวกัน อาจจะทำให้ผลการวิเคราะห์เกิดความผิดพลาดเพราะไม่สนใจความแตกต่างระหว่างหน่วยในการวิเคราะห์ ซึ่งไม่สามารถคำนวณค่าความแปรปรวนภายในหน่วยหรือกลุ่ม (within group variability) จึงเป็นการละเลยต่อการศึกษาปฏิสัมพันธ์ระหว่างตัวแปรที่อยู่ต่างระดับกัน ถึงแม้จะมี

การออกแบบการวิจัยอย่างรอบคอบก็ตาม ดังนั้นในการทดสอบ หรือประเมินผลทางการศึกษาควรคำนึงถึงโครงสร้างข้อมูลที่มีลักษณะลดหลั่น ด้วยเหตุผลนี้ Cho และ Cohen (2010) จึงได้มีการนำเสนอโมเดลการวัดที่เรียกว่า โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT: MMixIRT) เพื่อแก้ปัญหาดังกล่าว

ระยะแรกโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ได้ถูกนำไปประยุกต์ใช้กับการตรวจสอบการทำหน้าที่ต่างกันของแบบวัด หรือข้อสอบ เพื่อตรวจสอบคุณภาพของแบบวัดหรือข้อสอบ ว่ามีความโน้มเอียงไปยังกลุ่มแฝงกลุ่มใดกลุ่มหนึ่งหรือไม่ ซึ่งโมเดลที่ Cho และ Cohen เสนอนี้เป็นโมเดลที่ประยุกต์ใช้กับโมเดลเอกมิติ มีการบูรณาการโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมที่มี 1 พารามิเตอร์ คือพารามิเตอร์ความยากของข้อสอบ ต่อมา Tay และคณะ (2011) ได้เสนอโมเดลการวัดแบบผสมพหุระดับ (Multilevel Mix-Measurement IRT : MMMIRT) ซึ่งเป็นโมเดลในลักษณะที่คล้ายกันแต่เป็นโมเดลที่ประยุกต์ใช้กับโมเดลพหุมิติ มีการบูรณาการโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมที่มี 2 พารามิเตอร์ คือพารามิเตอร์ความยากของข้อสอบ และอำนาจจำแนกของข้อสอบ และนำไปประยุกต์ใช้กับการจัดกลุ่มแฝงของบุคคลในองค์กรตามการรายงานอารมณ์ของตนเองระดับบุคคล และในระดับข้ามประเทศ (Tay, Diener, Drasgow, & Vermunt, 2011) ในการประเมินผู้เรียนที่ผ่านมายังไม่มีการเปรียบเทียบกันระดับกลุ่มผู้เรียนว่าแต่ละคนอยู่ในระดับใดเมื่อเทียบกับผู้เรียนคนอื่น และไม่มีการเปรียบเทียบข้ามระดับประเทศ หรือระดับโรงเรียน ดังนั้นการจำแนกกลุ่มความสามารถที่สัมพันธ์กับคุณลักษณะแฝงของแต่ละบุคคลแบบข้ามระดับจะทำให้เกิดผลการประเมินที่ชัดเจนขึ้น

จากที่กล่าวข้างต้นนี้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จึงน่าจะเหมาะสมกับการจำแนกนักเรียนหรือบุคคลว่ามีคุณลักษณะแฝงในกลุ่มแฝงใด ซึ่งเป็นการวัดระดับองค์กร หรือระดับโรงเรียน และเป็นประโยชน์ในการวินิจฉัยระดับกลุ่ม ซึ่งผลจากการวินิจฉัยจะเป็นส่วนหนึ่งของการตัดสินใจในการออกแบบการเรียนการสอน ซึ่งสามารถเชื่อมโยงระหว่างความแตกต่างของนักเรียนเป็นรายกลุ่มแฝง หรือรายบุคคล และสามารถเชื่อมโยงถึงอิทธิพลที่อาจทำให้มีความเสี่ยงในการออกแบบการเรียนการสอนที่ผิดพลาดได้ต่อไป

ในการวิเคราะห์จำแนกกลุ่มนักเรียนที่มีความสามารถต่างกันสามารถจำแนกโดยใช้การทดสอบในโครงการประเมินผลนักเรียนระดับนานาชาติ (PISA) การทดสอบในโครงการการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระดับนานาชาติ (TIMSS) หรือการทดสอบทางศึกษาแห่งชาติขั้นพื้นฐาน (O – NET) ซึ่งเป็นการทดสอบมาตรฐาน แต่จากลักษณะของการทดสอบมาตรฐานในปัจจุบันส่วนใหญ่เป็นแบบเลือกตอบ จึงมีจุดอ่อนที่สำคัญคือเปิดโอกาสให้มีการตอบถูกโดยการเดา ปัญหาในการตอบแบบเดาสุ่มจึงเป็นแหล่งความคลาดเคลื่อนในการวัดที่สำคัญ ซึ่งจะส่งผลให้ผลการวัดที่ได้ไม่สามารถให้สารสนเทศที่เพียงพอในการจำแนกผู้ตอบในระดับต่าง ๆ ของความรู้ระหว่างผู้ที่รู้จริง ผู้ที่มี

ความรู้บางส่วน หรือผู้ที่ไม่รู้ นอกจากนี้การเดาข้อสอบอาจจะส่งผลให้ผู้สอบในบางกลุ่มมีโมเดลการวัดแตกต่างจากที่ไม่มีการเดาข้อสอบ ทั้งนี้โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมที่ผ่านมายังไม่สามารถลดความคลาดเคลื่อนเรื่องการเดาข้อสอบ และไม่คำนึงถึงความแตกต่างของบุคคลข้ามระดับ เนื่องจากลักษณะของการจัดการศึกษาเป็นข้อมูลหลายระดับลดหลั่นกัน ซึ่งในแต่ละบุคคลอาจมีโมเดลการวัดแตกต่างกันข้ามระดับ หรือในระดับเดียวกัน จึงจำเป็นต้องมีการพัฒนาโมเดลตอบสนองข้อสอบแบบผสมพหุระดับ ที่มีการลดความคลาดเคลื่อนในเรื่องการเดาข้อสอบ ซึ่งเป็นการรวมกันของโมเดลพหุระดับ (Multilevel model) โมเดลกลุ่มแฝง (Latent class model) และ โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3PL IRT model) ซึ่งจะเป็นประโยชน์มาก เนื่องจากรูปแบบการทดสอบในประเทศไทยส่วนใหญ่เป็นแบบเลือกตอบ ในการพัฒนาโมเดลนี้ ผู้วิจัยตรวจสอบโมเดลด้วยการจำลองข้อมูล จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวกับการวิเคราะห์ข้อสอบ ระบุว่าจำนวนข้อสอบในการวิเคราะห์ข้อมูลส่งผลต่อการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์ข้อมูลด้วยโมเดลการตอบสนองของข้อสอบ (Lord, 1986; Swaminathan & Gifford, 1982, 1985, 1986; เกศมณี พยัคฆ์, 2543; รัตนา ศรีเหรัญ, 2539; วิชุดา บัวคง, 2532) และพบว่า ถ้าจำนวนข้อสอบเพิ่มขึ้น จะทำให้มีการประมาณค่าพารามิเตอร์มีความถูกต้องเพิ่มมากขึ้น ดังนั้นผู้วิจัยจึงเลือกเงื่อนไขในการจำลองข้อมูล ในครั้งนี้ โดยพิจารณาจาก จำนวนข้อสอบเป็นเงื่อนไขในการจำลองข้อมูล การพัฒนาโมเดลดังกล่าวจะก่อให้เกิดองค์ความรู้ใหม่เกี่ยวกับสถิติวิเคราะห์ที่สามารถใช้ในการจำแนกความแตกต่างของแต่ละบุคคลได้ ซึ่งเป็นประโยชน์ต่อ การวิจัย การพัฒนาการศึกษาต่อไป

1.2 คำถามวิจัย

1. โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการของ โมเดลกลุ่มแฝง โมเดลพหุระดับ และโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ควรเป็นอย่างไร

2. โมเดลการประมาณค่าของทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการ โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ ที่พัฒนาขึ้น เมื่อนำไปใช้การจำแนกกลุ่มคุณลักษณะแฝงของบุคคลในข้อมูลจำลอง จะมีความถูกต้องแม่นยำเพียงใด

3. โมเดลการประมาณค่าของทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการ โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ ที่พัฒนาขึ้น เมื่อนำไปใช้การจำแนกกลุ่มคุณลักษณะแฝงของบุคคลในข้อมูลเชิงประจักษ์ จะได้ผลเป็นอย่างไร

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.3 วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาโมเดลการประมาณค่ากลุ่มความสามารถแฝงของบุคคลตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ

2. เพื่อตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลจำลอง

3. เพื่อวิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

1.4 ขอบเขตการวิจัย

1. การประมาณค่าพารามิเตอร์ความสามารถของผู้สอบในครั้งนี้ ใช้โมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3 PL IRT) เป็นการขยายองค์ความรู้จากที่มีอยู่เดิมเป็นแบบ 2 พารามิเตอร์ และ แบบ 1 พารามิเตอร์ เพื่อเป็นประโยชน์ในการวิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล

2. ในงานวิจัยนี้ศึกษากับข้อมูลที่มีการให้คะแนนแบบสองค่า (Dichotomous Scoring) โดยทำการจำลองผลการตอบข้อสอบตามโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture Item Response Theory Model) สำหรับจำแนกกลุ่มแฝง

3. ข้อมูลที่ใช้ในการศึกษาครั้งนี้ ใช้ข้อมูลจำลองภายใต้เงื่อนไขขนาดของโรงเรียนและความสามารถในการตอบข้อสอบของนักเรียน ที่ได้จากโปรแกรม R แพคเกจ IRTOYS รุ่น 0.1.7 ในงานวิจัยนี้มีเงื่อนไขในการจำลองข้อมูลคือ กลุ่มแฝงระดับโรงเรียน 6 กลุ่มแฝง ประกอบด้วย กลุ่มที่คละความสามารถ 3 กลุ่ม กลุ่มที่มีความสามารถปานกลางในทุกสาระ 1 กลุ่ม กลุ่มที่มีความสามารถเก่งทุกสาระ 1 กลุ่ม และกลุ่มที่มีความสามารถอ่อนทุกสาระ 1 กลุ่ม โดยแต่ละกลุ่มแฝงจำลองผลการตอบข้อสอบสองสถานการณ์ คือ จำนวนข้อสอบ 6 ข้อ และ 15 ข้อ ของนักเรียนในโรงเรียน 4 ขนาด ได้แก่ โรงเรียนขนาดเล็ก โรงเรียนขนาดกลาง โรงเรียนขนาดใหญ่ และโรงเรียนขนาดใหญ่พิเศษ

4. การประมาณค่าพารามิเตอร์ ในงานวิจัยนี้ประมาณด้วยวิธีการประมาณค่าแบบ Empirical Bayes โดยใช้การเขียนคำสั่งประมวลผลด้วยแพ็คเกจ ltm รุ่น 1.0-0 ในโปรแกรม R และในการวิเคราะห์กลุ่มแฝงใช้วิธี EM algorithm ประเมินความเหมาะสมของการจำแนกกลุ่ม

ความสามารถแฝงโดยใช้ Likelihood ratio test และดัชนีของเบส์เซียน (BIC) โดยใช้ แพ็คเกจ Mclust รุ่น 5.0.1 โปรแกรม R สำหรับงานวิจัยนี้การประมาณค่าพารามิเตอร์จะกำหนดให้มีการสุ่มค่าเริ่มต้นจำนวน 10 ค่า แล้วเลือกค่าพารามิเตอร์ a, b, c และ θ ที่ทำให้ค่า log likelihood มีค่าสูงสุด มาใช้ในการคำนวณต่อไป

5. ข้อมูลเชิงประจักษ์ที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลทุติยภูมิจากฐานข้อมูลของสำนักทดสอบทางการศึกษาแห่งชาติ (สทศ.) โดยเป็นผลการตอบข้อสอบของการทดสอบ O-NET วิชาคณิตศาสตร์ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2555 และดำเนินการสุ่มตัวอย่างระดับโรงเรียน ใช้วิธีการสุ่มโรงเรียนแบบแบ่งชั้น (stratified random sampling) โดยในทุกสังกัดของโรงเรียน จะใช้ขนาดของโรงเรียนเป็นชั้นของการสุ่ม 4 ขนาด คือ ขนาดเล็ก ขนาดกลาง ขนาดใหญ่ และขนาดใหญ่มาก

1.5 นิยามศัพท์ที่ใช้ในการวิจัย

1. โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ หมายถึงรูปแบบการวิเคราะห์ข้อมูลสำหรับการจำแนกกลุ่มความสามารถแฝง และระบุความแตกต่างของบุคคลจากตัวแปรสาเหตุที่ไม่สามารถสังเกตได้ เมื่อตัวแปรผลลัพธ์เป็นตัวแปรต่อเนื่อง โดยการบูรณาการโครงสร้างโมเดลพหุระดับ (Multilevel model) เข้ากับโมเดลการตอบสนองข้อสอบแบบผสม ซึ่งโมเดลการตอบสนองข้อสอบแบบผสม เป็นโมเดลที่มีการผสมโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และโมเดลกลุ่มแฝง โดยลักษณะของข้อมูลที่นำมาวิเคราะห์ เป็นข้อมูลผลการตอบข้อสอบของนักเรียนซึ่งสอดแทรกอยู่ในโรงเรียน และมีตัวแปรตามอยู่ในระดับบุคคล (individual class) โดยผลการวิเคราะห์ในระดับหน่วยระดับบุคคล นำมาใช้ในการวิเคราะห์ในระดับโรงเรียน ทำให้สามารถอธิบายอิทธิพลที่ส่งผลของตัวแปรที่อยู่ในระดับที่ต่างกันได้

2. กลุ่มความสามารถแฝง หมายถึง ความสามารถแท้จริงที่ไม่สามารถสังเกตได้โดยตรง แสดงถึงความแตกต่างของแต่ละบุคคล โดยกลุ่มความสามารถแฝงแบ่งออกเป็น 2 ส่วน คือกลุ่มความสามารถแฝงระดับบุคคล และระดับโรงเรียน โดยกลุ่มความสามารถแฝงระดับบุคคล คือระดับความสามารถที่แท้จริงรายบุคคลที่ส่งผลต่อความน่าจะเป็นในการตอบข้อสอบถูกในแต่ละข้อตามทฤษฎีการตอบสนองข้อสอบ โดยประมาณค่าพารามิเตอร์ และความสามารถโดยวิธีการ Empirical Bayes และกลุ่มความสามารถแฝงระดับโรงเรียน คือระดับความสามารถที่แท้จริงรายโรงเรียนที่ประมาณค่าได้จากสัดส่วนของจำนวนกลุ่มแฝงในระดับบุคคล ที่สอดแทรกอยู่ในแต่ละโรงเรียน โดยจำแนกกลุ่มโรงเรียนตามวิธีการ EM algorithm

3. การจำแนกกลุ่มความสามารถแฝง หมายถึง การจำแนกประชากรออกเป็นกลุ่มตามหลักการวิเคราะห์กลุ่มแฝง ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ซึ่งโมเดลทางคณิตศาสตร์ที่ใช้ในการจำแนกกลุ่มความสามารถแฝงสามารถแบ่งออกเป็น 2 ส่วน คือการจำแนกในระดับบุคคล และการจำแนกกลุ่มความสามารถแฝงในระดับโรงเรียน โดยในระดับบุคคลกลุ่มแฝงจะถูกแบ่งตามความน่าจะเป็นในการตอบถูกในแต่ละข้อโดยอ้างอิงจากโมเดลทฤษฎีการตอบสนองข้อสอบแบบสามพารามิเตอร์ (3PL-IRT) ส่วนการจำแนกกลุ่มความสามารถแฝงในระดับโรงเรียนพิจารณาจากสัดส่วนความสามารถของกลุ่มแฝงในระดับบุคคลที่สอดแทรกอยู่ในแต่ละโรงเรียน โดยโรงเรียนที่มีสัดส่วนความสามารถแฝงระดับบุคคลใกล้เคียงกันจะถูกจัดอยู่กลุ่มแฝงเดียวกัน

4. ความถูกต้องแม่นยำของการจำแนกกลุ่มความสามารถแฝง หมายถึง ความสอดคล้องของผลการจำแนกกลุ่มความสามารถแฝงที่จำลองตามระดับความสามารถของผู้สอบ กับผลของการจำแนกกลุ่มความสามารถแฝงที่ได้จากโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น โดยการคำนวณค่าดัชนีเกณฑ์สารสนเทศของเบส์เซียน (BIC) ซึ่งเป็นดัชนีที่ใช้กำหนดจำนวนกลุ่มแฝงที่เหมาะสม

5. ดัชนีเกณฑ์สารสนเทศของเบส์เซียน (Bayesian information criterion: BIC) คือ ดัชนีที่ใช้ในการบอกถึงความเหมาะสมของจำนวนกลุ่มที่ถูกจำแนก ค่า BIC มีที่มาจากการเปรียบเทียบค่า Maximum Likelihood ระหว่าง 2 โมเดล ซึ่งหลายโปรแกรมแสดงให้เห็นว่าหากค่า BIC มีค่าน้อยที่สุดจำนวนกลุ่มที่ถูกแบ่งออกมาในกรณีดังกล่าวจะมีความเหมาะสมมากที่สุด แต่สำหรับในงานวิจัยนี้ได้ใช้ แพ็คเกจ MCLUST ใน โปรแกรม R ซึ่งมีการคำนวณความเหมาะสมของจำนวนกลุ่มต่างจากโปรแกรมอื่น โดยในโปรแกรมนี้กำหนดว่าหากค่า BIC มีค่าสูงสุดแสดงว่าจำนวนกลุ่มที่ถูกแบ่งออกมาในกรณีดังกล่าวจะมีความเหมาะสมมากที่สุด (Haughton, Legrand, & Woolford, 2012)

1.6 ประโยชน์ที่ได้รับ

ประโยชน์เชิงวิชาการ

1. ได้รูปแบบวิธีการวิเคราะห์การจำแนกกลุ่มความสามารถแฝงโดยใช้โมเดลการวัดตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งเป็นประโยชน์ในการลดความคลาดเคลื่อนในการวัดที่เกิดจากการเดาข้อสอบ และเป็นประโยชน์ต่อวงการศึกษา ในการวิเคราะห์กลุ่มแฝง

2. ได้ต้นแบบการวิจัยที่ใช้การวิเคราะห์การจำแนกกลุ่มคุณลักษณะแฝง ในบริบททางการศึกษา

ประโยชน์เชิงปฏิบัติ

ผลที่ได้จากการจำแนกกลุ่มความสามารถแฝงจากข้อมูลเชิงประจักษ์นี้ได้สารสนเทศทั้งระดับโรงเรียนและระดับนักเรียน

สำหรับระดับโรงเรียน ผลการจำแนกกลุ่มความสามารถแฝงทำให้ทราบว่ามีโรงเรียนใดบ้างอยู่ในกลุ่มแฝงเดียวกัน และโรงเรียนที่อยู่ในกลุ่มแฝงเดียวกันมีลักษณะใดสอดคล้องกัน และโรงเรียนที่อยู่ในกลุ่มแฝงที่มีความสามารถสูงมีลักษณะอย่างไร กลุ่มแฝงที่มีความสามารถต่ำมีลักษณะอย่างไร

สำหรับระดับนักเรียน ผลการจำแนกกลุ่มความสามารถแฝงทำให้ทราบลักษณะว่านักเรียนที่มีลักษณะการตอบแบบใดอยู่ในกลุ่มแฝงเดียวกัน และกลุ่มแฝงนั้นมีลักษณะเด่นอย่างไร

ผลการจำแนกกลุ่มความสามารถแฝงนี้เป็นประโยชน์ต่อหน่วยงานที่เกี่ยวข้องกับการจัดการศึกษา ในการตัดสินใจ กำหนดนโยบาย วางแผนงบประมาณ ตลอดจนกำหนดแนวทางการส่งเสริมการพัฒนาระบบการศึกษา และพัฒนาโรงเรียน เพื่อส่งเสริมหรือลดอิทธิพลของตัวแปรที่มีผลต่อคุณภาพการจัดการศึกษา และเป็นประโยชน์ต่อผู้บริหารโรงเรียนและครู ในการกำหนดนโยบาย ทิศทางในการพัฒนาโรงเรียน แนวทางการจัดการเรียนการสอน เพื่อส่งเสริมให้โรงเรียน และนักเรียนมีผลสัมฤทธิ์ทางการเรียนในระดับที่สูงขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การนำเสนอสาระสำคัญของรายงานเกี่ยวกับแนวคิด ทฤษฎี และเอกสารรายงานการวิจัยที่เกี่ยวข้องกับวิธีการวิเคราะห์และการประมาณค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบทั้งจากการวิเคราะห์ข้อมูลจริง และข้อมูลจำลอง ผู้วิจัยได้ศึกษาค้นคว้าและสังเคราะห์จากหนังสือ เอกสาร บทความ รายงานการวิจัยทั้งในและต่างประเทศ รวมทั้งฐานข้อมูลการวิจัยที่สำคัญ ผู้วิจัยแบ่งการนำเสนอแนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้องออกเป็น 5 ตอน ซึ่งมีรายละเอียดในแต่ละตอนดังต่อไปนี้

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ

1.1 ทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT)

1.1.1 ทฤษฎีการตอบสนองข้อสอบ

1.1.2 โมเดลการวิเคราะห์กลุ่มแฝง

1.1.3 ทฤษฎีการตอบสนองข้อสอบแบบผสม

1.2 ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT)

1.2.1 การวิเคราะห์พหุระดับ

1.2.2 โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

1.3 การประมาณค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบ

1.4 การเลือกโมเดล (Model selection)

ตอนที่ 2 การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

2.1 ความเป็นมาของเทคนิคมอนติคาร์โล

2.2 ขั้นตอนการจำลองข้อมูล

2.3 การประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบ

2.4 ประสิทธิภาพของการจำลองข้อมูล

2.5 ข้อดีและข้อจำกัดของการศึกษาด้วยเทคนิคการจำลองข้อมูล

ตอนที่ 3 การทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O– NET)

3.1 ความสำคัญของการสอบ O-NET

3.2 ลักษณะทั่วไปของข้อสอบ O – NET

3.3 วิธีการคิดคะแนนของการสอบ O-NET วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6

3.4 ลักษณะข้อสอบ O-NET วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

4.1 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

4.2 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม

ตอนที่ 5 กรอบแนวคิดในการวิจัย

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT model: MMixIRT) ได้พัฒนาจากแนวคิดของโมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT model: MixIRT) ซึ่งเป็นการบูรณาการแนวคิดโมเดลการตอบสนองข้อสอบ (IRT model) กับ โมเดลกลุ่มแฝง (Latent Class model) แนวคิดโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) เกิดขึ้นโดยให้ความสำคัญกับมิติที่แตกต่างกันในเชิงคุณภาพของคุณลักษณะที่อยู่ระหว่างชั้น (between class) และมิติที่แตกต่างกันเชิงปริมาณที่อยู่ภายในชั้น (within class) (Hernández, Drasgow, & González-Romá, 2004) จึงได้มีการพัฒนาแนวคิดพหุระดับ (multilevel) ขึ้น

ในองค์กร หรือองค์กรทางการศึกษา มีการใช้แนวคิดพหุระดับในหลายมิติ เช่น การวิเคราะห์องค์ประกอบ (factor analysis) และ ทฤษฎีการตอบสนองข้อสอบ (item response theory) โดยมีจุดประสงค์สำคัญ คือ การวัดโครงสร้าง (construct measurement) และการตรวจสอบความเท่าเทียมกันของการวัด (measurement equivalence) ระหว่างกลุ่ม แนวคิดเกี่ยวกับการวัดแบบดั้งเดิมจะวัดความแตกต่างจากข้อมูลที่สังเกตได้ โดยการสมมุติให้ประชากรมีโมเดลการวัดเดียวกัน อย่างไรก็ตาม การพัฒนาวิทยาการ และองค์ความรู้ ทำให้มีการสังเคราะห์แนวคิดพหุมิติ (dimensional) และกลุ่มแฝง (latent class) โดยจำแนกกลุ่มความสามารถแฝงจากโมเดลการวัด (measurement models) ที่แตกต่างกัน (Rost, 1990, 1991, 1997)

สำหรับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT: MMix-IRT) นำเสนอโดยแบ่งออกเป็น 4 ประเด็นดังนี้

1.1 ทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT) เพื่อทำความเข้าใจพัฒนาการของทฤษฎี ซึ่งโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT model) เป็นการบูรณาการของ

โมเดลการตอบสนองข้อสอบ และโมเดลกลุ่มแฝง ดังนั้นจะกล่าวถึงรายละเอียดของโมเดลการตอบสนองข้อสอบ และ โมเดลกลุ่มแฝง

1.2 ทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) เป็นการบูรณาการโมเดลการตอบสนองข้อสอบแบบผสม กับโมเดลพหุระดับ รายละเอียดในส่วนนี้จะแบ่งออกเป็น 4 ประเด็นด้วยกัน กล่าวคือ

ประเด็นแรกจะกล่าวถึงโมเดลพหุระดับ (Multilevel model) และโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ซึ่งเป็นการขยายองค์ความรู้จากโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) โดยการสร้างแบบจำลองกลุ่มแฝงระดับลดหลั่น (hierarchical classes) ภายใต้คุณลักษณะที่ระดับล่าง เป็นการบูรณาการแนวคิดโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) กับ โมเดลพหุระดับ (Multilevel model)

ประเด็นที่สอง เสนอข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT)

ประเด็นที่สาม กล่าวถึงการประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) และการจำลองข้อมูล

ประเด็นที่สี่ ประเด็นของความแตกต่างของประชากร ซึ่งในการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ มีความจำเป็นต้องเข้าใจเรื่องความแตกต่างที่สังเกตได้ และความแตกต่างที่สังเกตไม่ได้ด้วย

1.3 การประมาณค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบ ซึ่งกล่าวถึงการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood estimation) และการประมาณค่าแบบเบส์ (Baysian estimation) ซึ่งทั้งสองวิธีนี้เป็นวิธีการประมาณค่าที่เหมาะสมกับข้อมูลพหุระดับ

1.4 การเลือกโมเดล กล่าวถึงการเลือกโมเดลที่สอดคล้องจากดัชนีที่ใช้ในการกำหนดจำนวนกลุ่มแฝง

1.1 ทฤษฎีการตอบสนองข้อสอบแบบผสม (Mixture IRT : MixIRT)

ในทฤษฎีการตอบสนองข้อสอบ กลุ่มคุณลักษณะแฝงของบุคคลถูกประมาณจากการตอบสนองข้อสอบ โมเดลทฤษฎีการตอบสนองข้อสอบ (IRT model) ระบุว่ากลุ่มคุณลักษณะแฝงและคุณสมบัติของข้อสอบสัมพันธ์กับการตอบสนองข้อสอบของบุคคลอย่างไร กลุ่มคุณลักษณะแฝงถูกประมาณในบริบทของโมเดลทฤษฎีการตอบสนองข้อสอบ ดังนั้นทฤษฎีการตอบสนองข้อสอบเป็นการวัดที่มีโมเดลเป็นพื้นฐาน แต่สำหรับโมเดลกลุ่มแฝงจะอธิบายความสัมพันธ์ระหว่างการตอบสนองของค่าที่สังเกตได้ ซึ่งการวิเคราะห์กลุ่มแฝงจะเป็นการวิเคราะห์จัดกลุ่มที่เป็นพฤติกรรมหรือคุณลักษณะ

ร่วมกัน ซึ่งเป็นตัวแปรไม่ต่อเนื่อง ดังนั้นในการบูรณาการกันของโมเดลกลุ่มแฝง และโมเดลทฤษฎีการตอบสนองข้อสอบจึงเป็นประโยชน์ในการจัดกลุ่มตัวแปรคุณลักษณะแฝงที่เป็นตัวแปรต่อเนื่อง (Collins & Lanza, 2010) ในหัวข้อนี้ ผู้วิจัยนำเสนอรายละเอียดเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ โมเดลการวิเคราะห์กลุ่มแฝง และทฤษฎีการตอบสนองข้อสอบแบบผสม ตามลำดับ

1.1.1 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มีความแตกต่างจากการวิเคราะห์องค์ประกอบ คือ การวิเคราะห์องค์ประกอบ ซึ่งกำหนดจำนวนที่น้อยที่สุดขององค์ประกอบที่สามารถลดข้อมูลในเมทริกซ์การตอบข้อสอบได้ สำหรับการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบได้สร้างความสัมพันธ์ระหว่างคุณลักษณะของผู้สอบและลักษณะของข้อสอบ ทฤษฎีการตอบสนองข้อสอบ เป็นเทคนิคทางสถิติที่ใช้ในการทดสอบทางการศึกษาและจิตวิทยาอย่างกว้างขวาง (Bergman & Magnusson, 1997) ซึ่งสามารถจัดการกรอบแนวคิดที่มีประโยชน์สำหรับวิเคราะห์การตอบสนองข้อสอบ โดยยืนยันความสัมพันธ์ระหว่างการตอบสนองข้อสอบแต่ละข้อ และตัวแปรแฝง ซึ่งบางครั้งเรียก คุณลักษณะแฝง หรือ ความสามารถ ความสัมพันธ์ระหว่างการตอบสนองข้อสอบและคุณลักษณะแฝงถูกระบุโดย หน้าที่การตอบสนองข้อสอบ (item response function) หรือโค้งคุณลักษณะข้อสอบ(item characteristic curve : ICC) โดยทั่วไป ความแตกต่างของโมเดลทฤษฎีการตอบสนองข้อสอบ มีความแตกต่างทางสูตรคณิตศาสตร์ในการตัดสินโค้งคุณลักษณะข้อสอบ กับสัมประสิทธิ์ของโมเดลที่สอดคล้องกับพารามิเตอร์คุณลักษณะข้อสอบ สำหรับรายละเอียดเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยนำเสนอแยกเป็นหัวข้อดังต่อไปนี้ (1) ข้อตกลงเบื้องต้น (2) ประเภทของโมเดลทฤษฎีการตอบสนองข้อสอบ (3) โมเดลการตอบสนองข้อสอบสำหรับข้อสอบที่มีการให้คะแนนสองค่า ซึ่งนำเสนอโมเดลโลจิสติกแบบ 3 พารามิเตอร์ 2 พารามิเตอร์ และ 1 พารามิเตอร์ ตามลำดับ และฟังก์ชันสารสนเทศของข้อสอบ

(1) ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ในหลายกรณีของโมเดลทฤษฎีการตอบสนองข้อสอบ มีข้อตกลงเบื้องต้นที่สำคัญดังนี้ (Hambleton & Swaminathan, 1985)

1. ความเป็นอิสระของการตอบข้อสอบ (local independence) หมายถึงการที่ผู้สอบมีความเป็นอิสระ ในการตอบข้อสอบและเป็นความสัมพันธ์กับข้อคำถามหรือผู้ตอบที่เหมาะสมกับโมเดลทฤษฎีการตอบสนองข้อสอบ ซึ่งความอิสระของการตอบข้อสอบนี้มีความจำเป็นที่จะต้องมีทั้งในแบบทดสอบที่เป็นเอกมิติ และพหุมิติ (unidimensional และ multidimensional)

2. โค้งคุณลักษณะข้อสอบ (ICC) จะมีรูปแบบที่เป็นลักษณะเฉพาะซึ่งจะบรรยายถึงความสัมพันธ์ที่จะเปลี่ยนแปลงไปของระดับความสามารถภายในของผู้สอบแต่ละบุคคล กับการ

เปลี่ยนแปลงของโอกาสหรือความน่าจะเป็นที่จะตอบถูกต้อง ซึ่งโค้งคุณลักษณะข้อสอบโดยทั่วไปจะมีลักษณะเป็นรูปตัวเอส (s-shape) และสามารถบ่งชี้พารามิเตอร์ต่าง ๆ ได้

นอกจากนี้ Hambleton (1991) ได้เพิ่มข้อตกลงเบื้องต้นดังนี้คือ ความเป็นเอกมิติของคุณลักษณะภายในที่มุ่งวัด หมายถึงคุณลักษณะภายในเป็นสิ่งที่กำหนดพฤติกรรมการตอบข้อสอบเพียงมิติเดียว นั่นคือ ข้อสอบแต่ละข้อมีคุณสมบัติในการวัดคุณลักษณะเดียวกัน ซึ่งคุณสมบัติข้อนี้ต่อมาในภายหลัง ทฤษฎีการตอบสนองข้อสอบสามารถประมาณค่าได้ในลักษณะพหุมิติ (multidimensional) เช่น โมเดลพหุมิติของราสช์ (multidimensional Rasch model)

(2) ประเภทของโมเดลทฤษฎีการตอบสนองข้อสอบ ในที่นี้จำแนกประเภทจากการให้คะแนน ซึ่งแบ่งออกเป็น

ประเภทที่ 1 โมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบมากกว่าสองค่า (models for polytomous item) Thissen และ Steinberg (1986) สรุปโมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบมากกว่าสองค่าตามคุณลักษณะเป็น 2 ประเภท คือ difference models และ divided-by-total models

สำหรับ difference models ได้แก่ (1) Graded response model เสนอโดย Samejima (1969) เป็นโมเดลทั่วไปของโมเดลการตอบสนองข้อสอบที่มี 2 พารามิเตอร์ ใช้กับแบบสอบหรือแบบวัดที่แต่ละข้อคำถามมีรายการตำตอบแบบมาตรเรียงลำดับ (Ordered categorical responses) ซึ่งแต่ละข้ออาจมีจำนวนรายการคำตอบที่แตกต่างกันได้ เช่น มาตรประมาณค่า (rating scale) ที่ไม่จำเป็นต้องมีรายการคำตอบเท่ากันทุกข้อ และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบแบบทางอ้อม (Indirect IRT Method) โดยใช้โปรแกรม MULTILOG (D Thissen, Chen, & Bock, 2003) (2) rating scale model ตามแนวคิดของ Rijmen and De Boeck (2003) โมเดลนี้คล้ายกับ Partial Credit Model (PCM) อยู่บนพื้นฐานของโมเดลการตอบสนองข้อสอบที่มี 1 พารามิเตอร์ และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method) ซึ่งถูกกำหนดความน่าจะเป็นของคะแนนที่เป็นแบบจัดกลุ่ม

สำหรับ divided-by-total models ได้แก่ (1) partial credit model เสนอโดย Masters (1982) เป็นโมเดลที่พัฒนาต่อจากโมดลการตอบสนองข้อสอบที่มี 1 พารามิเตอร์ สำหรับใช้ในการวิเคราะห์ข้อสอบหรือข้อคำถามที่มีกระบวนการตอบหลายลำดับขั้น ซึ่งจำเป็นต้องมีการตรวจให้คะแนนการตอบถูกต้องหรือการตอบถูกบางส่วนในแต่ละลำดับขั้นของกระบวนการตอบ เช่น ข้อสอบวัดผลสัมฤทธิ์ แบบวัดเจตคติ เชาว์ปัญญา เป็นต้น และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method) โดยใช้โปรแกรม PARSCALE (B. Muthén, 2004) และ (2) general partial credit model เสนอโดย Muraki

(1992) เป็นโมเดลทั่วไปของ PCM โดยยอมให้ข้อคำถามแต่ละข้อสามารถมีค่าพารามิเตอร์ความชันแตกต่างกันได้ และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method) ความน่าจะเป็นของการตอบสนองในกลุ่มถูกกำหนดโดยการคำนวณ โดยใช้โปรแกรม PARSCALE (B. Muthén, 2004)

นอกจากนี้ยังมีโมเดลการตอบสนองข้อสอบสำหรับตรวจให้คะแนนมากกว่า 2 ค่า โมเดลอื่นได้แก่ (1) Modified Graded Response Model (M-GRM) พัฒนาโดย Muraki (1990) เป็นโมเดลเฉพาะของโมเดล GRM โดยสามารถนำไปใช้กับข้อคำถามที่มีจำนวนรายการคำตอบที่เท่ากันหรือมีรูปแบบการตอบที่คงที่สำหรับทุกข้อคำถาม Muraki ได้ปรับปรุงโมเดล GRM เพื่อให้สะดวกแก่การใช้กับแบบวัดประเภทมาตรประมาณค่า (rating scale) เช่น รูปแบบการตอบเป็นมาตรประมาณค่า 5 ระดับของลิเคิร์ท (Likert rating scale) และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบแบบทางอ้อม (Indirect IRT Method) โดยใช้โปรแกรม PARSCALE (Muraki, 1993) (2) *Nominal Response Model* (NRM) พัฒนาโดย Formann (1992) เป็นโมเดลทั่วไปที่ใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบโดยตรงแบบขั้นตอนเดียว (Direct IRT Method) ใช้วิเคราะห์ข้อสอบหรือข้อคำถามที่รายการคำตอบไม่จำเป็นต้องถูกจัดเรียงลำดับ เช่น อกข้อสอบแบบหลายตัวเลือ (Multiple choices) ข้อคำถามวัดเจตคติ ข้อคำถามสำหรับประเมินบุคลิกลักษณะ เป็นต้น

ประเภทที่ 2 โมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบสองค่า (dichotomous IRT models) ซึ่งถูกจำแนกตามจำนวนพารามิเตอร์ที่ประกอบอยู่ในโมเดลที่มี 3 พารามิเตอร์ ได้แก่ ความยากของข้อสอบ (item difficulty: b) อำนาจจำแนกของข้อสอบ (item discrimination: a) และโอกาสในการเดา (pseudo-guessing parameter: c) ซึ่งแบ่งออกเป็นโมเดลโลจิสติกแบบสามพารามิเตอร์ (three-parameter logistic model: 3PL) โมเดลโลจิสติกแบบสองพารามิเตอร์ (two-parameter logistic model: 2PL) และโมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ (one-parameter logistic model: 1PL) ซึ่งในที่นี้จะเสนอเฉพาะโมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบสองค่า เนื่องจากเป็นโมเดลที่ใช้ในงานวิจัยนี้และเป็นฟังก์ชันที่คำนวณง่าย และสะดวกทนทานต่อความคลาดเคลื่อนและเป็นโมเดลที่นิยมใช้ (Lord, 1986)

(3) โมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบสองค่า (models for dichotomous item) โมเดลสำหรับข้อสอบที่มีการให้คะแนนแบบสองค่าแบ่งออกเป็น

1.1 โมเดลโลจิสติกแบบสามพารามิเตอร์ (three-parameter logistic model : 3PL) เป็นคุณลักษณะที่มี 3 พารามิเตอร์ คือ a, b และ c พัฒนามาจาก 2 พารามิเตอร์ เพื่อให้เหมาะสมกับแบบทดสอบที่มีอิทธิพลจากการเดาเข้ามาแฝงอยู่ด้วย (Birnbaum, 1968) ในโมเดล 3PL ความ

น่าจะเป็นในการตอบถูก (x_i=1) สำหรับคนที่ j ในข้อที่ i เมื่อให้ ระดับความสามารถเป็น θ_j (ค่า θ_j มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ผลการวิเคราะห์มักปรับให้มีค่าระหว่าง -3 ถึง +3) แสดงได้ดังนี้

$$P_{ij}\left(x=1|\theta_j\right)=c_i+\left(1-c_i\right)\left[\frac{\exp\left(a_i\left(\theta_j-b_i\right)\right)}{1+\exp\left(a_i\left(\theta_j-b_i\right)\right)}\right] \quad \text{.....(2.1)}$$

เมื่อ c_i แทนค่าพารามิเตอร์การเดา ข้อสอบข้อที่ i ได้ถูกต้อง ซึ่งเป็นความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมาก ๆ จะทำข้อสอบข้อที่ i ได้ถูกต้อง โดย c_i เป็นค่ากำกับต่ำสุด (lower asymptotes ของ ICC) และค่า c_i มีพิสัยอยู่ระหว่าง 0 ถึง 1

b_i แทนค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุดที่ θ ที่มีโอกาสตอบข้อสอบถูก $\frac{1+c_i}{2}$ โดยค่า b มีพิสัยอยู่ระหว่าง $-\infty$ ถึง ∞ แต่ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า b_i อยู่ระหว่าง -2.5 ถึง 2.5 ค่า b_i ที่มีค่าใกล้ -2.5 แสดงว่าเป็นข้อสอบที่ง่าย และ ค่า b_i ที่มีค่าใกล้ 2.5 แสดงว่าเป็นข้อสอบที่ยาก

a_i แทนค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าความชันของ ICC ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่า a_i ระหว่าง 0.5 ถึง 2.5 โดย a_i ที่สูงแสดงว่าข้อสอบมีอำนาจจำแนกที่สูงสามารถจำแนกผู้สอบที่มีความสามารถแตกต่างกันได้ดี

โดยมีการแปลความหมาย a_i ดังนี้ (Baker, 2001)

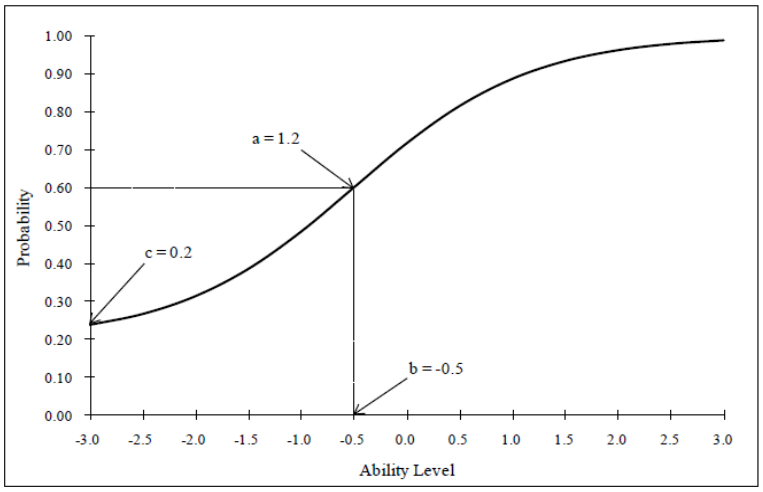
0.01 – 0.34 จำแนกได้ต่ำมาก

0.35 – 0.64 จำแนกได้ต่ำ

0.65 – 1.34 จำแนกได้ปานกลาง

1.35 – 1.69 จำแนกได้สูง

ตั้งแต่ 1.7 ขึ้นไป จำแนกได้สูงมาก



ภาพที่ 2.1 โค้งคุณลักษณะข้อสอบสำหรับโมเดล 3 พารามิเตอร์ (ศิริชัย กาญจนวาสี, 2555)

ลักษณะเฉพาะของโค้งของข้อสอบเมื่อระดับความสามารถของผู้สอบ (θ_j) เปลี่ยนแปลงไป จะมีรูปแบบแตกต่างกัน รูปแบบต่าง ๆ เหล่านี้เปลี่ยนแปลงไปตามฟังก์ชันการตอบสนองรายข้อ หรือที่นิยมเรียกกันทั่วไปว่าโค้งลักษณะเฉพาะของข้อสอบ ซึ่งรูปแบบนี้มีลักษณะเฉพาะดังนี้

(1) เส้นกำกับบนจะมีค่าสูงที่สุดได้เกือบเท่ากับ 1 แต่จะไม่เท่ากับ 1 เพราะโอกาสในการตอบข้อสอบถูกมีได้ไม่เท่ากับ 1 หรือ 100% เพราะไม่มีอะไรแน่นอนที่สุด

(2) เมื่อผู้สอบมีความสามารถมากขึ้น โอกาสในการตอบข้อสอบได้ถูกต้องมีมากขึ้น เส้นโค้งลักษณะเฉพาะของข้อสอบจะสูงขึ้นเรื่อย ๆ เมื่อเคลื่อนจากซ้ายไปขวา และเส้นโค้งมีความชันที่สุดคือจุดที่ดีที่สุดในการแบ่งผู้สอบที่มีความสามารถต่างกัน (จุดนี้เรียกว่าจุด a ซึ่งเป็นค่าอำนาจจำแนก)

(3) เส้นกำกับล่าง จะมีค่าได้น้อยที่สุดเกือบเท่ากับ 0 แต่ไม่เท่ากับ 0 เพราะการทดสอบย่อมมีการเดาไม่มากก็น้อย จะไม่มีการเดาเลยย่อมเป็นไปไม่ได้

1.2 โมเดลโลจิสติกแบบสองพารามิเตอร์ (two-parameter logistic model : 2PL) เป็นคุณลักษณะที่มี 2 พารามิเตอร์ คือ a และ b (Birnbaum, 1968) ในโมเดล 2PL ความน่าจะเป็นในการตอบถูก ($x_i = 1$) สำหรับคนที่ j ในข้อที่ i เมื่อให้ ระดับความสามารถเป็น θ_j แสดงได้ดังนี้

$$P_{ij}\left(x=1|\theta_j\right)=\left[\frac{\exp\left(a_i\left(\theta_j-b_i\right)\right)}{1+\exp\left(a_i\left(\theta_j-b_i\right)\right)}\right] \qquad \text{.....(2.2)}$$

เมื่อ b_i แทนค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุดที่ θ ที่มีโอกาสตอบข้อสอบถูก 0.5

a_i แทนค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าความชันของ ICC ณ ตำแหน่ง b_i

1.3 โมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ (one-parameter logistic model : 1PL) หรือที่เรียกกันว่า Rasch model (Rasch, 1960) เป็นโมเดลที่ประหยัดที่สุดของ IRT ใน 1PL การประมาณค่าของระดับความสามารถของผู้สอบได้จากการตอบสนองข้อสอบ ซึ่งวัดจากหนึ่งพารามิเตอร์ คือ พารามิเตอร์ความยาก (Rasch, 1960; Wright & Panchapakesan, 1969) ในโมเดล 1PL ความน่าจะเป็นในการตอบถูก (x_i=1) สำหรับคนที่ j ในข้อที่ i เมื่อให้ ระดับความสามารถเป็น θ_j แสดงได้ดังนี้

$$P_{ij}\left(x=1\middle|\theta_j\right)=\left[\frac{\exp\left(\theta_j-b_i\right)}{1+\exp\left(\theta_j-b_i\right)}\right] \qquad \text{.....(2.3)}$$

เมื่อ b_i แทนค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุดที่ θ ที่มีโอกาสตอบข้อสอบถูก 0.5

ฟังก์ชันสารสนเทศ (Information Function)

สารสนเทศของข้อสอบเป็นดัชนีผสมเพื่อบ่งชี้คุณภาพของข้อสอบ ประกอบด้วยค่าพารามิเตอร์ a, b, c ซึ่ง Birnbaum (1968) กล่าวถึงแนวคิด สำหรับโมเดลการตอบสนองข้อสอบที่มีการให้คะแนนแบบสองค่าดังนี้

$$I_i\left(\theta_j\right)=\frac{\left[P_i'\left(\theta_j\right)\right]^2}{P_i\left(\theta_j\right)\left(1-P_i\left(\theta_j\right)\right)} \qquad \text{.....(2.4)}$$

เมื่อ $I_i\left(\theta_j\right)$ แทนสารสนเทศของ Fisher (Fisher information) สำหรับคนที่ j ในข้อที่ i ให้ระดับความสามารถแทนด้วย θ_j และ $P_i\left(\theta_j\right)$ เป็นความน่าจะเป็นของการตอบข้อสอบถูกของคนที่ j ในข้อที่ i เมื่อให้ ความสามารถแทนด้วย θ_j และ $1-P_i\left(\theta_j\right)$ เป็นความน่าจะเป็นของการตอบข้อสอบผิดของคนที่ j ในข้อที่ i เมื่อให้ความสามารถแทนด้วย θ_j ฟังชันสารสนเทศของแบบสอบถูกกำหนดเป็นผลรวมของ ค่าสารสนเทศของข้อสอบ ค่าที่มากขึ้นของสารสนเทศของแบบสอบเป็นการวัดที่มีความแม่นยำและดีขึ้น สารสนเทศของแบบสอบแสดงดังนี้

$$I\left(\theta\right)=\sum_{i=1}^{n}I_i\left(\theta\right) \qquad \text{.....(2.5)}$$

Birnbaum (1968) แสดงถึงปริมาณของสารสนเทศของแบบสอบเชื่อมโยงกับความคลาดเคลื่อนมาตรฐานของ θ ดังนี้

$$SE\left(\theta\right)=\frac{1}{\sqrt{I\left(\theta\right)}} \qquad \text{.....(2.6)}$$

1.1.2 โมเดลการวิเคราะห์กลุ่มแฝง (Latent class Model) หรือ โมเดลโครงสร้างแฝง

การสำรวจความแตกต่างของประชากรในด้านพฤติกรรมศาสตร์ ในบริบทด้านการวิจัย กลุ่มตัวอย่างอาจประกอบไปด้วยกลุ่มตัวอย่างที่กำหนดไว้อย่างชัดเจนแล้ว เช่น กลุ่มทดลอง และกลุ่มควบคุม ซึ่งจะมีจุดมุ่งหมายคือการเปรียบเทียบกลุ่มนี้ ในทางความเป็นจริงแล้วแหล่งของความแตกต่างของประชากรอาจไม่สามารถสังเกตได้ เช่น คะแนนจากการทดสอบทางพุทธิปัญญา อาจสะท้อนถึงนักเรียน 2 แบบ คือ นักเรียนเก่งสามารถแก้ปัญหาได้ดี ในขณะที่นักเรียนอ่อนไม่สามารถแก้ปัญหาได้ การแบ่งกลุ่มประชากรที่เหมาะสมควรแบ่งกลุ่มจากสิ่งที่สังเกตได้ เช่น ตัวแปรภูมิหลัง หรือแบ่งกลุ่มจากคุณลักษณะแฝง เช่นคุณลักษณะที่แสดงลักษณะของนักเรียนเก่ง และนักเรียนอ่อนเป็นสิ่งที่น่าสนใจศึกษาต่อไป ซึ่งจะเป็นประโยชน์ในการออกแบบการเรียนการสอน ทางด้านจิตวิทยาความแตกต่างของประชากรที่ไม่สามารถสังเกตได้ อาจประกอบด้วย ความแตกต่างเชิงคุณภาพ ในการระบุประชากรกลุ่มย่อยเพื่อที่จะพัฒนาการระบุกลุ่ม หรือหาลักษณะที่เฉพาะกลุ่ม (Meehl, 1992) เสนอข้อถกเถียงอย่างครอบคลุมเกี่ยวกับความแตกต่างของประชากร ดังนี้

กลุ่มประชากรย่อยเป็นกลุ่มความแตกต่างของประชากร ซึ่งแบ่งเป็นความแตกต่างของประชากรที่สามารถสังเกตได้และสังเกตไม่ได้ สำหรับความแตกต่างที่สังเกตได้จะระบุประชากรย่อยจากตัวแปรสังเกตได้ เช่น เพศ ซึ่งแบ่งประชากรย่อยเป็น เพศชายและเพศหญิง ในบริบทของความแตกต่างที่สังเกตได้นี้ ประชากรย่อยถูกเรียกว่ากลุ่ม และสมาชิกในกลุ่มคือผู้เข้าร่วม และสามารถวิเคราะห์ข้อมูลโดยใช้โมเดลการวิเคราะห์กลุ่มพหุ (multiple group) ซึ่งจะเหมาะสมในกรณีที่สนใจการเปรียบเทียบกลุ่มที่ระบุได้อย่างชัดเจน เช่น กลุ่มเพศ กลุ่มอายุ หรือกลุ่มทดลอง กลุ่มควบคุม และสำหรับความแตกต่างที่ไม่สามารถสังเกตได้จะแตกต่างจากสถานการณ์กลุ่มพหุ โดยตัวแปรสาเหตุของความแตกต่างในข้อมูลเป็นตัวแปรที่ไม่สามารถสังเกตได้ ดังนั้น อาจไม่ทราบได้ว่าประชากรกลุ่มย่อยมีผู้ใดอยู่ในกลุ่มบ้าง และแทบจะเป็นไปไม่ได้ที่จะแบ่งกลุ่มตัวอย่างเข้ากลุ่ม แต่ในความเป็นจริงแล้วสามารถแบ่งกลุ่มประชากรย่อยได้จากข้อมูล ในบริบทความแตกต่างที่ไม่สามารถสังเกตได้ ประชากรย่อยถูกเรียกว่าเป็นกลุ่มแฝง เพราะว่าสมาชิกในประชากรย่อย ซึ่งไม่สามารถสังเกตได้แต่เป็นสิ่งที่แฝงอยู่ ในที่นี้จะใช้คำว่ากลุ่มแฝง หรือ กลุ่มความสามารถแฝง ในการระบุความแตกต่างที่ไม่สามารถสังเกตได้ และใช้คำว่ากลุ่มในบริบทความแตกต่างที่สังเกตได้ (เช่น เพศ จะระบุเป็น 2 กลุ่ม)

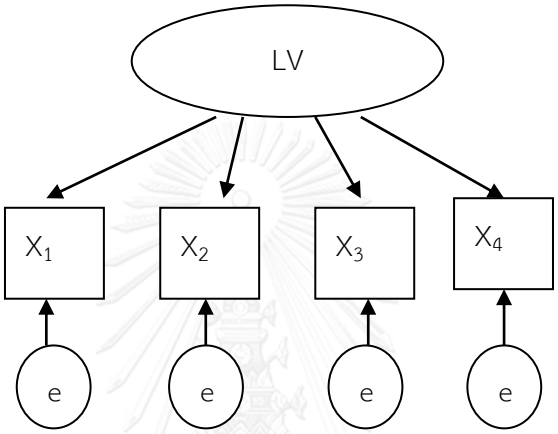
จากวิธีการที่แตกต่างกันนี้ จึงมีการพัฒนาการวิเคราะห์ความแตกต่างที่ไม่สามารถสังเกตได้โดยวิธีการวัดเดิม จะใช้คำถามเกี่ยวกับความแตกต่างของผู้เข้าร่วมกลุ่มในระดับที่เป็นไปตามพฤติกรรมที่สังเกตได้ หรือชนิดของคุณลักษณะที่แตกต่างกัน 2 แบบ (Meehl, 1992; Waller & Meehl, 1998) การวิเคราะห์กลุ่มแฝงแบบดั้งเดิม (Classic latent class) จะสามารถวิเคราะห์จำแนกได้มากกว่า 2 กลุ่ม แต่ยังมีข้อจำกัดกับการระบุตัวแปรที่สังเกตได้ภายในกลุ่ม นักการศึกษาจึง

ได้มีการพัฒนาโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม ซึ่งจะมีความยืดหยุ่นกับจำนวนกลุ่มแฝงและการระบุตัวแปรสังเกตได้ภายในกลุ่ม สำหรับวิธีการจำแนกกลุ่มความแตกต่างของประชากร เมื่อจำแนกตาม ตัวแปรที่ระบุความแตกต่างของประชากร และตัวแปรผลลัพธ์ แสดงดังตารางต่อไปนี้

ตารางที่ 2.1 การจำแนกวิธีการใช้ในการจำแนกความแตกต่างของประชากร (Lubke & Muthén, 2005)

ลักษณะโมเดล	**วิธีการ**	**ตัวแปรที่ระบุความแตกต่างของประชากร**	**ตัวแปรผลลัพธ์**
ไม่เป็นโมเดลตัวแปรแฝง	การวิเคราะห์จำแนก (Discriminant analysis)	สังเกตได้	ตัวแปรต่อเนื่อง
	การวิเคราะห์การถดถอยโลจิสติก (Logistic regression)	สังเกตได้	ตัวแปรจัดประเภท
	การวิเคราะห์ความแปรปรวนพหุ (MANOVA)	สังเกตได้	ตัวแปรต่อเนื่อง
	การจัดกลุ่มแบบ K-Means (K - Means clustering)	ไม่สามารถสังเกตได้	ตัวแปรต่อเนื่อง
เป็นโมเดลตัวแปรแฝง	การวิเคราะห์กลุ่มพหุ (Multigroup CFA)	สังเกตได้	ตัวแปรต่อเนื่อง/ตัวแปรจัดประเภท
	การวิเคราะห์กลุ่มแฝง (Latent class analysis)	ไม่สามารถสังเกตได้	ตัวแปรจัดประเภท
	การวิเคราะห์โปรไฟล์ (Latent Profile analysis)	ไม่สามารถสังเกตได้	ตัวแปรต่อเนื่อง
	การวิเคราะห์กลุ่มแฝงแบบผสม/ การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบผสม (Factor Mixture Modeling / Mixture IRT)	ไม่สามารถสังเกตได้	ตัวแปรต่อเนื่อง/ตัวแปรจัดประเภท

สำหรับการวิเคราะห์กลุ่มแฝง (latent class analysis) หรือการวิเคราะห์กลุ่มที่เป็นตัวแปรแฝงเป็นหนึ่งในเครื่องมือมาตรฐานที่ใช้ในการวิเคราะห์ข้อมูลทางสังคมศาสตร์ (Magidson & Vermunt, 2001) ซึ่งตัวแปรแฝง เป็นตัวแปรที่ไม่สามารถวัดได้โดยตรงด้วยค่าเฉลี่ยของตัวแปรสังเกตได้ ดังนั้น จะมีการวัดทางอ้อมโดยใช้ตัวแปรสังเกตได้ 2 หรือ 3 ตัวแปร ในการวิเคราะห์ทางสถิติเกือบทั้งหมดมีพื้นฐานว่าโมเดลที่มีตัวแปรแฝงจะมีการจำแนกตัวแปรแฝงและความคลาดเคลื่อนจากการวัดออกจากกัน และในการวิเคราะห์กลุ่มแฝง คำว่าตัวแปรสังเกตได้จะหมายถึงหน่วยในความคลาดเคลื่อน



ภาพที่ 2.2 ตัวแปรแฝง ที่มีตัวแปรสังเกตได้ 4 ตัว เป็นตัวชี้วัด (Collins & Lanza, 2010)

LV แทน ตัวแปรแฝง

X_1,... X_4 แทนตัวแปรบ่งชี้ที่สังเกตได้

e_1, ... , e_4 แทนความคลาดเคลื่อนที่สัมพันธ์กับ X_1,... X_4

จากภาพลูกศรจากตัวแปรแฝงไปยังตัวแปรบ่งชี้/ตัวแปรสังเกตได้ และลูกศรจากความคลาดเคลื่อนจะชี้ไปยังตัวแปรสังเกตได้ ซึ่งลูกศรจะแทนความสำคัญ และหลักการของการวิเคราะห์ตัวแปรแฝง สังเกตได้ว่า ความสัมพันธ์เชิงสาเหตุจะเริ่มจาก ตัวแปรแฝงไปยังตัวบ่งชี้ ซึ่งแตกต่างจากปกติที่ตัวแปรสังเกตได้ใช้วัดตัวแปรแฝง แต่ตัวแปรสังเกตได้แต่ละตัวแปรไม่ใช่สาเหตุของตัวแปรแฝง

โมเดลกลุ่มแฝง (latent class model หรือ LC model) นี้นำเสนอเริ่มแรกโดย Lazarsfeld และ Henry (1968) ซึ่งเป็นโมเดลสำหรับตัวแปรแฝงประเภททวิวิภาค (dichotomous latent variable) แต่ไม่มีการศึกษาเกี่ยวกับแนวคิดของการวิเคราะห์ตัวแปรแฝงแบบจัดประเภท ต่อมา Goodman (1974) ได้ประยุกต์ใช้วิธีการดังกล่าวกับตัวแปรแฝงระดับนามบัญญัติ (nominal latent

variable) และได้พัฒนาวิธีการวิเคราะห์กลุ่มแฝงโดยนำวิธีการประมาณค่าแบบ maximum likelihood มาใช้ในการวิเคราะห์ วิธีการของ Goodman ที่มีการประมาณค่านี้ มีความสัมพันธ์ใกล้เคียงกับหลักการ expectation-maximization (EM) (Dempster, Laird, & Rubin, 1977) ซึ่งหลักการนี้ถูกใช้มากในการวิเคราะห์กลุ่มแฝงจนถึงปัจจุบัน นอกจากนี้ Clogg (1988) Uebersax (1993) และ Heinen (1996) อ้างถึงใน Vermunt และ Margidson, 2001 นำวิธีการดังกล่าวไปใช้กับกลุ่มตัวแปรแฝงที่มีลักษณะเป็นตัวแปรแฝงเรียงอันดับ (ordinal latent variable) และ Wolfe (1970) McLachlan และ Basford (1988) Fraley และ Raftery (1998) อ้างถึงใน Vermunt และ Margidson (2001) ได้นำวิธีการดังกล่าวไปใช้กับกลุ่มตัวแปรแฝงที่มีลักษณะเป็นตัวแปรจัดประเภทแต่มีระดับการวัดแบบตัวแปรต่อเนื่อง นอกจากนี้การพัฒนาเทคนิคการวิเคราะห์กลุ่มแฝง สามารถใช้วิเคราะห์ตัวแปรสังเกตได้ที่มีระดับการวัดแตกต่างกันมาวิเคราะห์ร่วมกันได้ด้วย (Lawrence and Krzanowski, 1996, Moustaki, 1996, Hunt and Jorgensen, 1999 อ้างถึงใน Vermunt and Margidson, 2001)

กล่าวได้ว่าการวิเคราะห์กลุ่มแฝง เป็นเทคนิคการวิเคราะห์ตามลักษณะการวิเคราะห์องค์ประกอบประเภทแรกที่ได้รับการพัฒนาให้ สามารถวิเคราะห์ตัวแปรทวิวิภาค (Dichotomous variable) ได้ จากนั้นจึงมีการพัฒนาให้สามารถวิเคราะห์ตัวแปรจัดประเภทได้ทุกชนิด รวมทั้งสามารถวิเคราะห์ตัวแปรที่มีระดับการวัดต่างกันได้

โมเดลกลุ่มแฝง (Latent Class model) จะอธิบายความสัมพันธ์ระหว่างการตอบสนองของค่าสังเกตได้และข้อสอบ โมเดลกลุ่มแฝงประกอบไปด้วย ตัวแปรที่มีลักษณะเป็นกลุ่มซึ่งแทนกลุ่มที่ถูกระบุ และจำนวนของตัวแปรอื่น ๆ ซึ่งแทนคุณลักษณะของสิ่งนั้น (Zhang, 2004) ในการวิเคราะห์กลุ่มแฝง (Latent class analysis) สามารถพิจารณาข้อมูลเชิงคุณภาพ เหมือนกับการวิเคราะห์องค์ประกอบ (factor analysis) ซึ่งทำให้นักวิจัยสามารถระบุตัวแปรแฝงไม่ต่อเนื่อง (discrete latent variables) โดยการสังเกตจากตัวแปรสังเกตได้ไม่ต่อเนื่อง (discrete observed variables) (Green, 1951, 1952 อ้างถึงใน Mccutcheon, 1987) นอกจากนี้ยังมีเทคนิคการวิเคราะห์โครงสร้างแฝง (Latent structure methods) อีก 2 เทคนิค คือ การวิเคราะห์คุณลักษณะแฝง (Latent trait analysis) เป็นการวิเคราะห์ที่ทำให้สามารถบอกคุณลักษณะของตัวแปรแฝงต่อเนื่อง (continuous latent variables) จากตัวแปรสังเกตได้ไม่ต่อเนื่อง (discrete observed variables) การวิเคราะห์โปรไฟล์แฝง (latent profile analysis) เป็นการวิเคราะห์ที่ทำให้สามารถบอกลักษณะตัวแปรแฝงไม่ต่อเนื่อง (discrete latent variables) จากตัวแปรสังเกตได้ต่อเนื่อง (continuous observed variables) (Mccutcheon, 1987) จากการวิเคราะห์โครงสร้างทั้ง 4 ประเภทดังกล่าว สามารถแสดงได้ดังตารางที่ 2.2

ตารางที่ 2.2 ความแตกต่างของโมเดลที่มีตัวแปรแฝง (Collin & Lanza, 2011)

ตัวแปรแฝง	**ตัวแปรสังเกตได้**	
	ตัวแปรต่อเนื่อง	**ตัวแปรไม่ต่อเนื่อง / ตัวแปรจัดประเภท**
ตัวแปรต่อเนื่อง	การวิเคราะห์องค์ประกอบ	การวิเคราะห์โปรไฟล์แฝง
ตัวแปรไม่ต่อเนื่อง/ ตัวแปรจัดประเภท	การวิเคราะห์คุณลักษณะแฝง หรือ ทฤษฎีการตอบสนองข้อสอบ	การวิเคราะห์กลุ่มแฝง

ในการวิเคราะห์กลุ่มแฝงนั้นมีจุดประสงค์ที่สำคัญ 2 ประการ คือ ประการแรกเพื่อจำแนกความแตกต่างของแต่ละตัวแปร (variable-oriented) ต้องการอธิบายความสัมพันธ์ระหว่างตัวแปรโดยอธิบายความสัมพันธ์ของตัวแปรข้ามบุคคลทั้งหมด เช่น การวิเคราะห์องค์ประกอบ (factor analysis) เนื่องจากจุดเน้นของการวิเคราะห์องค์ประกอบนั้น คือเป็นการระบุความแตกต่างของโครงสร้างที่สร้างองค์ประกอบขึ้น สำหรับความสัมพันธ์เชิงเส้นตรงระหว่างกลุ่มของตัวแปรสังเกตได้ และจุดประสงค์ประการที่สองเพื่อจำแนกความแตกต่างของแต่ละบุคคล (person-oriented) เพื่อดูว่าลักษณะของแต่ละบุคคลที่เรากำลังศึกษาเป็นอย่างไร และจัดกลุ่มคนออกเป็นกลุ่ม และดูว่าแต่ละกลุ่มมีลักษณะอย่างไร (Bergman & Magnusson, 1997) ในหัวข้อนี้ผู้วิจัยนำเสนอรายละเอียดของโมเดลการวิเคราะห์กลุ่มแฝงเป็นหัวข้อดังนี้ (1) โมเดลการวิเคราะห์กลุ่มแฝง แบบจำลองทางคณิตศาสตร์ของการวิเคราะห์กลุ่มแฝง (2) หลักการและขั้นตอนของการวิเคราะห์กลุ่มแฝง (3) ข้อตกลงเบื้องต้นของการวิเคราะห์กลุ่มแฝง (4) ประเภทของโมเดลกลุ่มแฝง

(1) โมเดลกลุ่มแฝง หรือแบบจำลองทางคณิตศาสตร์ของการวิเคราะห์กลุ่มแฝง สามารถอธิบายได้จาก Rijmen และ De Boeck (2003) ดังนี้

$$P(Y_i|\phi_i = k) = \prod_{J-1}^{J} P(y_{ij}|\phi_i = k) \qquad \text{.....(2.7)}$$

เมื่อ $P(Y_i|\phi_i = k)(\phi = 1,...,k,...,K)$ เป็นความน่าจะเป็นของเวกเตอร์รูปแบบการตอบสนองของค่าสังเกต Y_i ของแต่บุคคล i ให้กลุ่มตัวแปรแฝง k กลุ่ม และ $P(y_{ij}|\phi_i = k)$ เป็นความน่าจะเป็นของการตอบสนอง y_{ij} สำหรับบุคคล i และข้อสอบ j โดยที่ j=1,...,J ในกลุ่มตัวแปรแฝง k ซึ่งโมเดลนี้มีข้อตกลงให้มีความเป็นอิสระกันข้ามข้อสอบในสมาชิกกลุ่มตัวแปรแฝง

ความน่าจะเป็นของเวคเตอร์ รูปแบบการตอบสนอง Y_i สำหรับแต่ละบุคคล i คือ

$$P(Y_i) = \sum_{\phi_i-1}^{k} \pi_{\phi-k} P(Y_i|\phi_i = k) \qquad \text{.....(2.8)}$$

เมื่อ $\pi_{\phi-k}$ เป็นสัดส่วนของสมาชิกกลุ่มสำหรับอธิบายและระบุกลุ่มแฝง k หรืออาจเรียกว่าเป็นความน่าจะเป็นโดยรวมของการเป็นสมาชิกในกลุ่มแฝง k

(2) หลักการ และขั้นตอนของการวิเคราะห์กลุ่มแฝง

การวิเคราะห์กลุ่มแฝง (latent class analysis) เป็นวิธีการทางสถิติจัดเป็นเทคนิคสถิติวิเคราะห์ตัวแปรพหุนาม (multivariate statistical analysis technique) สำหรับการวิเคราะห์จัดกลุ่มที่เป็นตัวแปรแฝงจากข้อมูลที่เป็นตัวแปรสังเกตได้ชนิดตัวแปรจัดประเภท เทคนิคการวิเคราะห์กลุ่มแฝงคล้ายคลึงกับการวิเคราะห์แบ่งกลุ่ม (cluster analysis) และการวิเคราะห์องค์ประกอบ (factor analysis) ซึ่งในการวิเคราะห์องค์ประกอบ เมตริกซ์ความแปรปรวนร่วมจะถูกวิเคราะห์ภายใต้โครงสร้างพื้นฐานที่แฝงอยู่ ในโมเดลการวิเคราะห์องค์ประกอบ ตัวแปรแฝงเป็นตัวแปรต่อเนื่อง การวิเคราะห์กลุ่มแฝงมีความแตกต่างกับการวิเคราะห์องค์ประกอบในเรื่องขั้นตอนของการดำเนินการ และตัวแปรแฝงจะเป็นตัวแปรไม่ต่อเนื่อง และแบ่งประชากรออกเป็นกลุ่มชัดเจน เช่น ผลสัมฤทธิ์ทางการเรียนที่แฝงอยู่ในการวิเคราะห์กลุ่มแฝงจะเป็นการแบ่งประชากรออกเป็นกลุ่ม เก่งมาก ปานกลาง และอ่อน ในการวิเคราะห์กลุ่มแฝงตัวแปรแฝงจะเป็นตัวแปรแบบคงที่ ดังนั้นจะมีการวัดครั้งเดียว และมีการระบุจำนวนกลุ่มแฝง นอกจากนี้ในการวิเคราะห์กลุ่มแฝง โครงสร้างของตัวแปรแฝงจะถูกสร้างมาจากตัวบ่งชี้ (indicator) ที่เป็นตัวแปรสังเกตได้ในโมเดลสมการโครงสร้าง (structural equation model) ประเภทโมเดลการวัด (measurement model) และใช้การแบ่งโครงสร้าง (construct or cluster) ตามหลักการวิเคราะห์องค์ประกอบ (factors analysis) โดยการวิเคราะห์กลุ่มแฝงมีขั้นตอนดังต่อไปนี้

การวิเคราะห์กลุ่มแฝงเป็นเทคนิคการวิเคราะห์โมเดลสมการโครงสร้างชนิดหนึ่ง ดังนั้นการวิจัยโดยใช้การวิเคราะห์กลุ่มแฝงมีขั้นตอนคล้ายคลึงกับการวิจัยโดยใช้การวิเคราะห์โมเดลสมการโครงสร้าง จากการศึกษาเอกสารของ B. Muthén (2004) สรุปเป็นขั้นตอนการวิเคราะห์กลุ่มแฝง 6 ขั้นตอน ดังต่อไปนี้

ขั้นตอนที่ 1 เป็นการพัฒนารูปแบบโมเดล โดยใช้ทฤษฎีเพื่อให้ได้โมเดลการวิเคราะห์กลุ่มแฝง

ขั้นตอนที่ 2 เป็นการกำหนดรายละเอียด และดำเนินการตรวจสอบข้อมูลรวมทั้งการตรวจสอบรูปแบบการตอบสนองตัวบ่งชี้รายตัว

ขั้นตอนที่ 3 เป็นการวิเคราะห์ประมาณค่าพารามิเตอร์ในโมเดล

3.1 การวิเคราะห์เพื่อทดสอบความกลมกลืนของการวิเคราะห์กลุ่มแฝงกับข้อมูลเชิงประจักษ์เมื่อกำหนดให้มีจำนวนกลุ่ม 2 กลุ่ม

3.2 การประมาณค่าความน่าจะเป็นของตัวแปรสังเกตได้แต่ละตัวของหน่วยตัวอย่างว่าควรจะอยู่ในกลุ่มใดใน 2 กลุ่ม

3.3 การตรวจสอบรูปแบบของตัวแปรหรือพฤติกรรมของหน่วยตัวอย่างในแต่ละกลุ่มโดยดูความสอดคล้องของความน่าจะเป็นในแต่ละรูปแบบ

3.4 การหาค่าความน่าจะเป็นเฉลี่ยของกลุ่ม (mean probability of class memberships)

3.5 การประเมินความแตกต่างระหว่างกลุ่ม เพื่อดูว่าแต่ละกลุ่มมีความแตกต่างกันอย่างไร (evaluating group differences) และมีจำนวนกลุ่มตัวอย่างเท่าไร

ขั้นตอนที่ 4 เป็นการดำเนินการตามขั้นตอนที่ 3.1-3.5 สำหรับการวิเคราะห์โมเดลกลุ่มแฝงเมื่อกำหนดให้มีจำนวนกลุ่มแฝงเป็น 3, 4,...,k กลุ่ม คัดเลือกเฉพาะโมเดลที่ข้อมูลเชิงประจักษ์สอดคล้องกับข้อมูลเชิงทฤษฎี (χ^2 / df มีค่าน้อยกว่า .05)

ขั้นตอนที่ 5 เปรียบเทียบผลการวิเคราะห์ที่ได้จากขั้นตอนที่ 4 ว่าโมเดลที่มีความเหมาะสมที่สุด มีจำนวนกลุ่มเป็นเท่าใด โดยพิจารณาจากค่าสถิติที่ใช้ในการทดสอบความสอดคล้องของโมเดล (model goodness of fit test) ได้แก่ค่า AIC, BIC, และค่า sample-size adjusted BIC กล่าวคือค่าสถิติที่ใช้ในการทดสอบความสอดคล้องของโมเดลมีค่าน้อยที่สุด หรือมีค่าเข้าใกล้ศูนย์ แสดงว่าเป็นโมเดลที่มีความเหมาะสมที่สุด

ขั้นตอนที่ 6 นำเสนอผลการวิเคราะห์ และแปลความหมาย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(3) ข้อตกลงเบื้องต้นในการวิเคราะห์กลุ่มแฝง

1. ประชากรประกอบด้วยกลุ่มประชากรย่อยที่มีคุณสมบัติเหมือนกันและมีความสัมพันธ์กันกับกลุ่มประชากรเหล่านี้จะจัดรูปแบบการจัดกลุ่มแฝงตามคำนิยามตัวแปรแฝง (latent variable) แทนด้วยสัญลักษณ์ θ จำนวนของกลุ่มแฝง (latent class) แทนด้วย t และความน่าจะเป็นของการตอบข้อสอบ แทนด้วย P_{θ_t} ซึ่ง $\sum tP_{\theta_t} = 1$ กลุ่มประชากรย่อยประกอบด้วยส่วนที่เหมือนกัน ในความหมายของความน่าจะเป็น สำหรับการให้คำตอบที่เหมือนกันของสิ่งใดสิ่งหนึ่งนั้นจะเป็นสมาชิกในกลุ่มแฝงเดียวกัน

2. ข้อตกลงเกี่ยวกับความเป็นอิสระเฉพาะที่ (local independence) ซึ่งกล่าวถึงความสัมพันธ์ระหว่างตัวแปรสังเกตได้ (Manifest variable) สามารถอธิบายได้โดยความไม่เป็นอิสระต่อกันของตัวแปรสังเกตได้ของการจัดกลุ่มแฝง (latent classification) ดังนั้นการจัดกลุ่มประชากร

ย่อยแฝง (latent subpopulation) ตัวบ่งชี้สังเกตได้ทั้งหมดจะเป็นอิสระต่อกันโดยวิธีการทางสถิติ หรืออีกนัยหนึ่ง จากการสร้างตารางการณ์จร (separate contingency table) ในการตรวจสอบว่าตัวแปรสังเกตได้ตัวแปรหนึ่งมีผลกับอีกตัวแปรหนึ่งหรือไม่ สำหรับตัวแปรที่สังเกตได้ที่สอดคล้องกับกลุ่มแฝงแต่ละกลุ่ม ตัวแปรที่สังเกตได้นั้นจะเป็นอิสระต่อกันในแต่ละตาราง แต่เนื่องจากตัวแปรแฝงไม่สามารถสังเกตได้โดยตรง ตารางการณ์จรจะถูกประมาณขึ้นมา (Heinen, 1996)

การวิเคราะห์กลุ่มแฝงเป็นพาราเมตริกโมเดลอย่างง่ายและใช้ข้อมูลจากการสังเกตเพื่อการประมาณค่าพารามิเตอร์ สำหรับโมเดลพารามิเตอร์ควรมีลักษณะดังนี้

(1) กลุ่มแฝง แต่ละกลุ่มต้องมีความสัมพันธ์กันในแต่ละกลุ่ม

(2) ความน่าจะเป็นของการส่งผลอย่างมีเงื่อนไข นั่นคือ ความน่าจะเป็นของการรวมกลุ่มแฝงแต่ละกลุ่มและแต่ละระดับการส่งผลของตัวแปร ต้องมีการเลือกสมาชิกแต่ละกลุ่มจากการสุ่ม (Formann, 1992)

(4) ประเภทของโมเดลกลุ่มแฝง (Latent Class Models)

Vermunt และ Magidson (2002) กล่าวถึงการจำแนกประเภทของโมเดลกลุ่มแฝงตามหลักการประยุกต์ทางสถิติทั่วไปของการวิเคราะห์กลุ่มแฝง (Latent Class analysis) จากลักษณะต่อไปนี้ (1) การจัดกลุ่มของกรณีต่าง ๆ (clustering of cases) (2) การลดลงของตัวแปร และการวัดโครงสร้าง (variable reduction and scale construction) (3) การทำนายตัวแปรตาม (prediction of a dependent variable) โดยจำแนกประเภทได้ดังนี้

(1) LC Cluster Models ระบุกลุ่มกรณีที่มีลักษณะ ความสนใจ คุณค่า พฤติกรรมคล้ายกันในกลุ่มเดียวกันโดยรวมตัวแปรแฝง k กลุ่ม และแต่ละตัวแปรจำแนก แทนที่เป็นกลุ่ม

(2) LC Factor Models ระบุปัจจัยซึ่งรวมตัวแปรที่มีแหล่งความแปรปรวนคล้ายกัน สามารถรวมปัจจัยแฝงที่เป็นลำดับที่เป็น 2 ลำดับหรือมากกว่า LC Factor Models คล้ายกับ maximum likelihood (ML) factor analysis ซึ่งอาจใช้ในการสำรวจ หรือยืนยันปัจจัย โดยอาจสมมติว่ามีความสัมพันธ์กันหรือไม่สัมพันธ์กันก็ได้

(3) LC Regression and Choice Models ใช้ในทำนายตัวแปรตามในลักษณะเดียวกับหน้าที่ของตัวทำนาย รวมตัวแปรแฝงที่เป็นตัวแปรจำแนก r กลุ่ม ในแต่ละกลุ่มจำแนกแทนที่ด้วยประชากรที่มีลักษณะเดียวกัน (class, segment) ความถดถอยที่แตกต่างถูกประมาณสำหรับแต่ละประชากร (for each latent segment) แยกกรณีเป็นส่วน ๆ และพัฒนาโมเดลการถดถอยไปพร้อม ๆ กัน

การใช้ LC models เป็นทางเลือกหนึ่ง ที่จะใช้เกี่ยวกับแนวคิดเครือข่ายประสาท (neural network) ของการจำลองที่ไม่เป็นเชิงเส้นตรง ตัวอย่างโปรแกรมคอมพิวเตอร์สำเร็จรูปที่ใช้ในการวิเคราะห์กลุ่มแฝง เช่น Latent GOLD® computer program (Vermunt & Magidson, 2005) Proc LCA (Lanza, Collins, Lemmon, & Schafer, 2007) MPLUS (L. K. Muthén & Muthén, 2008)

โดยสรุป การวิเคราะห์กลุ่มแฝงมีความสำคัญคือ ในโมเดลกลุ่มแฝง การตอบสนองที่สังเกตได้ของแต่ละบุคคล ถูกกำหนดโดยการรวมกันของกลุ่มแฝงแต่ละกลุ่ม และความคลาดเคลื่อนอย่างสุ่ม โมเดลกลุ่มแฝงอาจจะมีความแตกต่างทั้งเชิงคุณภาพและเชิงปริมาณระหว่างกลุ่มแฝง การวิเคราะห์กลุ่มแฝงสามารถช่วยให้ผู้วิจัยเข้าใจในความซับซ้อนที่ไม่แน่นอน จุดเด่นของกลุ่มแฝงคือความน่าจะเป็นของสมาชิกในแต่ละกลุ่มแฝง เพราะว่าแต่ละกลุ่มแฝงมีความสัมพันธ์กันเฉพาะ ซึ่งความน่าจะเป็นของการตอบสนองข้อสอบเป็นความน่าจะเป็นของการตอบสนองที่สังเกตได้เฉพาะต่อตัวแปรที่เฉพาะเจาะจงที่มีเงื่อนไขต่อสมาชิกของกลุ่มแฝงแต่ละกลุ่มแฝง ความน่าจะเป็นของการตอบสนองข้อสอบสำหรับกลุ่มของการตอบสนองที่เป็นไปได้ สำหรับข้อสอบเดียวมีผลรวมเป็น 1 ความน่าจะเป็นในการตอบสนองข้อสอบเป็นข้อมูลพื้นฐานสำหรับการระบุกลุ่มแฝง และการวิเคราะห์กลุ่มแฝงมีข้อตกลงเบื้องต้นของความเป็นอิสระทั่วไป ในการกำหนดเงื่อนไขต่อกลุ่มแฝง ตัวแปรที่สังเกตได้ เป็นอิสระต่อกัน และข้อตกลงเบื้องต้นนี้ไม่สามารถบ่งบอกถึงตัวแปรที่สังเกตได้ว่าเป็นอิสระในกลุ่มตัวอย่างที่สมบูรณ์

นอกจากนี้การวิเคราะห์กลุ่มแฝง เป็นการวิเคราะห์ที่ใช้สำหรับการจัดกลุ่ม เมื่อข้อมูลตัวแปรสังเกตได้มีระดับการวัดแบบตัวแปรไม่ต่อเนื่องหรือตัวแปรจัดประเภท (categorical variable) การวิเคราะห์ข้อมูลนี้ มีจุดเด่นตรงที่เป็นเทคนิคการวิเคราะห์ชนิดแรกที่ได้รับการพัฒนาตามหลักการวิเคราะห์องค์ประกอบที่สามารถใช้ได้กับข้อมูลที่เป็นตัวแปรจัดประเภทและความสัมพันธ์ระหว่างตัวแปรสังเกตได้ไม่ใช่ความสัมพันธ์เชิงเส้นตรง (Magidson & Vermunt, 2004) และปัจจุบันการวิเคราะห์กลุ่มแฝง (LCA) ถูกพัฒนาให้สามารถใช้วิเคราะห์จัดกลุ่มข้อมูลที่มีระดับการวัดแตกต่างกัน กล่าวคือ สามารถวิเคราะห์จัดกลุ่มกรณีที่ตัวแปรสังเกตได้มีระดับการวัดแบบผสมทั้งแบบตัวแปรต่อเนื่อง และตัวแปรจัดประเภท นอกจากนี้วิธีการวิเคราะห์กลุ่มแฝง(LCA) ยังสามารถนำมาประยุกต์ใช้ในการศึกษาการเปลี่ยนแปลงเกี่ยวกับประวัติศาสตร์ (event history) (McCutcheon, 1987) Muthen และ Muthen (2004) ได้พัฒนาเทคนิคการวิเคราะห์กลุ่มแฝงให้สามารถวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต่อเนื่องที่เป็นสาเหตุของตัวแปรแฝง ที่ได้จากการวิเคราะห์กลุ่มแฝงอีกด้วย โดยเรียกเป็นโมเดลการตอบสนองข้อสอบแบบผสม (Mixture Item Response Theory) รวมทั้งสามารถวิเคราะห์กลุ่มแฝงกรณีที่เป็นข้อมูลระยะยาว

1.1.3 ทฤษฎีการตอบสนองข้อสอบแบบผสม (Mixture Item Response Theory: MixIRT)

ทฤษฎีการตอบสนองข้อสอบแบบผสม (MixIRT) เป็นวิธีการหนึ่ง ที่สามารถระบุกลุ่มที่สังเกตได้ของผู้ตอบกับความสัมพันธ์การตอบสนองแฝงที่แตกต่างกัน (R. Mislevy & Huang, 2007) ในส่วนนี้นำเสนอดังหัวข้อต่อไปนี้ (1) นิยามทั่วไปของโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) (2) ข้อตกลงเบื้องต้น (3) การประมาณพารามิเตอร์ของข้อสอบในการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบผสม

(1) โมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) เป็นการบูรณาการโมเดลการตอบสนองข้อสอบ (IRT model) และโมเดลกลุ่มแฝง (LC model) ถูกเสนอโดย Cho and Cohen (2007) บางครั้งโมเดลนี้อาจเรียกว่าเป็นโมเดลการวิเคราะห์องค์ประกอบแบบผสม (Muthen & Asparouhov, 2006) ความแตกต่างของทั้งสองโมเดลคือในโมเดลกลุ่มแฝงตัวแปรจะเป็นตัวแปรจัดประเภท ส่วนในโมเดลการตอบสนองข้อสอบตัวแปรจะเป็นแบบต่อเนื่อง ดังนั้นโมเดลการตอบสนองข้อสอบแบบผสมจะมีประโยชน์ที่สำคัญคือ ยอมให้มีการผันแปรภายในกลุ่มของตัวแปรแฝงได้ หรืออาจกล่าวได้ว่า มีความโน้มเอียงสำหรับการตอบสนองในแต่ละข้อของแบบสอบ ที่แม้ว่าผู้สอบในกลุ่มเดียวกันจะมีความผันแปรในความสามารถแฝง แต่อย่างไรก็ตาม ผู้สอบในกลุ่มแฝงนั้นต้องมีความเป็นเอกพันธ์ในการตอบข้อสอบ และไม่มีการผันแปรในความสามารถแฝง นอกจากนี้ในการเพิ่มความยาวของข้อสอบในแบบสอบ เมื่อความยาวของแบบสอบเพิ่มขึ้น เมทริกซ์ของรูปแบบการตอบสนองจะเพิ่มการกระจายตัวขึ้น ซึ่งทำให้ประมาณค่าพารามิเตอร์ได้ยาก โมเดลการตอบสนองข้อสอบสามารถจัดการกับการกระจายตัวนั้นได้ โดยการประยุกต์ใช้ โมเดลพาราเมทริกซ์ (เช่น ฟังก์ชันโอจีฟหรือโลจิสติกส์)

โมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) สามารถตรวจสอบกลุ่มที่สังเกตได้ในพารามิเตอร์ข้อสอบ ซึ่งโมเดลการตอบสนองข้อสอบ (IRT model) แบ่งออกเป็นหลายโมเดล ดังนั้นโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) จึงสามารถแบ่งออกเป็นหลายโมเดลด้วยกัน เช่น Mixture Rasch Model (MRM) เป็นโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) แบบหนึ่งที่เป็นการบูรณาการโมเดลการตอบสนองข้อสอบ (IRT model) แบบ Rasch model กับโมเดลกลุ่มแฝง (LC model) และ Mix PCM model เป็นโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) ที่เป็นการบูรณาการโมดลการตอบสนองข้อสอบ (IRT model) แบบ Rasch สำหรับข้อสอบที่มีการให้คะแนนบางส่วน (Partial credit scoring: PCM) ซึ่ง PCM model เป็น Polytomous Rasch model โดยมีข้อตกลงว่าความยากของข้อสอบแต่ละข้อมีค่าเท่ากัน MRM มีการเพิ่มการแจกแจงผสมเข้าไปโดยออกแบบสำหรับกลุ่มตัวแปรแฝงโดยเฉพาะในการแจกแจงความสามารถซึ่งถูกนำมาใช้โดย Cho (2007) นอกจากนี้โมเดลการตอบสนองข้อสอบแบบผสมถูกนำไปประยุกต์ใช้ในด้านต่าง ๆ เช่น

Rost (1997) ประยุกต์ใช้ Mixture Rasch Model (MRM) ในการตรวจสอบความแตกต่างกันของคุณลักษณะการตอบสนองระหว่างกลุ่มแฝงของผู้สอบ Mixture linear test model ถูกใช้ในการตรวจสอบพฤติกรรมการเดาต่อ การทดสอบแบบเลือกตอบ (R. J. Mislevy & Verhelst, 1990) นอกจากนี้ Bolt, Cohen และ Wollack (2001) ใช้ Mixture nominal response model ในการตรวจสอบความแตกต่างกันของแต่ละบุคคลในการจำแนกกลุ่มจากการตอบสนองในข้อสอบแบบเลือกตอบ โดยจำแนกบุคคลเพื่อวินิจฉัยทางปัญญา (cognitive diagnostic) และ Cohen, Cho และ Kim (2005) ใช้ MRM สำหรับการวิเคราะห์ผลของข้อสอบ (analyzing testlet effects) Wilson (1989) ใช้ MRM ในการจำลองการเรียนรู้ (modeling learning) Von Davier และ Yamamoto (2004) ใช้สำหรับการสังเกตการผสม (observed mixture) Bolt, Cohen และ Wollack, (2002) Yamamoto และ Everson (1997) ใช้ทดสอบความเร็ว (test speededness) Reise และ Gomel (1995) ใช้ MRM ในการระบุความผันแปรเชิงคุณภาพ (identifying qualitative variation) Smit, Kelderman และ Von Der Flier (2000) เสนอ Mixed Birnbaum model ซึ่งเป็นการรวมกันของโมเดลกลุ่มแฝง กับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ เพื่อใช้ในการปรับเทียบข้อมูลในการรวมกลุ่มที่มีคุณลักษณะแฝงคล้ายกัน นอกจากนี้ Bilir (2011) ได้ออกแบบ Mixture IRT-MIMIC (MixIRT-MIMIC) Model เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ทั้งในกลุ่มตัวแปรสังเกตได้และตัวแปรแฝง

ในที่นี้จะขอกล่าวถึง Mixture Rasch Model ที่มีการใช้ Rasch model และ Mixture model ที่มีการใช้โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ เนื่องจากในการพัฒนาโมเดล Multilevel Mixture IRT ที่ผ่านมาได้มีการศึกษาโดยการบูรณาการโมเดลนี้

Chulalongkorn University

1. *Mixture Rasch Model (MRM)* เป็นโมเดลหนึ่งที่นำมาใช้สำหรับตัวแปรแฝงซึ่งไม่สามารถสังเกตได้ เช่น ความสามารถของบุคคล ค่าความยากของข้อสอบ เป็นต้น ซึ่ง Cohen และ Cho (2005) ได้นำโมเดลนี้มาตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) จากกลุ่มตัวแปรแฝง ซึ่งกลุ่มตัวแปรแฝง (class) เป็นตัวแปรร่วมที่ศึกษาได้จากข้อสอบแต่ละข้อโดยตรงและจากความสามารถใน เช่น ผลสัมฤทธิ์ทางคณิตศาสตร์ ความสามารถในการแก้ปัญหา เป็นต้น

สำหรับโมเดลการตอบสนองข้อสอบแบบผสม (MRM) นี้ Cho และคณะ (2010) ได้เสนอโมเดลดังนี้

ความน่าจะเป็นของการตอบสนองที่ถูกในรูปแบบ logistic ของ MRM สามารถแสดงดังนี้

$$P\left(y_{ijg}=1\middle|g,\theta_{jg}\right)=\frac{1}{1+\exp\left[-\left(\theta_{jg}-\beta_{ig}\right)\right]} \qquad \text{.....(2.9)}$$

เมื่อ g เป็นดัชนีสำหรับ กลุ่มแฝง g = 1,... , G ผู้สอบ j = 1,...,N

θ_{jg} เป็นความสามารถแฝงของผู้สอบ j ภายในกลุ่มที่ g

β_{ig} เป็นพารามิเตอร์ความยากแบบราสช์ (Rasch difficulty) ของข้อสอบ i ในกลุ่มที่ g

โดยที่โครงสร้างของความสามารถใน MRM คือ

$$\theta_{jg} \sim N(\mu_{g,}\sigma_g^2)$$

เมื่อ μ_g เป็นค่าเฉลี่ยเฉพาะกลุ่มของความสามารถ

σ_g^2 เป็นความแปรปรวนเฉพาะกลุ่มของความสามารถ

2. Mixed-Measurement IRT model (MM-IRT) เป็นโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม (Mixture IRT model) แบบสองพารามิเตอร์ (2PL model) เป็นโมเดลหนึ่งที่นำมาใช้สำหรับตัวแปรแฝง สามารถอธิบายกลุ่มของโมเดลที่บูรณาการแนวคิดพหุมิติ และโมเดลกลุ่มแฝง (Rost, 1990, 1991) MM-IRT ระบุกลุ่มแฝงที่มีกลุ่มพารามิเตอร์ข้อสอบเดียวกัน ดังนั้นมีความเท่าเทียมกันของการวัดภายในกลุ่ม แต่ไม่เท่าเทียมกันระหว่างกลุ่ม โดยเฉพาะอย่างยิ่ง ตัวแปรต่อเนื่องถูกจัดวางในการอธิบายสำหรับการแทรกซ้อนระหว่างข้อสอบภายในกลุ่ม

สำหรับโมเดลการตอบสนองข้อสอบแบบสองพารามิเตอร์ (2PL model) ซึ่งถูกเสนอโดย Smit, Kelderman และ Van der Flier (2000) ได้เสนอโมเดลดังนี้

ความน่าจะเป็นของการตอบสนองที่ถูกแสดงดังนี้

$$P\left(y_{jig}=1\middle|g,\theta_{jg}\right)=\frac{1}{1+\exp\left(-\eta_{jig}\right)} \qquad \text{.....(2.10)}$$

เมื่อ $\eta_{jig}=\beta_{ig}+\lambda_{ig}\theta_j$ แสดงถึงความแตกต่างข้ามกลุ่มซึ่งได้ผลโดยจุดตัดของข้อ และน้ำหนักความเฉพาะในกลุ่ม

β_{ig} เป็นพารามิเตอร์ความยากของข้อที่ i ในกลุ่มที่ g

λ_{ig} เป็นพารามิเตอร์อำนาจจำแนกของข้อที่ I ในกลุ่มที่ g และ

θ_{jg} เป็นความสามารถแฝงของคนที่ j ในกลุ่มที่ g

นอกจากนี้สำหรับ Mixed-Measurement IRT model (MM-IRT) Tay และคณะ (2011) ได้เสนอโมเดลดังนี้

ความน่าจะเป็นอย่างมีเงื่อนไขขึ้นอยู่กับกลุ่มแฝง และ คุณลักษณะแฝง θ_j : $P(y_{ji}=1|k,\theta_j)$ ความน่าจะเป็น $P(y_{ji}|k)$ ของเวคเตอร์การตอบสนองสำหรับคนที่ j สามารถแสดงดังนี้

$$P(y_{ji}|k)=\int\prod_{i=1}^{I}P(y_{ji}|k,\theta_j)f(\theta_j)d\theta_j \quad \text{.....(2.11)}$$

เมื่อ $f(\theta_j)$ เป็นความหนาแน่นปกติมาตรฐาน

(2) ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบแบบผสม

ความน่าจะเป็นที่ได้ผลการตอบข้อสอบสามารถจะถูกกำหนดได้จากผลของความน่าจะเป็นในการตอบของแต่ละบุคคล โดยปกติจะถูกเรียกว่าข้อตกลงเบื้องต้นของความเป็นอิสระเฉพาะที่ของการตอบข้อสอบ (local independence) และข้อตกลงเบื้องต้นที่ว่าค่าในโครงสร้างสมมติฐานที่จะถูกประเมินมีค่าเพิ่มขึ้นแล้วความน่าจะเป็นได้มาจากการตอบสนองที่ถูกในข้อคำถามนั้นจะเพิ่มขึ้น โดยปกติจะเรียกว่า Monotonicity assumption หรือรูปแบบของฟังก์ชันทางเดียว ซึ่งเป็นไปตามความน่าจะเป็นของการตอบสนองที่ถูกต้องต่อความสามารถ ซึ่งเป็นข้อตกลงเบื้องต้นที่เหมือนกันของแต่ละบุคคล และข้อสอบ จากข้อตกลงเบื้องต้นนี้นำไปสู่การพัฒนาโมเดลให้มีโครงสร้างโมเดลอย่างง่าย ซึ่งช่วยให้โมเดลมีความเหมาะสมสำหรับการประยุกต์ที่เฉพาะเจาะจง การพัฒนาโมเดลจะสมมติให้ตัวแปรบุคคลเป็นตัวแปรต่อเนื่องโดยอาจเป็น unidimensional หรือ multidimensional ในทางตรงกันข้ามข้อตกลงเกี่ยวกับความต่อเนื่องนี้ จะให้ผลลัพธ์ที่แตกต่างกันอย่างไม่มีขอบเขต

(3) การประมาณค่าพารามิเตอร์ของโมเดลการตอบสนองข้อสอบแบบผสม

กระบวนการการประมาณค่าที่ใช้ในปัจจุบันคือโมเดลตัวแปรบุคคลเป็นพารามิเตอร์แทรกซ้อน (nuisance parameter) ใช้การประมาณพารามิเตอร์ที่ไม่ต่อเนื่องเป็นการแจกแจงพารามิเตอร์บุคคลที่ต่อเนื่อง สำหรับเงื่อนไข Rasch model ใช้การประมาณความน่าจะเป็นสูงสุดแบบมีเงื่อนไข (conditional maximum likelihood) ซึ่งจะเป็นประโยชน์เกี่ยวกับจำนวนของความสามารถที่สามารถสังเกตได้ประมาณให้เท่ากันกับคะแนนรวม (sum score)

สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 และ 3 พารามิเตอร์ (2- 3 PL IRT model) ใช้การประมาณความน่าจะเป็นสูงสุดส่วนริม (the marginal maximum likelihood) ในการประมาณตัวแปรไม่ต่อเนื่องของการรวมกันที่ต่อเนื่องเป็น marginalize ซึ่งการประมาณนี้สามารถประมาณพารามิเตอร์ข้อสอบโดยปราศจากการประมาณพารามิเตอร์ความสามารถร่วมกัน

1.2 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT:MMixIRT)

โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT: MMixIRT) เป็นโมเดลใหม่ เสนอโดย Cho และ Cohen (2010) เป็นวิธีวิทยาการที่เป็นประโยชน์ในการระบุรูปแบบของกลุ่มแฝง โดยเป็นแบบ 1 มิติ (one dimensional Rasch model) ปรากฏกับกรอบแนวคิดโมเดลตัวแปรแฝงทั่วไป ซึ่งสามารถใช้กับโปรแกรมที่มีตัวแปรแฝงได้ (Vermunt & Magidson, 2005) และในภายหลัง Tay และคณะ (2011) ได้เสนอโมเดลที่คล้ายกับ Cho และ Cohen แต่เป็นแบบพหุมิติ (multidimensional 2 PL model) โดยเสนอโมเดลการตอบสนองข้อสอบการวัดแบบผสมพหุระดับ (Multilevel Mixed-measurement IRT : MMMIRT)

จากมโนทัศน์เกี่ยวกับการวัดโดยทั่วไปจะวัดจากตัวแปรที่สังเกตได้ โดยสมมติให้มีโมเดลการวัดเพียงโมเดลเดียวสำหรับประชากร อย่างไรก็ตาม วิธีวิทยาการที่สูงขึ้นนำไปสู่การสังเคราะห์ของแนวคิดเกี่ยวกับมิติ (dimension) และกลุ่มแฝง (latent class) ซึ่งนำไปใช้ในการวิเคราะห์กลุ่มแฝง ที่มีโมเดลการวัดแตกต่างกัน (Rost, 1990, 1991) ในทางตรงข้าม แนวคิดเกี่ยวกับ การใช้กลุ่มก่อนหน้า แนวคิด bottom-up (bottom-up approach) ถูกประยุกต์ใช้ในการวิเคราะห์ และผลแสดงถึงระดับการวัดที่แตกต่าง ในองค์กร หรือในโรงเรียน ซึ่งมีความเป็นระดับขั้นโดยแท้จริง (Kozlowski & Klein, 2000) แนวคิดเกี่ยวกับพหุระดับ ซึ่งมีชื่อว่า multilevel mixture IRT analysis (MMixIRT) ถูกใช้สำหรับข้อมูลที่สอดแทรก (nested data) ในระดับกลุ่มของแต่ละบุคคล กับโมเดลการวัดปกติครอบคลุมหน่วยข้ามระดับ มากไปกว่านี้การจำแนกกลุ่ม จำแนกจากการแบ่งส่วนที่คล้ายกันของกลุ่มในระดับบุคคล (individual-level classes) ซึ่งส่งผลถึงกลุ่มแฝงที่ระดับแตกต่างกัน แนวคิดของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) เป็นแหล่งที่มาของกลุ่มการวัดโครงสร้างของความสนใจ (เช่น บุคลิกภาพ เจตคติ และบรรยากาศ) ซึ่งไม่เกี่ยวกับหน่วยในกลุ่มแฝงที่ถูกสังเกต นอกจากนี้ความแตกต่างที่อาจเกิดขึ้นในแต่ละบุคคลในการนิยามโครงสร้างหรือการใช้มาตรวัด และความสอดคล้องกัน (commonalities) ระหว่างหน่วยของระดับชั้น จะแสดงได้จากการวิเคราะห์ด้วยโมเดลดังกล่าว (Eid & Rauber, 2000; Hernández et al., 2004; Maij-de Meij, Kelderman, & van der Flier, 2008; Zickar, Gibby, & Robie, 2004) โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับเป็นโมเดลที่มีประโยชน์ มีคุณค่าในการวิจัยในองค์กร เป็นการพัฒนาวิธีวิทยาการ กรอบแนวคิดใหม่และระบุประเด็นทางทฤษฎี

ในที่นี้จะเสนอแนวคิดเกี่ยวกับการวิเคราะห์พหุระดับ เพื่อทำความเข้าใจเกี่ยวกับโมเดล จากนั้นจะนำเสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixture IRT)

1.2.1 การวิเคราะห์พหุระดับ

ลักษณะทั่วไปของการจัดการองค์กรทางการศึกษา นักเรียนจะถูกสอนร่วมกันเป็นชั้นเรียนภายในโรงเรียน ซึ่งโรงเรียนก็จะอยู่ในแต่ละสังกัด แต่ละภูมิภาค แต่ละอำเภอแตกต่างกันไป ข้อมูลทางการศึกษาจึงมักเป็นข้อมูลที่เป็นหลายระดับ (multilevel data) สอดแทรกกันอยู่ (hierarchical nested data) ซึ่งเป็นผลมาจากที่ระบบทางการศึกษาเป็นระบบที่มีความซับซ้อน การบริหารงานมีโครงสร้างเป็นหน่วยที่ซ้อนกันเป็นลำดับชั้น โดยหน่วยที่อยู่ระดับล่างสุดจะได้รับอิทธิพลร่วมกันจากตัวแปรที่อยู่ในระดับที่สูงกว่า (Raudenbush & Bryk, 1986) Cronbach และ Webb (1975) ได้ศึกษาอิทธิพลของตัวแปรภายนอกที่มีอิทธิพลต่อการเรียนรู้ของผู้เรียน และพบว่าความถนัดมีอิทธิพลปฏิสัมพันธ์กับตัวแปรความแตกต่างระหว่างห้องเรียนมากกว่าในระดับบุคคล ซึ่งแตกต่างจากงานวิจัยก่อนหน้าที่พบว่ามีความสัมพันธ์ของตัวแปรความถนัดกับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ว่ามีปฏิสัมพันธ์มากกับผู้เรียน ซึ่งในอดีตมีข้อจำกัดเกี่ยวกับสถิติวิเคราะห์ ทำให้มีความคลาดเคลื่อนของการสรุปผลการวิจัยเกิดขึ้น ดังนั้นหากเกิดความผันแปรขึ้นไปตามบริบทของโรงเรียนหรือห้องเรียน การวิเคราะห์ในแบบเดิมที่มีการบีบบังคับให้ตัวแปรต่างระดับให้เสมือนอยู่ในระดับเดียวกัน อาจจะทำให้ผลการวิเคราะห์เกิดความผิดพลาดเพราะไม่สนใจความแตกต่างระหว่างหน่วยในการวิเคราะห์ ซึ่งไม่สามารถคำนวณค่าความแปรปรวนภายในหน่วยหรือกลุ่ม (within group variability) จึงเป็นการละเลยต่อการศึกษาปฏิสัมพันธ์ระหว่างตัวแปรที่อยู่ต่างระดับกัน ถึงแม้จะมีการออกแบบการวิจัยอย่างรอบคอบก็ตาม

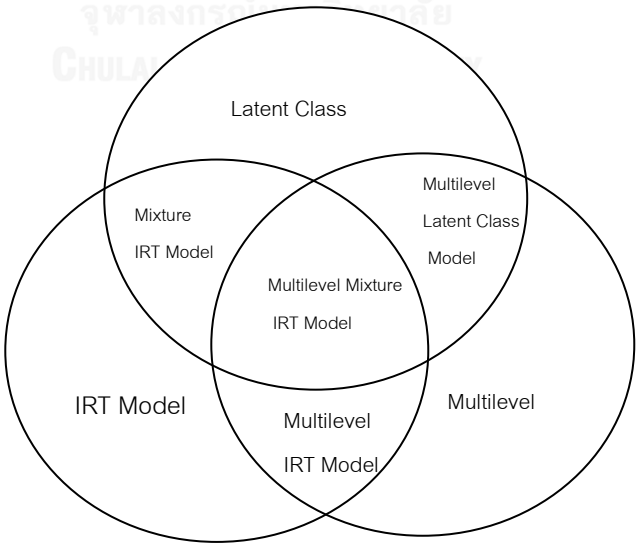
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1.2.2 โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel mixture IRT)

Multilevel model หรือที่เรียกว่า Hierarchical linear model (HLM) จะวิเคราะห์โครงสร้างระดับที่เป็นธรรมชาติของข้อมูลทางการศึกษาและจิตวิทยา (Bryk & Stephen, 2002) การบูรณาการของโมเดลพหุระดับ โมเดลการตอบสนองข้อสอบ และโมเดลกลุ่มแฝง เป็นประโยชน์มากซึ่งสามารถประมาณค่าที่ถูกต้องของความคลาดเคลื่อนมาตรฐานของโมเดล (J. P. Fox, 2005) การบูรณาการนี้นำไปสู่การพัฒนาโมเดลการวัดทางจิตมิติ สำหรับข้อมูลการตอบสนองข้อสอบซึ่งประกอบด้วยโครงสร้างลำดับขั้น ดังนั้นนักวิจัยจึงศึกษาถึงผลกระทบของตัวทำนายที่แตกต่างกัน ซึ่งเป็นการลดข้อจำกัด เช่น โรงเรียน หลักสูตร ที่มีต่อระดับที่ต่ำกว่า เช่น นักเรียน (Kamata, 2001) ตามทฤษฎีการตอบสนองข้อสอบแบบผสม ข้อจำกัดที่ว่าไม่สามารถจัดการสารสนเทศเกี่ยวกับสมาชิกในกลุ่มโดยให้ตัวทำนายในระดับลดหลั่นนี้ ซึ่งเป็นตัวทำนายที่เด่นชัดรวมในโมเดล

การบูรณาการโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้ เป็นการพัฒนาโมเดลโดยบูรณาการโมเดลพหุระดับ กับโมเดลการตอบสนองข้อสอบ (J.-P. Fox & Glas, 2001) และโมเดล

กลุ่มแฝง (Vermunt & Magidson, 2003) การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้ สามารถแบ่งออกเป็น 3 ลักษณะ ได้แก่ (a) MM-IRT ถูกพัฒนาโดยการสร้างแบบจำลองแบบเป็นลำดับขั้น (hierarchical classes) (b) multilevel latent class model (MLC) (Vermunt, 2008a; Vermunt & Magidson, 2003) ซึ่งถูกพัฒนาโดยการเพิ่มคุณลักษณะ ที่ระดับล่าง และ (c) multilevel mixture IRT Model (Vermunt, 2008b) ถูกพัฒนาโดยการเพิ่มการประมาณค่าที่ระดับบุคคล (individual classes) จากที่ 3 โมเดลข้างต้น มีเพียงส่วนประกอบหลัก 2 ส่วน ของโมเดล คือ individual classes individual traits และ hierarchical classes แต่ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับจะมีทั้ง 3 ส่วน ทั้งนี้ Cho และ Cohen (2010) และ Tay และคณะ (2011) ได้เสนอโมเดลที่คล้ายกัน แต่มีแนวคิดที่แตกต่างกันโดย Cho และ Cohen จำกัดขอบเขตที่โมเดล 1 พารามิเตอร์ (one-dimensional Rasch models) และใช้การประมาณค่าแบบเบส์เซียน (Bayesian) ส่วน Tay และคณะ จะจำกัดขอบเขตเป็นโมเดล 2 พารามิเตอร์ (multidimensional 2PL model) และใช้การประมาณค่าความน่าจะเป็นสูงสุด (maximum likelihood :ML) ซึ่งในการศึกษานี้จะใช้กับ โมเดล 3 พารามิเตอร์ (multidimensional 3PL model) และใช้การประมาณค่าโดยวิธีเอ็มไพริคัล เบส์ (Empirical Bayes) และพิจารณาความเหมาะสมของการประมาณค่าความน่าจะเป็นสูงสุด ซึ่งสามารถกล่าวได้ว่าโมเดลนี้เป็นโมเดลที่ใหม่ และเหมาะสมกับข้อมูลที่มีลักษณะสอดแทรกกัน (nested data) และมีการให้คะแนนแบบสองค่า และใช้ในการสำรวจกลุ่มแฝงของแต่ละบุคคล ซึ่งมีโมเดลการวัดเดียวกันข้ามกลุ่มสังกัด ซึ่งการรวมกันของ 3 โมเดลเสนอเป็นแผนภาพเวนน์ได้ดังนี้



ภาพที่ 2.3 แผนภาพเวนน์แสดงการบูรณาการกันเป็นโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Cho , 2007)

ในการพัฒนาโมเดลผู้วิจัยได้ศึกษารายละเอียดในหลายองค์ประกอบที่เกี่ยวข้องของโมเดลการตอบสนองข้อสอบ โมเดลกลุ่มแฝง โมเดลการตอบสนองข้อสอบแบบผสม และโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ที่พัฒนาโดย Cho และ Cohen (2010) และ Tay et al. (2011) ดังหัวข้อต่อไปนี้ (1) โครงสร้างความสามารถ (2) โครงสร้างสัดส่วนของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (3) โครงสร้างค่าความยาก (4) โครงสร้างอำนาจจำแนก (5) โครงสร้างความน่าจะเป็นของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (6) ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (7) ประเด็นความแตกต่างของประชากร

(1) โครงสร้างความสามารถ Cho และ Cohen (2010) ได้เสนอโครงสร้างความสามารถดังรายละเอียดต่อไปนี้

โมเดลการตอบสนองข้อสอบ ตัวแปรแฝง (ความสามารถ ; θ) ในโมเดลการตอบสนองข้อสอบ แสดงดังนี้

$$\theta_j \sim N(0,1)$$

เมื่อ j = 1,..., J

โมเดลกลุ่มแฝง ตัวแปรแฝง (ความสามารถ ; θ) ในโมเดลกลุ่มแฝงแสดงดังนี้

$$\theta_j = \sum_{g=1}^{G} \pi_g \cdot \theta_g \qquad \text{.....(2.12)}$$

เมื่อ θ_g เป็นความสามารถในกลุ่มแฝงที่ g

π_g เป็นสัดส่วนของนักเรียนที่ตอบถูกในชั้น g

โมเดลพหุระดับ

โมเดลพหุระดับเป็นโมเดลที่ไม่มีตัวแปรแฝงสำหรับการรวมความคลาดเคลื่อนทางการวัดเข้าด้วยกัน อย่างไรก็ตามผลกระทบอย่างสุ่ม (random effect) ในโมเดลพหุระดับ ถูกสมมติให้มีการแจกแจงแบบปกติ ในแบบเดียวกันกับพารามิเตอร์ความสามารถ

ในโมเดลการตอบสนองข้อสอบ การบูรณาการระหว่างโมเดลการตอบสนองข้อสอบ และโมเดลพหุระดับ เป็นโมเดลการตอบสนองข้อสอบพหุระดับ (Multilevel IRT) สำหรับกรณีที่ตัวแปรแฝงถูกสมมติให้เป็นตัวแปรไม่ต่อเนื่อง ลักษณะนี้เหมือนกับกรณีของโมเดลกลุ่มแฝง ซึ่งเป็นการบูรณาการกันกับโมเดลพหุระดับ เป็นโมเดลกลุ่มแฝงพหุระดับ (Multilevel LCM) การสมมติตัวแปรลักษณะนี้สามารถนำสู่การบูรณาการกันของโมเดลซึ่งคล้ายกับโมเดลการตอบสนองข้อสอบพหุระดับ ซึ่งเป็นโมเดลที่รวมกันระหว่างโมเดลพหุระดับแบบไม่เป็นเส้นตรง (multilevel nonlinear model) และโมเดลการตอบสนองข้อสอบ (IRT Model) และโมเดลสมการโครงสร้างพหุระดับ (multilevel SEM) ซึ่งเป็นโมเดลที่บูรณาการระหว่างโมเดลพหุระดับแบบเป็นเส้นตรง (multilevel

linear model) และโมเดลสมการโครงสร้าง (SEM) ในที่นี้จะสนใจโมเดลพหุระดับแบบไม่เป็นเส้นตรง (multilevel nonlinear model)

โมเดลการตอบสนองข้อสอบแบบผสม ตัวแปรแฝง (ความสามารถ ; θ) ในโมเดลการตอบสนองข้อสอบแบบผสม แสดงดังนี้

$$\theta_{jg} \sim N\left(\mu_g, \sigma_g^2\right)$$

เมื่อ j หมายถึง นักเรียนคนที่ j โดยที่ j = 1, ... , J และ

g หมายถึงกลุ่มแฝงระดับนักเรียน

μ_g หมายถึง ค่าเฉลี่ยของความสามารถสำหรับกลุ่มที่ g

σ_g^2 หมายถึง ความแปรปรวนของความสามารถสำหรับกลุ่มที่ g

โมเดลการตอบสนองข้อสอบพหุระดับ โมเดลของตัวแปรแฝง (ความสามารถ ; θ) ในโมเดลการตอบสนองข้อสอบพหุระดับ แสดงดังนี้

$$\theta_{jt} = \gamma_{00} + u_{jt} + v_t \qquad \text{.....(2.13)}$$

เมื่อ j หมายถึง นักเรียนคนที่ j โดยที่ j = 1, ... , J

t หมายถึง เป็นค่าเฉลี่ยความสามารถของโรงเรียนที่ t โดยที่ t= 1,..., T

u_{jt} เป็นค่าเฉลี่ยความสามารถข้ามระดับนักเรียน u_{jt} จะเป็นไปตาม $N(0,\tau)$ เมื่อ τ เป็นความแปรปรวนของความสามารถที่ระดับนักเรียน

v_t จะเป็นไปตาม $N(0,\zeta)$ เมื่อ ζ เป็นความแปรปรวนของความสามารถที่ระดับโรงเรียน ความแปรปรวนรวมของความสามารถในโมเดล Multilevel IRT ถูกจำแนกไปในความแปรปรวนของความสามารถระดับนักเรียน (i.e., τ) และความแปรปรวนของความสามารถระดับโรงเรียน (i.e., ζ)

โมเดลกลุ่มแฝงพหุระดับ โมเดลของตัวแปรแฝง (ความสามารถ ; θ) ในโมเดลกลุ่มแฝงพหุระดับแสดงดังนี้

$$\theta_{jt} = \sum_{k=1}^{K}\sum_{g=1}^{G} \pi_k \cdot \pi_{g|k} \cdot \theta_{gk} \qquad \text{.....(2.14)}$$

เมื่อ j หมายถึง นักเรียนคนที่ j โดยที่ j = 1, ... , J

t หมายถึงโรงเรียนที่ t โดยที่ t= 1,..., T

g หมายถึง กลุ่มแฝงในระดับนักเรียน โดยที่ g= 1,...,G และ

k หมายถึงกลุ่มแฝงระดับโรงเรียน โดยที่ k=1,...,K

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ โมเดลของตัวแปรแฝง (ความสามารถ; θ) ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับในการศึกษาครั้งนี้ แสดงดังนี้

$$\theta_{jtgk} \sim N\left(\mu_{gk}, \sigma^2_{gk}\right)$$

เมื่อ j หมายถึง นักเรียนคนที่ j โดยที่ j = 1, ... , J

t หมายถึงโรงเรียนที่ t โดยที่ t = 1,..., T

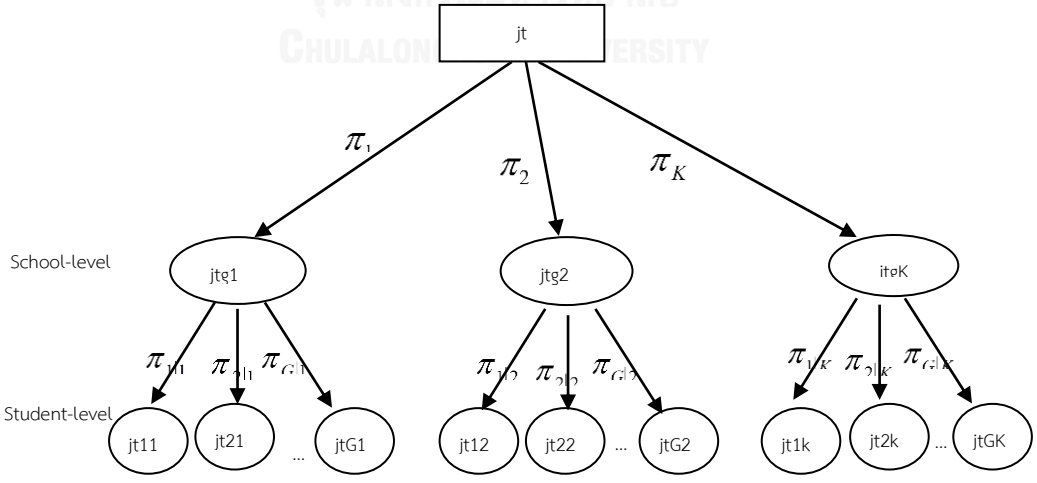
g หมายถึง กลุ่มแฝงในระดับนักเรียน โดยที่ g= 1,...,G

k หมายถึงกลุ่มแฝงระดับโรงเรียน โดยที่ k=1,...,K

μ_{gk} หมายถึง ค่าเฉลี่ยของความสามารถสำหรับกลุ่มที่ g และ k

σ^2_{gk} หมายถึง ความแปรปรวนของความสามารถสำหรับกลุ่มที่ g

จากโมเดลของตัวแปรแฝง (ความสามารถ; θ) ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับข้างต้น เป็นการอธิบายโครงสร้างของข้อมูลในเทอมของนักเรียน ที่ซ้อนกันภายในโรงเรียน กับความน่าจะเป็นของการผสมที่ระดับนักเรียน และระดับโรงเรียน นักเรียนแต่ละคนถูกจำแนกอย่างชัดเจน ตาม j และ t โดยนักเรียนคนที่ j ซ้อนอยู่ในโรงเรียน t ในเงื่อนไขนี้ นักเรียนจะมีการแยกคุณลักษณะแฝงตาม g และ k ความสามารถของนักเรียน (θ_{jtgk}) มีความแปรปรวนสำหรับแต่ละหน่วยการผสมของ g และ k ซึ่ง มีค่าเฉลี่ย μ_{gk} และ ความแปรปรวน σ^2_{gk} เทียบกับความน่าจะเป็นในระดับนักเรียนของการผสม $\pi_{g|k}$ และความน่าจะเป็นในระดับโรงเรียนของการผสม π_k โครงสร้างนี้สามารถแสดงดังภาพ



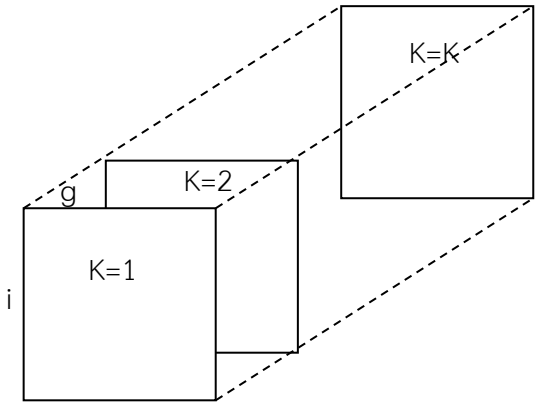
ภาพที่ 2.4 โครงสร้างของข้อมูลพหุระดับ (Cho & Cohen, 2010)

Cohen (2010) ได้เสนอโครงสร้างสัดส่วนของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ดังตารางต่อไปนี้ เมื่อโครงสร้างของ $\pi_{1:G|k}$ ซึ่งแสดงในตารางต่อไปมีความน่าจะเป็น K ในแถว $\pi_{1:G|k}$ เมื่อ k=1,...,K และ G เป็นมิติของแต่ละแถว

ตารางที่ 2.3 โครงสร้างสัดส่วนของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Cho & Cohen, 2010)

	K=1	K=2	. .	K=K
G=1	$\pi_{1\|1}$	$\pi_{1\|2}$		$\pi_{1\|K}$
G=2	$\pi_{2\|1}$	$\pi_{2\|2}$		$\pi_{2\|K}$
:	:	:		:
G=G	$\pi_{G\|1}$	$\pi_{G\|2}$		$\pi_{G\|K}$
รวม	$\sum_{g=1}^{G}\pi_{g\|1}=1$	$\sum_{g=1}^{G}\pi_{g\|2}=1$		$\sum_{g=1}^{G}\pi_{g\|K}=1$

(3) โครงสร้างค่าความยากของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ Cho และ Cohen (2010) เสนอโครงสร้างค่าความยาก และกล่าวว่าโมเดลโดยทั่วไป ของ โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ พารามิเตอร์ความยากมีค่าเฉพาะของกลุ่มทั้งในระดับนักเรียน และระดับโรงเรียน สามารถเขียนแทนด้วยสัญลักษณ์ β_{igk} ความสัมพันธ์นี้สามารถแสดงได้ดังภาพที่ 2.5 ซึ่งหลัก (column) แสดงข้อที่ i และ แถว (row) แทนการผสมระดับนักเรียน g และ การผสมระดับโรงเรียน k นี้แสดงให้เห็นว่า ความยากของข้อสอบของกลุ่มที่ระดับนักเรียน เพิ่มมากขึ้นที่การผสมระดับโรงเรียน



ภาพที่ 2.5 ความชัดเจนของความยากของข้อสอบในกลุ่ม (Cho & Cohen, 2010)

(4) โครงสร้างค่าอำนาจจำแนกของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

Cho และ Cohen (2010) เสนอโครงสร้างค่าอำนาจจำแนก และกล่าวว่าโมเดลโดยทั่วไปของ โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ พารามิเตอร์อำนาจจำแนกมีค่าเฉพาะของกลุ่มทั้งในระดับนักเรียน และระดับโรงเรียน สามารถเขียนแทนด้วยสัญลักษณ์ λ_{igk}

(5) โครงสร้างความน่าจะเป็นของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

สำหรับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ เป็นการรวมกันของโมเดล 3 โมเดล ได้แก่ โมเดลการตอบสนองข้อสอบ โมเดลกลุ่มแฝง และโมเดลพหุระดับ โดย Cho และ Cohen (2010) เสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) โดยเสนอวิธีการ One-dimensional Rasch model ในการศึกษาตัวแปรแฝงความสามารถทางคณิตศาสตร์ในระดับนักเรียน และระดับโรงเรียน ส่วน Tay และคณะ (2011) เสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Mix-Measurement IRT) โดยเสนอวิธีการ multidimensional 2-PL model ในการศึกษาตัวแปรแฝงการรายงานอารมณ์ตนเองในระดับบุคคล และระดับประเทศ ดังนั้นในที่นี้จึงเสนอรายละเอียดของโครงสร้างความน่าจะเป็นของ Cho และ Cohen (2010) และของ Tay และคณะ (2011) ดังนี้

(5.1) โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบ Rasch model

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบ Rasch model เสนอโดย Cho and Cohen (2010) สำหรับความน่าจะเป็นในการตอบข้อสอบถูก ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) แบบ 1 พารามิเตอร์ (Rasch model) แสดงดังนี้

$$P\left(y_{ijt}=1\middle|gk,\theta_{jtgk}\right)=\frac{1}{1+\exp\left[-\left(\theta_{jtgk}-\beta_{igk}\right)\right]} \qquad \text{.....(2.15)}$$

เมื่อ g = 1,... , G เป็นดัชนีสำหรับกลุ่มแฝงระดับนักเรียน

k = 1,... K เป็นดัชนีสำหรับกลุ่มแฝงระดับโรงเรียน

j = 1,...,J เป็นดัชนีสำหรับผู้สอบ

t = 1,..., T เป็นดัชนีสำหรับโรงเรียน

i = 1,...,I เป็นข้อในแบบสอบ

θ_{jtgk} เป็นความสามารถแฝงของผู้สอบ j ในโรงเรียน t และในกลุ่มแฝง g และ k

β_{igk} เป็นพารามิเตอร์ความยากของข้อสอบ i สำหรับกลุ่มแฝง g และ k

(5.2) โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบสองพารามิเตอร์

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบสองพารามิเตอร์ เสนอโดย Tay, Diener, et al. (2011) ซึ่ง Tay และคณะ นำเสนอเป็นโมเดลการตอบสนองข้อสอบแบบผสมโดยนำไปใช้กับข้อมูลซึ่งมีการวิเคราะห์แบบพหุระดับ ดังรายละเอียดต่อไปนี้

โดยทั่วไปโมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT model) อธิบายกลุ่มของโมเดล ซึ่งรวมแนวคิดพหุมิติ (dimensional) และ โมเดลกลุ่มแฝง (Latent Class) (Rost, 1990, 1991) โมเดลการตอบสนองข้อสอบแบบผสมการวัด (Mixed-Measurement model) หมายถึงกลุ่มแฝงในแต่ละกลุ่มของบุคคล ซึ่งอยู่ในเซตเดียวกันของพารามิเตอร์ข้อสอบ ดังนั้นมีความไม่แปรเปลี่ยนของการวัดภายในกลุ่มซึ่งไม่เท่าเทียมกันระหว่างกลุ่ม โดยเฉพาะอย่างยิ่งสำหรับตัวแปรต่อเนื่องซึ่งถูกนำมาคำนวณร่วมกับข้อสอบภายในกลุ่ม อย่างไรก็ตามการบูรณาการโมเดลการตอบสนองข้อสอบ และโมเดลกลุ่มแฝง สามารถเรียกว่าเป็นโมเดลการการตอบสนองข้อสอบแบบผสม (Mixture IRT)

โมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT model) แสดงดังสมการด้านล่าง อย่างไรก็ตาม ความน่าจะเป็นแบบมีเงื่อนไขขึ้นอยู่กับกลุ่มแฝงและลักษณะแฝง (latent trait) $\theta_j : P\left(y_{ji}=1\middle|k,\theta_j\right)$ ดังนั้นสำหรับกลุ่มข้อสอบ ความน่าจะเป็น $P\left(\underset{\sim}{y_j}\middle|k\right)$ ของเวกเตอร์การตอบสนองสำหรับบุคคลที่ j สามารถเขียนแทนด้วย $\int\prod_{i=1}^{I}P\left(y_{ji}\middle|k,\theta_j\right)f\left(\theta_j\right)d\theta_j$ เมื่อ $f\left(\theta_j\right)$ แทนความหนาแน่นปกติมาตรฐาน (standard normal density)

สำหรับความน่าจะเป็นในการตอบสนองที่ถูก ในโมเดลการตอบสนองข้อสอบแบบผสมแบบพหุมิติ (2PL model) แสดงดังนี้

$$P\left(y_{ji}=1\middle|k,\theta_j\right)=\frac{1}{1+\exp\left(-\eta_{jik}\right)} \qquad \text{.....(2.16)}$$

ในที่นี้ $\eta_{jik}=\beta_{ik}+\lambda_{ik}\theta_j$ ถูกยอมให้แตกต่างกันข้ามกลุ่มแฝง (Latent Class) โดยยอมให้ item intercept และ loading เป็นกลุ่มเฉพาะ (k)

และ $\theta_{jk}\sim N\left(\mu_k,\sigma_k^2\right)$ เมื่อ j หมายถึง นักเรียนคนที่ j ซึ่ง j = 1, ..., J และ k หมายถึงกลุ่มแฝงในระดับนักเรียน μ_k แทนค่าเฉลี่ยของความสามารถของชั้นที่ k และ σ_k^2 หมายถึงความแปรปรวนของความสามารถสำหรับชั้นที่ k

เนื่องจากโมเดลการตอบสนองข้อสอบแบบผสมเป็นการบูรณาการโมเดลการตอบสนองข้อสอบ และโมเดลกลุ่มแฝง Tay และคณะ (2011) จึงเสนอรายละเอียดเกี่ยวกับทั้งสองโมเดลดังนี้

โมเดลการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์

ความน่าจะเป็นของโมเดลการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ แสดงดังนี้

$$P\left(y_{ji}=1\middle|\theta_j\right)=\frac{1}{1+\exp\left(-\eta_{ji}\right)} \qquad \text{..... (2.17)}$$

โดยที่ $\eta_{ji}=\beta_i+\lambda_i\theta_j$ และ สัมประสิทธิ์ β_i และ λ_i แทนความง่ายของข้อสอบ และอำนาจจำแนกของข้อสอบ

โมเดลนี้จะคล้ายกับ item intercept และ item loading ในการวิเคราะห์องค์ประกอบ (Factor analysis) ซึ่งความแตกต่างหลัก ๆ คือ logistic function ในที่นี้แทนโมเดลเชิงเส้นสำหรับข้อสอบ เพื่อความสอดคล้องกัน ในที่นี้จะหมายถึง β_i และ λ_i เป็น item intercept และ item loading ตามลำดับ

ความน่าจะเป็นสำหรับทุกการตอบสนองของบุคคล j $\left(y_j\right)$ มีข้อตกลงว่า มีความเป็นอิสระของ θ_j ซึ่งข้อตกลงเบื้องต้นคือ local independence ในที่นี้ $P\left(y_j\middle|\theta_j\right)=\Pi_{i=1}^{I}P\left(y_{ji}\middle|\theta_j\right)$ และ $P\left(y_j\right)$ ได้จากการรวม (integrating) อย่างครอบคลุมความน่าจะเป็นของ θ_j

โมเดลกลุ่มแฝง

ความน่าจะเป็นของโมเดลกลุ่มแฝง แสดงดังนี้

$$P\left(\underset{\sim}{y_j}\right)=\sum_{k=1}^{K}\pi_k P\left(\underset{\sim}{y_j}\middle|k\right) \qquad \text{..... (2.18)}$$

สมการนี้ความน่าจะเป็นส่วนน้อยของการตอบสนองแต่ละบุคคล $\left(y_j\right)$ ถึงกลุ่มคำตอบข้อ i ให้กลุ่มแฝง k (k= 1,...,K) ซึ่งขึ้นอยู่กับสองส่วนคือ ความน่าจะเป็น π_k ของความสัมพันธ์กับกลุ่มแฝง k และ ความน่าจะเป็นแบบมีเงื่อนไข $P\left(\underset{\sim}{y_j}\middle|k\right)$ ของการตอบสนองของกลุ่มของข้อสอบ i ให้กลุ่มแฝงเป็น k

หมายเหตุ ผลรวมความน่าจะเป็นของโมเดลกลุ่มแฝง (Latent Class) เป็น 1 $\left(\sum_{k=1}^{k}\pi_k=1\right)$ คล้ายกับโมเดลการตอบสนองข้อสอบ (IRT Model) ความน่าจะเป็น $P\left(y_j\middle|k\right)$ ถูกจำกัดโดยข้อตกลงว่าการตอบสนอง i เป็น local independent ของอีกตัวหนึ่ง (ในที่นี้ความอิสระที่เกิดร่วมกันให้เป็นสมาชิกกลุ่มเดียวกัน) เป็น

$$P\left(\underset{\sim}{y_j}\middle|k\right)=\prod_{i=1}^{I}P\left(y_{ji}\middle|k\right) \qquad \text{.....(2.19)}$$

(6) ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT Model)

Tay และ คณะ (2011) ได้กล่าวถึงข้อตกลงเบื้องต้นดังรายละเอียดต่อไปนี้

1. โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT Model) มีข้อตกลงเบื้องต้นเหมือนกับโมเดลการตอบสนองข้อสอบแบบผสม (MixIRT) และ โมเดลกลุ่มแฝง (Latent Class) สำหรับการวิเคราะห์ในหน่วยระดับบุคคล (individual – level unit) ที่มีโมเดลการวัดที่แตกต่างกัน อธิบายเพิ่มเติมใน Vandenberg และ Lance (2000) ว่าโมเดลการวัดที่แตกต่างกันส่งผลที่แตกต่างในนิยามตามทฤษฎีเกี่ยวกับโครงสร้าง เช่น คะแนนไม่สามารถเทียบกันระหว่างกลุ่มแฝง (Latent classes) เนื่องจากมีความแตกต่างกันเชิงคุณภาพ ในทางตรงกันข้าม การเปรียบเทียบภายในกลุ่มแฝงสามารถทำได้ ข้อตกลงเบื้องต้นทางสถิติ คือ การตอบสนองข้อสอบของแต่ละบุคคล มีเงื่อนไขที่มีความเป็นอิสระอย่างสมบูรณ์บนคุณสมบัติและสมาชิกในกลุ่มแฝง

2. ความไม่เป็นอิสระที่ซ้อนกัน (Nested dependencies) กลุ่มแฝงของระดับลดหลั่น (hierarchical units) มีอิทธิพลต่อความน่าจะเป็นของกลุ่มแฝงระดับบุคคล (individual-level LCs) แต่ไม่มีผลต่อการตอบสนองข้อสอบโดยตรง หรืออีกนัยหนึ่งผลของหน่วยระดับลดหลั่น (hierarchical units) ที่มีต่อการตอบสนองข้อสอบ เป็นตัวแปรสอดแทรกอย่างสมบูรณ์ (fully mediated) โดยสมาชิกของกลุ่มแฝงระดับบุคคล (individual-level class membership) เช่น ในสภาพแวดล้อมเกี่ยวกับองค์กร คุณลักษณะเกี่ยวกับแผนกอาจมีอิทธิพลต่อมุมมองของแต่ละคน

3. หน่วยระดับลดหลั่น (hierarchical units) กลุ่มแฝงระดับลดหลั่น จะพิจารณาสำหรับการใช้ตัวแปรแฝงระดับนามบัญญัติที่ระดับลดหลั่นที่เป็นกลุ่มแฝงระดับเดียวกัน โดยในระดับลดหลั่นจะจำแนกกลุ่มความสามารถแฝงตามสัดส่วนของระดับบุคคลที่คล้ายกัน สำหรับข้อตกลงเบื้องต้นเกี่ยวกับหน่วยระดับลดหลั่นอธิบายเพิ่มเติมได้ดังตัวอย่างต่อไปนี้ การจำแนกกลุ่มความสามารถแฝงระดับลดหลั่นจำแนกตามสัดส่วนที่คล้ายกัน เช่น ในแผนกขายขององค์กร มี 80% ของทั้งหมดอธิบายโดยโมเดลการวัดเกี่ยวกับความพอใจในงานแบบที่ 1 และ 20% ของทั้งหมด อธิบายโดยโมเดลการวัดเกี่ยวกับความพอใจในงานแบบที่ 2 จะถูกจำแนกรวมกับแผนกโฆษณาและแผนกความสัมพันธ์ส่วนรวม ซึ่งมีสัดส่วนของโมเดลการวัดที่คล้ายกัน ในขณะที่แผนกบัญชีมี 20 % อธิบายโดยโมเดลการวัดเกี่ยวกับความพอใจในงาน แบบที่ 1 และ 80 % อธิบายโดยโมเดลการวัดเกี่ยวกับความพอใจในงาน แบบที่ 2 ซึ่งจะถูกจำแนกรวมกับแผนกวิศวะ ซึ่งมีสัดส่วนของโมเดลการวัดที่คล้ายกัน

จากรายละเอียดข้างต้น จะเห็นว่าจะต้องมีการจัดการสารสนเทศ และวิเคราะห์ข้อตกลงเบื้องต้นอย่างครบถ้วน สำหรับการใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ(MMixIRT)

(7) ประเด็นของความแตกต่างของประชากร

ในการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ มีความจำเป็นที่จะต้องเข้าใจเรื่องของความแตกต่างที่สังเกตได้ และสังเกตไม่ได้ การจำแนกสิ่งที่สังเกตได้ และสังเกตไม่ได้ ทำให้เข้าใจเกี่ยวกับความแตกต่างที่ถูกใช้ และการเป็นตัวแทนของระดับพหุในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

ความแตกต่างหรือความเป็นวิวิธพันธ์ (heterogeneity) ในที่นี้หมายถึง ประชากรย่อย (subpopulation) ที่แตกต่างกัน ซึ่งไม่จำเป็นในการเชื่อมโยงกับความผันแปรของคะแนน ตัวอย่างเช่น นักวิจัยอาจจะรวบรวมข้อมูลจากประชากรที่แตกต่างกันประกอบด้วย 2 กลุ่มประชากรย่อย (เช่น เพศชาย และเพศหญิง) ผลที่ตามมาคือ คุณลักษณะของเพศในกลุ่มประชากรย่อยนี้มีความเหมือนกัน หรือเป็นเอกพันธ์ (homogeneous) โดยทั่ว ๆ ไป ข้อมูลเกี่ยวกับองค์กรเป็นตัวอย่างจากประชากรที่มีความแตกต่างกัน ซึ่งจะแบ่งออกเป็น 2 รูปแบบ คือความแตกต่างที่สังเกตได้ (observed) กับความแตกต่างที่ไม่สามารถสังเกตได้ (unobserved) (Skrondal & Rabe-Haasbetg, 2004 cited in Tay et.al., 2011)

ความแตกต่างที่สังเกตได้ (observed) ภายในประชากร สามารถถูกกำหนดโดยการใช้ตัวแปรที่สังเกตได้ ซึ่งถูกจัดไว้ก่อนล่วงหน้า ตัวอย่างเช่น ตัวแปรเกี่ยวกับประชากรแต่ละคน ถูกระบุเพศไว้ก่อนล่วงหน้า หรือ ถูกระบุชาติพันธ์วรรณะ สีผิว ซึ่งเป็นการเปรียบเทียบกลุ่มพหุที่ระบุจากตัวแปรที่สังเกตได้ อาจระบุบนพื้นฐานของบริบทเกี่ยวกับกฎหมาย การเมือง (เช่น กลุ่มเสียงข้างมาก หรือกลุ่มเสียงข้างน้อย) การออกแบบการวิจัย (เช่น กลุ่มทดลอง กลุ่มควบคุม) การพิจารณาตามทฤษฎี (เช่น ประเทศ A ประเทศ B) ความแตกต่างที่สังเกตได้ถูกใช้บ่อย ๆ สำหรับการอธิบายและการวินิจฉัยกลุ่มประชากรย่อย

อย่างไรก็ตาม ความแตกต่างที่ไม่สามารถสังเกตได้ (unobserved heterogeneity) จะถูกสรุปจากข้อมูล เนื่องจากไม่สามารถทราบว่ากลุ่มประชากรย่อยไหนที่แต่ละคนเป็นเจ้าของ แต่ละคนอาจจะขาดข้อมูลที่บ่งบอกถึงตัวแปรที่สังเกตได้ หรือเลือกที่จะไม่ถูกกำหนดกลุ่มย่อยด้วยตัวแปรที่สังเกตได้ และให้ข้อมูลเป็นตัวบ่งบอกหรือตัดสินกลุ่มย่อยเอง ไม่ใช้การตัดสินตามสิ่งที่สังเกตได้ เนื่องจากความแตกต่างที่ไม่สามารถสังเกตได้ (unobserved heterogeneity) ถูกกำหนดมาก่อนล่วงหน้าความแตกต่างที่สังเกตได้ ซึ่งสามารถโต้เถียงได้ว่าเป็นวิธีการ หรือแนวทางที่เหมาะสมมากกว่าในการค้นพบหรือเปิดเผยกลุ่มย่อยโดยอ้างอิงจากโครงสร้างของความสนใจ (construct of interest) ของแต่ละคน ในกรณีของโมเดลแบบผสม (Rost, 1990, 1991) รูปแบบของการตอบสนองข้อสอบถูกใช้เพื่อที่จะอ้างอิงถึงกลุ่มย่อยที่ถูกแฝงไว้และมีโมเดลการวัดเดียวกันตามที่คาดไว้ กลุ่มย่อยแฝงนี้อาจจะมีความสอดคล้องหรือไม่สอดคล้องกับความแตกต่างที่สังเกตได้ ตัวอย่างเช่น อาจจะไม่

มีความสอดคล้องกันอย่างถูกต้องระหว่างกลุ่มย่อยซึ่งมีลักษณะกลุ่มของการวัดเชิงทัศนคติ (attitude) ที่เหมือนกัน และกลุ่มย่อยเพศ (Eid & Rauber, 2000) โดยแท้จริงแล้ว ความสัมพันธ์ระหว่างความแตกต่างที่สังเกตได้ และไม่สามารถสังเกตได้ มีแนวโน้มที่จะเป็นแนวความคิดเกี่ยวกับความแตกต่างในระดับเบื้องต้น (เช่น การศึกษาเกี่ยวกับประชากร) และความแตกต่างในระดับลึก (เช่น โครงสร้างทางจิตวิทยา เจตคติ ซึ่งเป็นคุณลักษณะแฝง) (Harrison, Price, & Bell, 1998 cited in Tay et.al., 2011) กลุ่มย่อยซึ่งมีความแตกต่างกันในโครงสร้างที่ถูกวัด ไม่จำเป็นที่จะต้องแตกต่างกันตามสิ่งที่เห็นได้ชัด

แม้ว่ากระบวนการตรวจสอบความไม่แปรเปลี่ยนของการวัด(measurement equivalence) ใช้เพียงแหล่งที่สังเกตได้ของความแตกต่าง ส่วนโมเดลแบบผสม (MM-IRT) จะใช้เพียงแหล่งที่ไม่สามารถสังเกตได้ของความแตกต่างเท่านั้น แต่โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับสามารถใช้ได้กับทั้งแหล่งที่สังเกตได้ และไม่สามารถสังเกตได้ แหล่งที่สังเกตได้ของความแตกต่างในกรณีนี้ประกอบด้วย หน่วยที่เป็นลำดับลดหลั่น (hierarchical unit) เช่น ทีม องค์กร ประเทศ ในการวิจัยพหุระดับนั้น การแบ่งกลุ่มที่มีนัยสำคัญทางทฤษฎี (Klein, Dansereau, & Hall, 1994 cited in Tay et.al., 2011) โดยทั่ว ๆ ไปแล้วจะสัมพันธ์กับความผันแปรระหว่างกลุ่ม และความไม่เป็นอิสระที่อยู่ในกลุ่ม (nested dependencies) (S. W. Raudenbush & Bryk, 2002) ความแตกต่างที่ไม่สามารถสังเกตได้ ถูกทำให้มีกรอบแนวคิดในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ที่ 2 ระดับ ระดับบุคคล และระดับลดหลั่น

ระดับบุคคล กลุ่มประชากรย่อยแฝง ไม่ถูกกำหนดโดยกลุ่มลดหลั่นอย่างสมบูรณ์ ตัวอย่างเช่น การพิจารณาว่าระดับลดหลั่นขององค์กร (ทีม แผนก) เป็นวิธีที่สะดวกที่สุดของการแบ่งกลุ่มคนซึ่งอาจจะมีการรับรู้เหมือนกัน เครือข่ายการทำงานซึ่งไม่เป็นทางการ และมีการติดต่อกันภายในองค์กรสามารถที่จะมีการรับรู้เหมือนกัน (Newman, Hanges, Duan, & Ramesh, 2008) ดังนั้น โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ใช้ข้อมูลจากการแบ่งกลุ่มโดยการใช้การสังเกตได้แบบลดหลั่น แต่ขณะเดียวกันกลุ่มย่อยของ แต่ละคนมีโมเดลการวัดเดียวกัน

ระดับลดหลั่น ความแตกต่างที่ไม่สามารถสังเกตได้นี้ จะถูกสมมติ ถูกแทนด้วยกลุ่มแฝงที่เป็นระดับลดหลั่นที่ต่างกัน และถูกพิจารณาที่จะส่งผลต่อสัดส่วนของกลุ่มย่อยระดับบุคคล ที่อยู่ในแต่ละระดับลดหลั่น หรือกล่าวอีกนัยหนึ่งคือ มีกลุ่มที่สังเกตไม่ได้ในหน่วยที่เป็นระดับลดหลั่น ซึ่งมีคุณสมบัติเหมือนกันภายใน แต่แตกต่างกันภายนอกกับกลุ่มอื่นในระดับเดียวกัน

ที่จุดเชื่อมต่อนี้เองแสดงให้เห็นว่า ถึงแม้ว่าการค้นพบความแตกต่างที่ไม่สามารถสังเกตได้เป็นวิธีที่มีความจำเป็นกับกระบวนการ bottom up ซึ่งการแบ่งกลุ่มแฝงถูกสรุป กระบวนการนี้มีพื้นฐานบนโมเดลแนวคิดขั้นต้นของปรากฏการณ์ทางจิตวิทยาองค์กรที่ถูกกำหนดไว้ก่อนล่วงหน้าโดยนักวิจัย

แต่ละโมเดลกรอบแนวคิดคือ ‘‘workable approximation of reality’’ (Wedel, 2002 cited in Tay et.al., 2011) ยังคงทำให้สมบูรณ์และเปรียบเทียบกับกรอบแนวคิดอื่น ๆ ตัวอย่างเช่น นักวิจัยอาจจะเสนอว่าสองกลุ่มประชากรย่อยในระดับบุคคลมีอยู่จริง แต่จะต้องเปรียบเทียบข้อเสนอนี้กับทางเลือกอื่น เช่น โมเดล 3 หรือ 4 ระดับกลุ่มประชากรที่แตกต่างกัน เป็นการเปรียบเทียบกับการวิเคราะห์องค์ประกอบแบบสำรวจซึ่งอาจจะเปรียบเทียบข้ามจำนวนองค์ประกอบที่แตกต่างกัน

โดยสรุปแล้วโมเดลที่เฉพาะเจาะจงซึ่งมีมโนทัศน์สำหรับโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับที่แตกต่างกันแบ่งออกเป็น 2 ชนิด แสดงดังตารางที่ 2.4 การแทนที่ของกลุ่มตัวแปรแฝงระดับลดหลั่น ซึ่งเป็นมิติแนวคิดของระดับลดหลั่น โดยหน่วยลดหลั่นนี้สามารถถูกจัดเป็นมิติต่อเนื่อง จากตารางแสดงการเปรียบเทียบโมเดลตามโครงสร้างที่แตกต่างกันนี้เป็นการเปรียบเทียบโมเดลตามลักษณะข้อมูลที่แตกต่างกัน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 2.4 โมเดลโครงสร้างที่แตกต่างกันภายในกรอบแนวคิดตัวแปรแฝงทั่วไป (Tay, Diener, et al., 2011)

<table>
<tr><th>ระดับการวิเคราะห์</th><th colspan="5">ชนิดข้อมูล</th></tr>
<tr><td rowspan="5">ระดับลดหลั่น
-ระดับบุคคล</td><td rowspan="3"></td><td colspan="4">ข้อมูลที่ไม่ซ้อนกัน (Non-nested data)</td></tr>
<tr><td colspan="2">ไม่มีโมเดลการวัด
(ไม่เป็นข้อมูลในมิติต่อเนื่อง)</td><td colspan="2">มีโมเดลการวัด
(เป็นข้อมูลในมิติต่อเนื่อง)</td></tr>
<tr><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td></tr>
<tr><td rowspan="2">ไม่ต่อเนื่อง
ต่อเนื่อง</td><td>-</td><td>Latent class analysis (LCA)</td><td>Item Response Theory (IRT)</td><td>Mixed-measurement (MM) IRT</td></tr>
<tr><td>-</td><td>Latent Profile Analysis (LPA)</td><td>Factor Analysis (FA)</td><td>Mixed-measurement (MM) FA</td></tr>
<tr><td rowspan="5">ระดับลดหลั่น
-ระดับบุคคล</td><td></td><td colspan="4">ข้อมูลที่ซ้อนกัน (Nested data)</td></tr>
<tr><td></td><td colspan="2">ไม่มีโมเดลการวัด
(ไม่เป็นข้อมูลในมิติต่อเนื่อง)</td><td colspan="2">มีโมเดลการวัด
(เป็นข้อมูลในมิติต่อเนื่อง)</td></tr>
<tr><td></td><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td></tr>
<tr><td rowspan="2">ไม่ต่อเนื่อง
ต่อเนื่อง</td><td>-</td><td>Multilevel LCA</td><td>Multilevel IRT</td><td>Multilevel MM-IRT</td></tr>
<tr><td>-</td><td>Multilevel LPA</td><td>Multilevel FA</td><td>Multilevel MM-FA</td></tr>
<tr><td rowspan="5">ระดับลดหลั่น
-ระดับบุคคล</td><td></td><td colspan="4">กลุ่มแฝงแบบพหุ (Multiple Latent Classes)</td></tr>
<tr><td></td><td colspan="2">ไม่มีโมเดลการวัด
(ไม่เป็นข้อมูลในมิติต่อเนื่อง)</td><td colspan="2">มีโมเดลการวัด
(เป็นข้อมูลในมิติต่อเนื่อง)</td></tr>
<tr><td></td><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td><td>กลุ่มแฝงเดียว</td><td>กลุ่มแฝงแบบพหุ</td></tr>
<tr><td rowspan="2">ไม่ต่อเนื่อง
ต่อเนื่อง</td><td></td><td>Multilevel Mixture LCA</td><td>Multilevel (mixture) IRT</td><td>Multilevel (mixture) MM-IRT</td></tr>
<tr><td></td><td>Multilevel Mixture LPA</td><td>Multilevel (mixture) FA</td><td>Multilevel (mixture) MM-FA</td></tr>
<tr><td colspan="6">หมายเหตุ Multilevel (mixture) MM-IRT ถูกเสนอโดย Tay et. al. (2011) เป็นเทคนิค ซึ่งสามารถเรียกว่า Multilevel (mixture) multidimensional MM-IRT แต่เพื่อให้ง่าย ในที่นี้ใช้ชื่อว่า Multilevel MM-IRT(MMM-IRT) Model</td></tr>
</table>

จากที่กล่าวข้างต้น จะเห็นได้ว่า โมเดลการตอบสนองข้อสอบ สามารถจัดการกรอบแนวคิดที่มีประโยชน์สำหรับวิเคราะห์การตอบสนองข้อสอบ โดยยืนยันความสัมพันธ์ระหว่างการตอบสนองข้อสอบแต่ละข้อ ส่วนการวิเคราะห์กลุ่มแฝง เป็นเทคนิคการวิเคราะห์ตามลักษณะการวิเคราะห์องค์ประกอบที่ได้รับการพัฒนาให้ สามารถวิเคราะห์ตัวแปรทวิวิภาคได้ วิเคราะห์ตัวแปรจัดประเภทได้ทุกชนิด รวมทั้งสามารถวิเคราะห์ตัวแปรที่มีระดับการวัดต่างกันได้ ซึ่งโมเดลกลุ่มแฝง อธิบายความสัมพันธ์ระหว่างการตอบสนองของค่าสังเกตได้และข้อสอบ ประกอบไปด้วย ตัวแปรที่มีลักษณะเป็นกลุ่มซึ่งแทนกลุ่มที่ถูกระบุ และจำนวนของตัวแปรอื่น ๆ ซึ่งแทนคุณลักษณะของสิ่งนั้น ส่วนโมเดลพหุระดับวิเคราะห์โครงสร้างระดับที่เป็นธรรมชาติของข้อมูลทางการศึกษาและจิตวิทยา การรวมกัน

ของโมเดลพหุระดับกับโมเดลการตอบสนองข้อสอบ เป็นประโยชน์มากสามารถจัดการประมาณค่าที่ถูกต้องของความคลาดเคลื่อนมาตรฐานของโมเดล แต่ข้อจำกัดของการตอบสนองข้อสอบแบบพหุระดับ จากมุมมองของการตอบสนองข้อสอบแบบผสม คือการตอบสนองข้อสอบแบบพหุระดับละเลยการจัดการสารสนเทศเกี่ยวกับสมาชิกในกลุ่มแฝงซึ่งมีความสำคัญรวมเข้ากับโมเดล ซึ่งต่อมาได้มีการบูรณาการโครงสร้างโมเดลพหุระดับเข้ากับ โมเดลการตอบสนองข้อสอบแบบผสมโดยเสนอเป็นโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ และถูกใช้ในการเปรียบเทียบกลุ่มแฝง และจำแนกกลุ่มความสามารถแฝงของข้อมูลซึ่งมีคุณลักษณะทางการวัดที่แตกต่างกัน ประโยชน์ของแนวคิดนี้แสดงถึงคุณลักษณะที่แท้จริงของกลุ่มแฝง แม้ว่าไม่สามารถสังเกตได้ทันที ซึ่งแนวคิดนี้ถูกนิยามจากความโน้มเอียงเกี่ยวกับการตอบสนองที่มีลักษณะร่วมกัน ซึ่งสามารถใช้อธิบายสาเหตุที่เป็นไปได้ของคุณลักษณะทางการวัดที่แตกต่างกัน แนวคิดนี้ถูกเสนอในการศึกษาการจัดการกับสารสนเทศของพหุระดับที่อธิบายองค์ประกอบของกลุ่มแฝงที่แตกต่างกัน

1.3 การประมาณค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) มีจุดมุ่งหมายหลักเพื่อวิเคราะห์หาสารสนเทศของการทดสอบ ซึ่งได้แก่ สารสนเทศของข้อสอบ และสารสนเทศของผู้สอบ การหาสารสนเทศดังกล่าวจะต้องมีการประมาณค่าพารามิเตอร์ผู้สอบซึ่งได้แก่ระดับความสามารถของผู้สอบและประมาณค่าพารามิเตอร์ของข้อสอบซึ่งได้ ค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสในการเดาคำตอบของข้อสอบถูก ดังนั้นการประมาณค่าพารามิเตอร์สำหรับการวิเคราะห์ข้อสอบจึงมีความสำคัญและมีความจำเป็น สำหรับการวิเคราะห์ข้อสอบ

การประมาณค่าพารามิเตอร์ในโมเดลตัวแปรแฝง มีวัตถุประสงค์เพื่อที่จะประมาณค่าพารามิเตอร์สำหรับตัวแปรที่ไม่สามารถสังเกตได้ และระบุค่าทั่วไปของโมเดลทฤษฎีการตอบสนองข้อสอบ เมื่อการแจกแจงความสามารถถูกสมมติให้มีการแจกแจงแบบปกติ นอกจากนี้สำหรับ โมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT) ก็จะมีการระบุค่าสเกลและพารามิเตอร์สำหรับประมาณค่า เช่น การประมาณค่าพารามิเตอร์โดยวิธีเบส์เซียน

การประมาณค่าพารามิเตอร์ เป็นการใช้ข้อเท็จจริงจากตัวอย่างสุ่มในรูปของสถิติในการประมาณว่าค่า θ มีค่าเท่าใด หรืออยู่ในช่วงใด ซึ่งการประมาณค่าแบ่งออกเป็น 2 ชนิด คือการประมาณค่าแบบจุดและการประมาณค่าแบบช่วง ซึ่งในการเลือกใช้สถิติเพื่อประมาณค่าพารามิเตอร์ θ นั้น จะพิจารณาจากการแจกแจงของกลุ่มตัวอย่างเป็นหลักในการพิจารณา โดยใช้สถิติหลายอย่างประกอบกันในการเลือก ตัวประมาณที่ดี นั่นคือความไม่เอนเอียง (unbiasedness) ความคงเส้นคงวา (consistency) ความพอเพียง (sufficiency) ความมีความแปรปรวนต่ำสุด (minimum variance)

ความมีประสิทธิภาพ (efficiency) และความมีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด (minimum mean square error) เป็นต้น

จากการพิจารณาถึงคุณสมบัติทางสถิติที่กล่าวมา มีวิธีหาตัวประมาณค่า θ อยู่หลายวิธี ซึ่งต้องหาค่า $\hat{\theta} = \hat{\theta}(X_1, X_2, ..., X_n)$ ซึ่งใช้เป็นตัวประมาณค่าพารามิเตอร์ θ หรือฟังก์ชันของพารามิเตอร์ θ เช่น $\tau_1(\theta), \tau_2(\theta), ..., \tau_r(\theta)$ ในทฤษฎีทางสถิติมีวิธีการประมาณค่าพารามิเตอร์หลายวิธีเช่น วิธีโมเมนต์ (Method of moments) วิธีภาวะน่าจะเป็นสูงสุด (Method of Maximum Likelihood) วิธีไคว์สแควร์ต่ำสุด (Method of Minimum Chi-square) วิธีกำลังสองต่ำสุด (Method of Least Square) วิธีระยะทางต่ำสุด (Minimum Distance Method) วิธีของเบส์ (Method of Bayes) วิธีมินิแม็กซ์ (Minimax Method) และวิธีของพิทแมน (Method of Pitman) เป็นต้น โดยวิธีโมเมนต์และวิธีภาวะน่าจะเป็นสูงสุดให้ตัวประมาณค่าพารามิเตอร์ (p) หรือค่าสัดส่วนทวินามเหมือนกันคือ $\hat{P} = \frac{X}{n}$ ซึ่งเป็นตัวประมาณที่มีคุณสมบัติไม่เอนเอียง และมีความแปรปรวนต่ำสุดอย่างเอกรูป (Uniformly Minimum Variance Unbiased Estimator : UMVUE)

สำหรับการประมาณค่าพารามิเตอร์ θ ในฟังชันก์การตอบสนองข้อสอบนั้นได้มีการนำวิธีการประมาณค่าพารามิเตอร์โดยวิธีเบส์ (Baysian estimation) มาใช้ โดยวิธีของเบส์จะมีแนวคิดแตกต่างจากวิธีดังที่ได้กล่าวมา คือ โดยทั่วไปจะถือว่าพารามิเตอร์ θ เป็นค่าคงที่ซึ่งไม่ทราบค่า แต่ตามวิธีของเบส์จะถือว่าค่า θ เป็นค่าของตัวแปรสุ่ม θ ที่มีการแจกแจงแสดงได้ในรูปการแจกแจงความน่าจะเป็น ซึ่งเรียกว่า การแจกแจงก่อน (prior distribution) เนื่องจากเป็นการแจกแจงที่กำหนดขึ้นก่อนที่จะมีการเก็บรวบรวมข้อมูล เมื่อมีการรวบรวมข้อมูลมาแล้วจะใช้ความรู้ที่ได้จากข้อมูลมาปรับปรุงการแจกแจงก่อน การแจกแจงที่ได้จากการปรับปรุงในภายหลังนี้ เรียกว่า การแจกแจงภายหลัง (posterior distribution) (Congdon, 2005)

ส่วนวิธีประมาณค่าพารามิเตอร์ของวิธีการวิเคราะห์พหุระดับ มีหลายวิธีดังนี้ (ศิริชัย กาญจนวาสี, 2550) การประมาณค่าส่วนประกอบความแปรปรวน (analysis of variance component estimation) วิธีกำลังสองน้อยที่สุด (OLS single equation approach) วิธีกำลังสองน้อยที่สุดแบบแบ่งสองสมการ (OLS separated equation approach) วิธีการประมาณค่าเป็นไปได้สูงสุด (maximum likelihood) การประมาณค่าพารามิเตอร์ด้วยวิธีการของเบส์ (Bayesian estimation) โดยในตอนนี้จะเป็นการนำเสนอวิธีการประมาณค่าพารามิเตอร์ซึ่งเป็นวิธีที่ใช้ในงานวิจัยนี้ เป็นวิธีการประมาณค่าที่เหมาะสมกับข้อมูลที่เป็นพหุระดับ ได้แก่ วิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood) และวิธีของเบส์ (Bayesian) (Hambleton & Cook, 1977; Lord, 2009; รัตนา ศรีเหรัญ, 2539; ศิริชัย กาญจนวาสี, 2555) แต่ละวิธีมีรายละเอียด ดังนี้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1.3.1 วิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood: ML)

การประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด เป็นการประมาณค่าพารามิเตอร์โดยอาศัยผลที่ได้จากตัวอย่างที่สุ่มเลือกมาจากการแจกแจงที่ทราบรูปแบบของ ฟังก์ชันความหนาแน่นแต่ไม่ทราบค่าพารามิเตอร์ ดังนั้นจึงใช้หลักของความน่าจะเป็นในการเลือกตัวอย่างและวัดค่าได้จากกลุ่มตัวอย่างที่ถูกเลือกมา พิจารณาค่าประมาณของค่าพารามิเตอร์ที่ต้องการ ซึ่งโดยทั่วไปสมการไลค์ลิฮูดไม่เป็นสมการเส้นตรง ดังนั้น การหาค่าความสามารถ ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา ที่ทำให้ฟังก์ชั่นไลค์ลิฮูดมีค่าสูงสุด จึงไม่สามารถหาได้ด้วยการใช้วิธีการอย่างง่าย แต่หาได้โดยใช้วิธีของ วิธีของ นิวตัน ราฟสัน (Newton-Raphson Procedure) (Lord, 2009) ซึ่งเป็นการหาค่าประมาณโดยการ ประมาณค่าซ้ำ (Iterative) จนค่าที่ได้มีค่าคงที่

การประมาณค่าด้วยวิธีการประมาณแบบไลค์ลิฮูดมี 2 วิธี คือ วิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุดส่วนริม (Marginal maximum likelihood) และวิธีการประมาณค่าความน่าจะเป็นสูงสุดร่วมกัน (Joint Maximum Likelihood Estimation Procedure) ในที่นี้จะแสดงเพียงข้อสรุปของทั้งสองวิธี เนื่องจากข้อมูลที่ใช้ในงานวิจัยนี้เป็นข้อมูลพหุระดับ จึงจะแสดงรายละเอียดเฉพาะวิธีการประมาณค่าความน่าจะเป็นสูงสุดในข้อมูลพหุระดับในหัวข้อถัดไป

สำหรับวิธีการประมาณค่าความน่าจะเป็นสูงสุดส่วนริม (Marginal maximum likelihood) เป็นวิธีการประมาณค่า โดยสามารถประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบได้พร้อมกัน ซึ่งการประมาณค่าด้วยวิธีการนี้จะไม่ได้อิงกับค่าพารามิเตอร์ของผู้สอบ นอกจากนี้ยังมีการแจกแจงความสามารถของผู้สอบ เพื่อให้ได้ฟังก์ชันความน่าจะเป็นของพารามิเตอร์ของข้อสอบ หลังจากนั้นจึงนำค่าที่ได้ไปใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบ การประมาณค่าด้วยวิธีนี้สามารถใช้ได้ กับแบบทดสอบที่มีทั้งจำนวนข้อมากและน้อยได้ นอกจากนี้ การประมาณค่าด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุดส่วนริม สามารถใช้โปรแกรมคำสั่งชื่อ BILOG ซึ่งพัฒนาโดย Mislevy และ Bock (1983) แต่ในการใช้โปรแกรมดังกล่าวต้องใช้กลุ่มผู้สอบที่มีขนาดใหญ่ เพื่อทำให้สามารถประมาณค่าพารามิเตอร์ของข้อสอบได้อย่างคงที่ นอกจากนี้จำนวนข้อสอบสำหรับการประมาณค่ายังต้องมีจำนวนมาก สำหรับการประมาณค่าพารามิเตอร์ของผู้สอบโดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3 PL) ต้องมีผู้สอบที่มีระดับความสามารถต่ำที่เพียงพอด้วย จึงจะสามารถทำการประมาณค่าพารามิเตอร์ค่าโอกาสในการเดาข้อสอบถูก (c) ได้อย่างมีความน่าเชื่อถือ นอกจากมีผลกระทบต่อค่าพารามิเตอร์โอกาสในการเดาข้อสอบถูก (c) แล้วยังอาจจะส่งผลกระทบถึงค่าพารามิเตอร์ความยากของข้อสอบ (b) และค่าพารามิเตอร์อำนาจจำแนกของ ข้อสอบ (a) อีกด้วย

ส่วนวิธีการประมาณค่าความน่าจะเป็นสูงสุดร่วมกัน (Joint Maximum Likelihood Estimation Procedure) มีการกำหนดค่าเริ่มต้นของความสามารถ (θ) โดยใช้ค่า log ของอัตราส่วนจำนวนข้อที่ตอบถูกต่อจำนวนข้อที่ตอบผิดสำหรับผู้ตอบแต่ละคนแปลงเป็นคะแนนมาตรฐานเพื่อใช้เป็นค่าเริ่มต้นของค่าความสามารถ (θ) จากนั้นทำประมาณค่าพารามิเตอร์ของข้อสอบจากค่าเริ่มต้นที่กำหนดขึ้นของค่าความสามารถ จากขั้นตอนแรก ในขั้นตอนที่สอง จากค่าพารามิเตอร์ของข้อสอบที่ได้จากขั้นตอนแรก ทำเสมือนทราบค่าพารามิเตอร์ของข้อสอบ เพื่อประมาณค่าความสามารถของผู้สอบ วิธีนี้จะกระทำซ้ำตามขั้นตอนทั้งสองจนกระทั่งได้ค่าประมาณ 2 ครั้งหลังที่ไม่เปลี่ยนแปลงโปรแกรมคอมพิวเตอร์สำหรับการประมาณค่าด้วยวิธีการนี้ ได้แก่ โปรแกรม LOGIST สำหรับการวิเคราะห์ข้อสอบตามโมเดลการตอบสนอง 1, 2 และ 3 พารามิเตอร์และโปรแกรม BICAL และโปรแกรม BIGSCALE สำหรับการวิเคราะห์ข้อสอบตามโมเดลการตอบสนอง 1 พารามิเตอร์ (ศิริชัย กาญจนวาสี, 2555) ซึ่งข้อพึงระวังสำหรับการประมาณค่าด้วยวิธีประมาณค่าความน่าจะเป็นสูงสุดร่วมกัน คือ การประมาณค่าด้วยวิธีการนี้ไม่สามารถประมาณค่าความสามารถได้ในกรณีมีผู้สอบได้คะแนนเต็มหรือได้ศูนย์คะแนน นั่นคือยังไม่สามารถประมาณค่าพารามิเตอร์ข้อสอบได้ ในกรณีข้อสอบที่มีผู้ตอบถูกหรือผิดหมดทุกคน และในการวิเคราะห์โมเดล 2 และ 3 พารามิเตอร์ค่าพารามิเตอร์ของโมเดลการตอบสนองข้อสอบจะมีความคงเส้นคงวาได้มากขึ้น ก็ต่อเมื่อมีจำนวนข้อสอบมากและกลุ่มผู้สอบมีขนาดใหญ่ (Hambleton, Swaminathan และ Rogers, 1991 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555)

เมื่อเปรียบเทียบวิธีการประมาณค่าความน่าจะเป็นสูงสุดทั้ง 2 วิธี พบว่าวิธีการประมาณค่าที่เป็นไปได้สูงสุดแบบวิธีการประมาณค่าความน่าจะเป็นสูงสุดส่วนริม มีข้อดีกว่าวิธีการประมาณค่าที่เป็นไปได้สูงสุดแบบร่วมกันที่สำคัญ คือ สามารถประมาณค่าผู้สอบโดยไม่ต้องประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ รวมทั้งความแม่นยำของการวิเคราะห์ เมื่อมีจำนวนผู้สอบ 1,000 ถึง 2,000 คน ทำข้อสอบ 40 ข้อ จะให้ค่าประมาณที่แตกต่างกันเล็กน้อย แต่ในกรณีที่ผู้สอบทำข้อสอบ 10 ถึง 15 ข้อ วิธีการประมาณค่าที่เป็นไปได้สูงสุดแบบร่วมกัน จะให้ค่าประมาณความสามารถที่ลำเอียงโดยเฉพาะในผู้สอบที่มีความสามารถต่ำ ดังนั้นจึงทำให้การประมาณค่าพารามิเตอร์มีความคลาดเคลื่อนแม้จะมีผู้สอบจำนวนมาก (Hambleton และ Swaminathan, 1985)

การประมาณค่าด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุดมีข้อจำกัด คือ ในการประมาณค่าพารามิเตอร์ในขั้นที่ 2 และ 3 โดยใช้ค่าอนุพันธ์อันดับที่ 2 ในกระบวนการนิวตัน ราฟสันนั้น มีโอกาสที่ค่าประมาณที่ได้จะไม่ลู่เข้าสู่ค่าคงที่ นอกจากนี้สำหรับการประมาณค่าในสมการไลค์ลิฮูดไม่ใช่สมการเชิงเส้นตรง จะทำให้การหารากของสมการที่ทำให้ฟังก์ชันไลค์ลิฮูดมีค่าสูงสุดได้หลายค่า แต่ค่าเหล่านี้ไม่สามารถนำไปใช้หรือประกันได้ว่าเป็นค่าพารามิเตอร์ที่แท้จริงได้ และในบางครั้ง

ค่าพารามิเตอร์หรือค่าที่ได้จากการประมาณไม่ตกอยู่ในขอบเขตของค่าพารามิเตอร์ อาจมีค่าใดค่าหนึ่งอยู่ภายนอกขอบเขตที่ยอมรับได้ ในกรณีเช่นนี้ต้องมีการกำหนดขอบเขตจำกัดของค่าประมาณไว้เพื่อให้ค่าประมาณที่ได้ไม่สูงหรือต่ำเกินไปนัก ซึ่งเป็นจุดอ่อนของกระบวนการประมาณค่าด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด โดยเฉพาะ 2PL และ 3PL จะมีปัญหาเกี่ยวกับความตรงของค่าที่ประมาณได้ และในการประมาณค่าพารามิเตอร์ใน 3PL มีสมการหลายสมการที่ต้องหารากที่ทำให้ฟังก์ชันไลค์ลิฮูดมีค่าสูงสุด จึงจำเป็นต้องใช้เครื่องคอมพิวเตอร์ขนาดใหญ่ และมีความจุมากในการทำงาน

ส่วนข้อดีของการประมาณค่าด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด คือ การประมาณค่าพารามิเตอร์ด้วยวิธีการนี้จะให้สารสนเทศของค่าพารามิเตอร์ที่ต้องการได้ทั้งหมด ไม่ว่าจะเป็นพารามิเตอร์ของผู้สอบ ซึ่งได้แก่ ค่าความสามารถที่แท้จริงของผู้สอบ หรือพารามิเตอร์ของข้อสอบ ซึ่งได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบและค่าโอกาสในการเดาข้อสอบได้ถูกต้องการประมาณค่าพารามิเตอร์ของวิธีนี้ขึ้นอยู่กับจำนวนของกลุ่มผู้เข้าสอบและข้อสอบถ้ามีจำนวนเพิ่มขึ้น การประมาณค่าก็จะมีความคงที่ไปสู่ค่าพารามิเตอร์เพิ่มมากขึ้น

การประมาณค่าด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุดกับข้อมูลพหุระดับ

พารามิเตอร์ของ โมเดลตัวแปรแฝงพหุระดับ (Multilevel Latent Variable Model) สามารถประมาณได้โดย Maximum Likelihood (ML) บนพื้นฐานของสมการการถดถอย ข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงความคลาดเคลื่อนสำหรับตัวแปรตอบสนอง และตัวแปรแฝง และโครงสร้างเชิงลำดับขั้นของโมเดลสามารถหาอนุพันธ์ได้ ฟังก์ชั่นความหนาแน่นที่เกี่ยวข้องกับเวกเตอร์ตอบสนองของการสังเกตที่อิสระ การสังเกตที่ระดับสูงของลำดับขั้น สำหรับคำจำกัดความที่ง่าย ให้คิดว่ามีโมเดลตัวแปรแฝงต่อเนื่องที่ สองระดับ ซึ่งปกติจะเป็นโมเดลสำหรับ ข้อมูล 3 ระดับที่เป็น univariate หรือ ข้อมูล 2 ระดับที่เป็น multivariate ในที่นี้จะใช้คำศัพท์ที่เกี่ยวข้องกับ โมเดล 3 ระดับ ซึ่งระดับของลำดับขั้นใช้ i , j , k

ฟังก์ชัน Log-likelihood

ฟังก์ชันไลค์ลิฮูด เป็นพื้นฐานต่อ ความหนาแน่นที่น่าจะเป็น (probability density) ของการสังเกตใน 3 ระดับ แสดงโดย $f\left(y_k|Z_k,W_k\right)$ ในที่นี้ y_k, Z_k และ W_k มีการตอบสนองและการออกแบบเวกเตอร์สำหรับทุกข้อสังเกตในระดับต่ำกว่าที่เป็นในระดับที่ 3 หรือกลุ่มที่ k เพื่อลดความซับซ้อนของสัญลักษณ์เงื่อนไขเวกเตอร์การออกแบบจะถูกแทนด้วยดัชนีที่สอดคล้องกันกับหน่วยที่

เกี่ยวข้อง โดยจะแสดงสั้น ๆ ด้วยสัญลักษณ์ $f_k(y_k)$ สำหรับความหนาแน่นที่น่าจะเป็นของระดับ 3 หน่วยที่ k

ค่า log likelihood จะเท่ากับ

$$\log L = \sum_{k-1}^{K} \log f_k(y_k) \quad \text{..... (2.20)}$$

โดยที่

$$f_k(y_k) = \int_{\nu^{(3)}} f_k\left(y_k \middle| \nu^{(3)}\right) f\left(\nu^{(3)}\right) d\nu^{(3)}$$

$$= \int_{\nu^{(3)}} \left\{ \prod_{j=1}^{n_k} f_{jk}\left(y_{jk} \middle| \nu^{(3)}\right) \right\} f\left(\nu^{(3)}\right) d\nu^{(3)} \quad \text{..... (2.21)}$$

และ

$$f_{ik}\left(y_{jk} \middle| \nu^{(3)}\right) = \int_{\nu^{(2)}} f_{jk}\left(y_{jk} \middle| \nu^{(2)}, \nu^{(3)}\right) f\left(\nu^{(2)}\right) d\nu^{(2)}$$

$$= \int_{\nu^{(2)}} \left\{ \prod_{i=1}^{n_{jk}} f_{ijk}\left(y_{ijk} \middle| \nu^{(2)}, \nu^{(3)}\right) \right\} f\left(\nu^{(2)}\right) d\nu^{(2)} \quad \text{..... (2.22)}$$

และการตอบสนองของ n_k ในระดับที่ 2 ภายในระดับ 3 หน่วยที่ k ซึ่งสมมติว่าเป็นอิสระของอีกตัวแปรหนึ่งที่ได้รับผลกระทบแบบสุ่ม $\nu^{(3)}$ และการตอบสนองของ n_{jk} ระดับที่ 1ภายในระดับ 2 หน่วยที่ jk ถูกสมมติว่าเป็นอิสระของอีกตัวแปรแฝงที่ได้รับผลกระทบแบบสุ่ม $\nu^{(2)}$ และ $\nu^{(3)}$ การอินทิเกรตทางขวาของสมการข้างต้น สามารถประเมินโดยใช้วิธีการ Gauss – Hermite quadrature numerical (Bock & Aitkin, 1981; Skrondal & Rabe-Hesketh, 2004; Stroud & Secrest, 1966) ซึ่งการแจกแจงแบบผสมปกติหลายตัวแปร ถูกประมาณโดยจำนวนที่จำกัดของ discrete point ความแม่นยำที่มากขึ้นคือการอินทิเกรตถูกแทนที่โดย ผลบวกบน $T^{(3)}$ และ $T^{(2)}$ (summation over $T^{(3)}$ and $T^{(2)}$ quadrature point)

$$f_k(y_k) = \sum_{s=1}^{T^{(3)}} P_k\left(y_k \middle| v_s^{(3)}\right) \pi\left(v_s^{(3)}\right) = \sum_{s=1}^{T^{(3)}} \left[\prod_{j=1}^{n_k} P_{jk}\left(y_{jk} \middle| v_s^{(3)}\right) \right] \pi\left(v_s^{(3)}\right)$$

$$= \sum_{s=1}^{T^{(3)}} \prod_{j=1}^{n_k} \sum_{s=1}^{T^{(2)}} \left[\left\{ \prod_{i=1}^{n_{jk}} P_{ijk}\left(y_{ijk} \middle| v_r^{(2)}, v_s^{(3)}\right) \right\} \pi\left(v_s^{(3)}\right) \right] \pi\left(v_s^{(3)}\right) \quad \text{..... (2.23)}$$

ในที่นี้ $v_r^{(2)}$ และ $v_s^{(3)}$ เป็น quadrature nodes $\pi\left(v_r^{(2)}\right)$ และ $\pi\left(v_s^{(3)}\right)$ เป็น quadrature weights ซึ่งสัมพันธ์กับความหนาแน่นปกติ (multivariate) ของความสนใจ เพราะว่าตัวแปรแฝงหรือผลกระทบอย่างสุ่มถูก orthogonalize สำหรับ nodes และ weights ของมิติที่แยกกันเท่ากับส่วนหนึ่งของ ความหนาแน่นปกติ (univariate) ซึ่งสามารถได้จากตารางปกติ (e.g., Stroud and crest,1966) สมมติว่าแต่ละ มิติถูกประมาณกับ Q quadrature nodes น้ำหนัก $T^{(2)} = Q^{R^{(2)}}$ และ $T^{(3)} = Q^{R^{(3)}}$ จะได้จากการคูณน้ำหนักของมิติที่แยกกัน การอินทิเกรต สามารถประมาณดีกรีของความแม่นยำได้อย่างแท้จริง โดยการตั้งค่า Q ให้ใหญ่เพียงพอ

การประมาณค่าแบบ Maximum Likelihood โดยประยุกต์ EM algorithm

โดยปกติการแก้ปัญหาการประมาณค่าความน่าจะเป็นสูงสุด สำหรับโมเดลตัวแปรแฝงพหุระดับ (multilevel latent variable model :MLVM) โดยขั้นตอนวิธีของ EM (Dempster et al., 1977) โดย E step ของ ขั้นตอนวิธี EM จะพัฒนาการคำนวณการคาดการณ์ (expectation) ของข้อมูลที่สมบูรณ์ log likelihood ซึ่งในโมเดล MLVM จะอยู่ในรูปแบบต่อไปนี้

$$\log L_c = \sum_{s=1}^{T^{(3)}}\sum_{r=1}^{T^{(2)}}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\sum_{i=1}^{n_{jk}} P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)\log f_{ijk}\left(y_{ijk}|, v_r^{(s)}, v_s^{(3)}\right) \quad \text{..... (2.24)}$$

ซึ่งแสดงให้เห็นว่าในความเป็นจริง การพัฒนา E step ได้รับการแจกแจงความน่าจะเป็นภายหลัง $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ ให้การประมาณปัจจุบันสำหรับ Model unknown parameter ใน M Step ของขั้นตอน EM Model unknown parameter ถูกทำให้ทันสมัยดังนั้น ข้อมูลที่สมบูรณ์ที่คาดไว้เป็น Maximized (หรือ improved) นี้สามารถทำให้สำเร็จโดยการใช้ขั้นตอนข้อมูลที่สมบูรณ์มาตรฐานสำหรับการประมาณ ML ของ generalized linear model

ปัญหาในส่วนการดำเนินการของ EM สำหรับ MLVM คือ E step ต้องได้รับ ความน่าจะเป็นภายหลัง (posterior probabilities) $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ การดำเนินการมาตรฐานของ E step ควรพัฒนาการคำนวณความคาดหมายที่เป็นเงื่อนไขร่วมกัน (joint conditional expectation) ของผลกระทบอย่างสุ่ม $n_k \cdot R^{(2)} + R^{(3)}$ สำหรับ level-3 หน่วยที่ k ซึ่งการแจกแจงภายหลังร่วมกัน $P_k\left(v_{r_1}^{(2)}, v_{r_2}^{(2)}, ..., v_{r_{n_k}}^{(2)}, v_s^{(3)} \middle| y_k\right)$ กับ $Q^{\left\{n_k \cdot R^{(2)} + R(3)\right\}}$

หมายเหตุ ปริมาณนี้จะคำนวณ การคาดการณ์ของทุกข้อมูลที่ขาดหาย สำหรับ หน่วยระดับที่ 3 ซึ่ง joint posteriors ควรจะถูกยุบเพื่อที่จะได้รับ marginal posterior probabilities ในภายหลังสำหรับแต่ละ level-2 หน่วยที่ j ภายใน level-3 หน่วยที่ k, $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ นี้ทำให้กระบวนการในการจัดเก็บในคอมพิวเตอร์ และเวลาเพิ่มขึ้นทวีคูณ กับจำนวนของหน่วยใน level-2 ซึ่งหมายถึง มันสามารถถูกใช้กับ n_k ที่มีจำนวนน้อย

อย่างไรก็ตามความเป็นไปได้ในการคำนวณการแจกแจงความน่าจะเป็นภายหลังของ n_k $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ โดยไม่มีการผ่านการแจกแจงภายหลังอย่างสมบูรณ์ ซึ่งการใช้ข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระจากกันอย่างมีเงื่อนไข ที่สัมพันธ์กับฟังก์ชั่นความหนาแน่นที่ระบุในสมการที่ (3) ด้วยเหตุนี้ กระบวนการนี้จึงคล้ายกับ forward-backward algorithm สำหรับการประมาณของ hidden Markov models กับจำนวนที่มากของจุดเวลา (Baum, Petrie, Soules, & Weiss, 1970; Frühwirth-Schnatter, 2006; Juang & Rabiner, 1991) ซึ่งกระบวนการนี้ Vermunt (2004) เรียกว่าเป็น upward-downward procedure หรือ โมเดลเครือข่ายของ Bayesian โมเดลตัวแปรแฝงแบบพหุระดับ จะถูกจำแนกเป็น single-connected network หรือ polytree ที่เกี่ยวข้องกับ

ความน่าจะเป็นอย่างมีเงื่อนไขเล็กน้อย (marginal conditional probabilities) ซึ่งสามารถได้รับจาก propagation algorithms Pearl และ Russel (1998) ซึ่งทั้ง forward-backward algorithm สำหรับ hidden Markov models และ upward-downward algorithm จะแสดงภายใต้ propagation algorithms

สำหรับ upward-downward algorithm ตัวแปรแฝงถูก integrate ออกไปจากระดับที่ต่ำกว่า ถึงระดับที่สูงกว่า สำหรับ marginal posterior probabilities ถูกคำนวณจากระดับที่สูงกว่าถึงระดับที่ต่ำกว่า เป็นผลให้กระบวนการในการจัดเก็บข้อมูลในคอมพิวเตอร์ และเวลาที่เพิ่มขึ้นเป็นแบบเชิงเส้น (linearly) ตามจำนวนของตัวแปรสังเกตได้ ในระดับ ที่ 2 แทนที่จะเป็นแบบทวีคูณ (exponentially) ตามที่จะได้รับในกรณีที่เป็นมาตรฐานของ EM algorithm ซึ่งขั้นตอนวิธีนี้ ถูกใช้ใน Latent Gold software package (Vermunt, 2007b; Vermunt & Magidson, 2005)

Marginal posterior probabilities $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ สามารถจำแนกดังนี้

$$P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right) = P_k\left(v_s^{(3)} \middle| y_k\right) P_{jk}\left(v_r^{(2)} \middle| y_k, v_m^{(3)}\right) \quad \text{..... (2.25)}$$

กระบวนการนี้คือ

$$P_{jk}\left(v_r^{(2)} \middle| y_k, v_s^{(3)}\right) = P_{jk}\left(v_r^{(3)} \middle| y_{jk}, v_s^{(3)}\right) \quad \text{..... (2.26)}$$

$v_r^{(2)}$ ไม่ขึ้นกับการตอบสนองที่สังเกตได้ของหน่วยในระดับที่ 2 ภายในระดับที่ 3 $v^{(3)}$ นี้เป็นผลจากความเป็นจริงที่การสังเกตของระดับที่ 2 เป็นอิสระกันร่วมกันให้ผลกระทบอย่างสุ่มในระดับที่ 3 ซึ่งแสดงในฟังก์ชั่นความหนาแน่นในสมการที่ 3 การใช้ผลที่สำคัญสามารถแสดงให้ง่ายขึ้นเล็กน้อยดังนี้

$$P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right) = P_k\left(v_s^{(3)} \middle| y_k\right) P_{jk}\left(v_r^{(2)} \middle| y_{jk}, v_s^{(3)}\right) \quad \text{..... (2.27)}$$

การคำนวณ marginal posterior probabilities จะลดการคำนวณ 2 term ทางด้านขวาของสมการ และ $P_{jk}\left(v_r^{(2)} \middle| y_{jk}, v_s^{(3)}\right)$ จะได้จาก

$$P_{jk}\left(v_r^{(2)} \middle| y_{jk}, v_s^{(3)}\right) = \frac{P_{jk}\left(y_{jk}, v_r^{(2)} \middle| v_s^{(3)}\right)}{P_{jk}\left(y_{jk} \middle| v_s^{(3)}\right)} \quad \text{..... (2.28)}$$

ซึ่ง $P_{jk}\left(y_{jk}, v_r^{(2)} \middle| v_s^{(3)}\right) = \pi\left(v_r^{(2)}\right) \prod_{i=1}^{n_{jk}} P_{ijk}\left(y_{ijk} \middle| v_r^{(2)}, v_s^{(3)}\right)$ .....(2.29)

$$P_{jk}\left(y_{jk} \middle| v_s^{(3)}\right) = \sum_{r=1}^{T^{(2)}} P_{ijk}\left(y_{ijk} \middle| v_r^{(2)}, v_s^{(3)}\right) \quad \text{.....(2.30)}$$

ในส่วนของ $P_k\left(v_s^{(3)} \middle| y_k\right)$ ได้จาก

$$P_k\left(v_s^{(3)} \middle| y_k\right) = \frac{P_k\left(y_k, v_s^{(3)}\right)}{P_k\left(y_k\right)} \quad \text{.....(2.31)}$$

ซึ่ง $$P_k\left(y_k, v_s^{(3)}\right) = \pi\left(v_s^{(3)}\right)\prod_{j=1}^{n_k} P_{jk}\left(y_k \middle| v_s^{(3)}\right)$$ .....(2.32)

$$P_k\left(y_k\right) = \sum_{s=1}^{T^{(3)}} P\left(y_k, v_s^{(3)}\right)$$ .....(2.33)

ดังนั้น อันดับแรก ความน่าจะเป็นภายหลัง ในระดับที่ 2 $P_{jk}\left(v_r^{(2)} \middle| y_{jk}, v_s^{(3)}\right)$ ได้จากข้อมูลในระดับที่ 1 $P_{ijk}\left(y_{ijk} \middle| v_r^{(2)}, v_s^{(3)}\right)$ และต่อมา ความน่าจะเป็นภายหลัง ในระดับที่ 3 $P_k\left(v_s^{(3)} \middle| y_k\right)$ ได้จากข้อมูลในระดับที่ 2

$P_{jk}\left(y_k \middle| v_s^{(3)}\right)$ นี้เรียกว่า upward step ของขั้นตอน เพราะว่า มีการขึ้นไปในโครงสร้างลดหลั่น ใน downward step จะคำนวณโดยใช้ $P_{jk}\left(v_r^{(2)}, v_s^{(3)} \middle| y_k\right)$ ในสมการที่ 7

วิธี upward-downward สามารถ generalized ไปยัง 3 ระดับ ได้อย่างง่าย ตัวอย่างเช่น ในแบบ 4 ระดับ จะคำนวณ 3 term คือ $P_m\left(v_o^{(4)} \middle| y_m\right)$, $P_{km}\left(v_r^{(3)} \middle| y_{km}, v_o^{(4)}\right)$ และ $P_{jkm}\left(v_r^{(2)} \middle| y_{jkm}, v_s^{(3)}, v_o^{(4)}\right)$ ซึ่ง m หมายถึง หน่วยในระดับที่ 4 และ o หมายถึง quadrature point สำหรับผลกระทบอย่างสุ่มในหน่วยระดับที่ 4 ทั้ง 3 term ได้จาก upward step และ ใช้การคำนวณ relevant marginal posteriors ใน downward step

หมายเหตุ นี้เป็นการอธิบายถึงการประมาณค่า ML สำหรับโมเดล ที่เป็นตัวแปรแฝงแบบต่อเนื่อง และการรวมตัวเลข ขั้นตอนที่เทียบเท่าคือ การใช้การใช้ตัวแปรแฝงที่ไม่ต่อเนื่อง ความแตกต่างเพียงอย่างเดียวก็คือ quadrature weights ไม่คงที่ แต่มีตัวแปรอิสระที่จะถูกประมาณ

ประเด็นเกี่ยวกับความคลาดเคลื่อนมาตรฐานและการเป็นเอกลักษณ์ (identification issue)

วิธี Newton-like ใน EM algorithm ไม่ให้ความคลาดเคลื่อนมาตรฐานของ model parameter เป็น by-product ความคลาดเคลื่อนมาตรฐานที่ประมาณแล้ว สามารถได้จากการคำนวณเมตริกซ์ข้อมูลที่สังเกตได้ เมตริกซ์ของการหาอนุพันธ์อันดับสอง ของฟังก์ชั่น log likelihood ไปยังทุก model parameters อินเวอร์สของเมตริกซ์นี้ เป็นเมตริกซ์ variance-covariance ที่ประมาณแล้ว โปรแกรม Latent GOLD จะคำนวณ การหาอนุพันธ์อันดับสองโดยใช้การวิเคราะห์อนุพันธ์อันดับหนึ่ง ซึ่งการหาอนุพันธ์อันดับหนึ่งถูกกำหนดโดย EM algorithm ข้างต้น

สำหรับการตรวจสอบการเป็นเอกลักษณ์ (identifiability) จะใช้ Jacobian matrix ซึ่งเป็นเมตริกซ์ ที่เป็นอนุพันธ์อันดับหนึ่งของ $f_k(y_k)$ ไปยัง model parameter ซึ่งได้มาเหมือนกับ by-product ของ EM iteration cycle เงื่อนไขที่เพียงพอสำหรับ local identification คือ Jacobian เป็นของ full column rank (Rothenberg, 1971)

ข้อสังเกตเกี่ยวกับการประมาณค่าแบบ Maximum Likelihood โดยประยุกต์ EM algorithm

Maximum likelihood สามารถแสดงโดยใช้ EM algorithm (Agresti, Booth, Hobert, & Caffo, 2000; Bock & Aitkin, 1981) ซึ่งสามารถใช้กับกรณีที่มีข้อมูลขาดหายได้ โดยปกติ EM algorithm สามารถใช้ประมาณค่าได้ดี แต่อย่างไรก็ตาม ไม่สามารถใช้กับ Mixed models ชนิดอื่นได้ เพราะว่าจำนวนของข้อมูลในการแจกแจงภายหลังมีขนาดใหญ่ ซึ่งเป็นวิธีที่ไม่มีประสิทธิภาพ โดยเฉพาะอย่างยิ่งเกี่ยวกับปัญหาที่สัมพันธ์กับ standard EM คือ โมเดล 3 ระดับ ซึ่ง E step เกี่ยวข้องกับการคำนวณ การแจกแจงภายหลัง ของ random intercept ซึ่งสัมพันธ์กับแต่ละหน่วยของ 3 ระดับ ซึ่ง joint distribution ของ random intercept ในโมเดล 3 ระดับกับ โมเดล 2 ระดับ เป็นสมาชิกของ หน่วย 3 ระดับ quadrature nodes 10 หน่วย และหน่วยในระดับ 2 20 หน่วยต่อ หน่วยในระดับ 3 ซึ่งเป็นการแจกแจงภายหลังกับ 10^{1+20} ข้อมูล นี้แสดงถึง ความจุคอมพิวเตอร์ และเวลาที่เพิ่มขึ้นอย่างมากกับจำนวนของหน่วยในระดับ 2 ภายใน หน่วยระดับ 3 ซึ่งจะเห็นว่า EM ไม่มีประสิทธิภาพ กับ หน่วยในระดับ 2 ต่อ หน่วยในระดับ 3 ในขณะที่ EM เป็นวิธีที่เสถียรมาก และ ค่อนข้างจะเป็นวิธีที่เร็ว

MIXOR , NLMIXED และ GLLAMM ทำให้ค่า log-likelihood มีค่าสูงสุด โดยใช้ขั้นตอนวิธี Newton – type algorithm Lesaffre และ Spiessens (2001) รายงานความยากกับ MIXOR และ NLMIXED Newton type algorithm ในข้อค้นพบเกี่ยวกับ ML ใน Nonlinear mixed (2 level) models ซึ่งมีชุดคำสั่งที่แตกต่าง และขั้นตอนวิธีที่แตกต่างกันนี้ อาจจะให้ผลแตกต่างกันสำหรับจำนวนของ quadrature point เดียวกัน คาดได้ว่าปัญหานี้จะกลายเป็นปัญหาที่แย่ลงในโมเดลที่มีมากกว่า 2 ระดับ สำหรับ Newton algorithm ถูกใช้โดย GLLAMM สามารถที่จะจัดการกับ ข้อมูลที่มากกว่า 2 ระดับที่มีลักษณะเป็น nested ได้ และมีความเสถียรมากขึ้นด้วย อย่างไรก็ตาม ใน GLLAMM ใช้ การหาอนุพันธ์อันดับ 1 และ 2 ของ ค่า log likelihood ได้ ซึ่งมีการคำนวณอย่างเข้มงวดกับมากกว่าในพารามิเตอร์ที่เพิ่มขึ้น แม้ว่าไม่สามารถคาดการณ์ได้ว่า EM จะแก้ปัญหาที่สัมพันธ์กับวิธีการ Newton-type ได้ แต่อาจเป็นประโยชน์ในการใช้ EM algorithm สำหรับ โมเดล nonlinear hierarchical เป็นเครื่องมือเพิ่มเติม ประโยชน์ที่สำคัญมากที่สุดของ EM algorithm คือ การลู่เข้าโดยไม่คำนึงถึงค่าเริ่มต้น และสามารถวิเคราะห์หาอนุพันธ์สำหรับ M Step ได้อย่างทันที แต่สามารถถูกเลือกจาก generalized linear modeling package ที่มีอยู่

โดยเฉพาะอย่างยิ่งสำหรับการแจกแจงผสมเป็นภาวะปกติแบบหลายตัวแปร อย่างไรก็ตาม แทนที่จะทำงานกับ การแจกแจงพาราเมตริกซ์ สำหรับ สัมประสิทธิ์อย่างสุ่ม มีความเป็นไปได้ว่าจะใช้ความเฉพาะเจาะจงแบบ nonparametric (Aitkin, 1999; Laird, 1978) ผลนี้จะถูกอ้างถึงการวิเคราะห์การถดถอยของกลุ่มแฝง (latent class regression) หรือ การวิเคราะห์การถดถอยแบบ

ผสม (finite mixture regression) (Vermunt & Van Dijk, 2001) ประโยชน์ของแนวคิด non parametric คือ ไม่จำเป็นต้องเพิ่มข้อตกลงเบื้องต้นที่ไม่เหมาะสม และไม่สามารถพิสูจน์ได้ เกี่ยวกับการแจกแจงของ random effect (Aitkin, 1999)

โมเดลการวิเคราะห์การถดถอยของกลุ่มแฝง (Latent class regression model) ถูกพัฒนาไปมาก อย่างไรก็ตามยังเป็นเพียงกับข้อมูลที่มีโครงสร้าง 2 ระดับเท่านั้น standard EM algorithm ไม่สามารถถูกใช้สำหรับ การประมาณค่า ML ของ nonparametric hierarchical model ที่มีมากกว่า 2 ระดับ แม้ว่าเป็นไปได้ในการประมาณค่าในโมเดล โดยวิธี Newton –Raphson ซึ่งเป็นวิธีที่รู้จักกันดี ซึ่งเรียกร้องค่าเริ่มต้นที่เหมาะสมแต่ค่าเริ่มต้นที่เหมาะสมนั้นก็หาได้ยาก

ดังนั้นวิธีการประยุกต์ EM algorithm สำหรับการประมาณค่าแบบ ML ของจะเป็นวิธีที่สามารถใช้กับ parametric และ nonparametric hierarchical nonlinear model ที่มีมากกว่า 2 ระดับได้ด้วย

1.3.2 การประมาณค่าพารามิเตอร์แบบเบส์

การประมาณค่าด้วยวิธีของเบส์ (Baysian estimation) จะมีแนวคิดคือ โดยทั่วไปในการประมาณค่าจะถือว่าพารามิเตอร์ θ เป็นค่าคงที่ซึ่งไม่ทราบค่า แต่ตามวิธีของเบส์จะถือว่าค่า θ เป็นค่าของตัวแปรสุ่ม θ ที่มีการแจกแจงแสดงได้ในรูปการแจกแจงความน่าจะเป็น ซึ่งเรียกว่า การแจกแจงก่อน (prior distribution) เนื่องจากเป็นการแจกแจงที่กำหนดขึ้นก่อนที่จะมีการเก็บรวบรวมข้อมูล เมื่อมีการรวบรวมข้อมูลมาแล้วจะใช้ความรู้ที่ได้จากข้อมูลมาปรับปรุงการแจกแจงก่อน การแจกแจงที่ได้จากการปรับปรุงในภายหลังนี้ เรียกว่า การแจกแจงภายหลัง (posterior distribution) (Congdon, 2005)

การประมาณค่าพารามิเตอร์ ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ สามารถใช้ Markov chain Monte Carlo (MCMC) algorithm ซึ่งเป็นเทคนิคสำหรับการจำลองกลุ่มตัวอย่างแบบสุ่ม (simulating random samples) เทคนิค MCMC นี้เป็นเทคนิคที่ไม่ยุ่งยากซับซ้อนมากนักในทางการปฏิบัติและมีต้นทุนต่ำ นอกจากนี้การศึกษาของ J.-P. Fox (2007) กล่าวว่าเทคนิค Markov chain Monte Carlo (MCMC) เป็นเทคนิคที่มีความยืดหยุ่นมาก และเป็นเทคนิคที่สามารถวิเคราะห์ตัวแปรตามแบบแฝง (latent dependent variable) สามารถใช้ได้กับข้อสอบแบบตรวจให้คะแนน 2 ค่า (dichotomous) และข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (polytomous) โดยสถิติเบส์เซียนจะถูกใช้ในขั้นตอนการประมาณค่าพารามิเตอร์โดยการเขียน algorithms ในโปรแกรม WinBugs 1.4 โดยเขียนคำสั่งและแสดงผลภายใต้โปรแกรม R

สำหรับประเด็นที่ต้องตัดสินใจในการประมาณค่าแบบเบส์คือ Label Switching และ Posterior Distribution

Label Switching เกิดขึ้นเมื่อกลุ่มแฝงเปลี่ยนแปลง ข้ามการการประมาณค่าที่เป็นลูกโซ่ (chain) ซึ่งยังเกิดขึ้นในการประมาณค่าประเภทอื่น ๆ เช่นกัน (เช่น การประมาณค่าความน่าจะเป็นสูงสุด) ซึ่งเป็นที่น่าสนใจ ในบริบทของ MCMC มี label switching 3 ประเภทที่สามารถเกิดขึ้นได้ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ทั้ง 3 ประเภทนี้จะคล้ายกันในความหมายของ label switching มีการเปลี่ยนแปลง แต่ยังแตกต่างกันในสาเหตุ และผลกระทบ ประเภทแรกคือ Label switching within MCMC chain ประเภทที่สองคือ label switching across chains และประเภทที่สามคือ ความผันแปรของชนิดที่ 2 ใน กลุ่มแฝงระดับบุคคล (ระดับนักเรียน) สลับภายในกลุ่มแฝงระดับที่สูงกว่า (ระดับโรงเรียน/ระดับสังกัด)

ประเภทแรกของ label switching สลับภายใน chain กับการคำนวณซ้ำต่อมา ซึ่งอาจเป็นปัญหาที่สำคัญในการประมาณค่าแบบเบส์ การสลับประเภทนี้จะเกิดขึ้นเมื่อมีการขาดหายของข้อมูลที่เป็นประโยชน์กับขั้นตอนในการจำแนกระหว่างกลุ่มแฝงของ mixture model ซึ่งเป็นของกลุ่มการวัดเดียวกัน (McLachlan & Peel, 2004) ปัญหานี้ยากที่จะจัดการสำหรับ MMixIRT เนื่องจากเป็นโครงสร้างพหุระดับ จึงใช้แนวคิด hierarchical mixtures of experts model (Jordan & Jacobs, 1992) และมุมมองของ MMixIRT ซึ่งมีโครงสร้างแบบผสมที่เป็น 2 ระดับ ระดับที่สูงกว่า (ระดับโรงเรียน) และสัดส่วนการผสมในกลุ่มแฝง เป็น π_k ที่ระดับต่ำกว่า (ระดับนักเรียน) สัดส่วนการผสมในกลุ่มแฝงเป็น $\pi_{g|k}$ เงื่อนไขที่จำเป็นสำหรับการระบุของโมเดลแบบผสมพหุระดับคือ โมเดลในระดับสูง มีโครงสร้างของ โมเดลกลุ่มแฝงที่ระบุได้ (Vermunt, 2007a) การระบุของระดับต่ำของโมเดลเป็นเงื่อนไขที่เพียงพอ แต่ไม่จำเป็น เมื่อจำนวนของระดับที่สูง ใหญ่กว่าในระดับที่ 1 เงื่อนไขที่จำเป็นสำหรับการระบุ คือ π_k สำหรับ กลุ่มแฝงที่ K ถูกระบุ จากที่กล่าวก่อนหน้า label switching สามารถสังเกตได้ เมื่อการกระโดดที่แตกต่างเกิดขึ้นอย่างมีนัยสำคัญของพารามิเตอร์ และ เมื่อความหนาแน่นสำหรับพารามิเตอร์เป็นแบบการคูณ (multiple mode) (Stephens, 2000) ถ้า multiple mode ไม่มีค่าสำหรับ π_k label switching ชนิดนี้ก็จะไม่แสดง และเงื่อนไขที่จำเป็นในการระบุจะมีความสอดคล้อง

ประเภทที่สอง ของ label switching เกิดขึ้นใน MCMC chain เช่น สำหรับการทำซ้ำที่แตกต่างกันในการศึกษาข้อมูลจำลอง หรือสำหรับค่าเริ่มต้นที่แตกต่างกัน ความผันแปรของ label switching ประเภทนี้ สามารถเกิดขึ้นภายในแต่ละการผสมในระดับโรงเรียนใน MMixIRT เพราะว่าสัดส่วนของระดับนักเรียนถูกจำลองภายในการผสมระดับโรงเรียน ตัวแปรภายหลังของ label switching นี้ เป็นประเภทที่สาม ของ label switching และ เกิดขึ้นเมื่อ สมาชิกของระดับนักเรียน

แตกต่างกันสำหรับการผสมในระดับโรงเรียน ถ้า label switching นี้เกิดขึ้นในการศึกษาข้อมูลจำลอง การประมาณค่าพารามิเตอร์สามารถเปรียบเทียบกับการ พารามิเตอร์โดยทั่วไปได้ เพื่อตัดสินใจว่า label นี้ควรจะถูก ประยุกต์เข้ากับแต่ละกลุ่มแฝง สำหรับแต่ละการผสมในระดับโรงเรียน สำหรับข้อมูลเชิงประจักษ์ สมาชิกในกลุ่มสามารถสอดคล้องข้าม chain และข้ามการผสมในระดับโรงเรียนโดยดูจาก รูปแบบของความสามารถ สัดส่วนการผสมและค่าความยาก

Posterior Distribution

ความน่าจะเป็นของการได้รับการตอบสนองที่ถูกต้องฟังก์ชั่นของ g, k และ θ_{jtgk} ความน่าจะเป็นในการระบุชั้น $P\left(y_{ijt}=1\middle|g,k,\theta_{jtgk}\right)$ ถูกแสดงสัดส่วนการผสม π_k และ $\pi_{g|k}$ ตัวแปรบ่งชี้ที่ไม่สามารถสังเกตได้ (ζ_{jtgk}) ถูกใช้ เพราะว่า แต่ละบุคคล j ในกลุ่ม t ถูกกำหนดในทั้งกลุ่มแฝงระดับนักเรียน g และกลุ่มแฝงระดับโรงเรียน k ในการทำซ้ำ l ฟังก์ชั่นไลค์ลิฮูด ของ MMixIRT คือ

$$L\left(g,k,\theta_{jtgk}\right)=\prod_{i=1}^{I}\prod_{j=1}^{J}\left[\left\{\sum_{k=1}^{K}\sum_{g=1}^{G}\pi_k\cdot\pi_{g|k}P\left(y_{ijt}=1\middle|g,k,\theta_{jtgk}\right)\right\}^{u_{ij}}\cdot\left\{1-\sum\sum\pi_k\cdot\pi_{g|k}P\left(y_{ijt}=1\middle|g,k,\theta_{jtgk}\right)\right\}^{1-u_{ij}}\right]^{\zeta_{jtgk}} \quad \text{.....(2.34)}$$

เมื่อ u_{ij} เป็นคะแนนแบบสองค่า คือ 0 และ 1 ζ_{jtgk} เป็น 1 เมื่อผู้สอบ j อยู่ในรูปการผสม g และ k และ ζ_{jtgk} เป็น 0 เมื่อในการทำซ้ำเป็นอย่างอื่น

$$\theta_{jtgk}=\sqrt{\sigma_{gk}^2}\cdot\eta_{jtgk} \quad \text{.....(2.35)}$$

การแจกแจงภายหลังร่วมกันสำหรับการใช้ก่อนหน้าบน $\pi_{g|k}$ และ π_k

$$S=\left\{g,k,\eta_{jtgk},\mu_{gk},\sigma_{gk},\beta_{igk},\pi_k,\pi_{g|k}\right\} \quad \text{.....(2.36)}$$

สามารถเขียนเป็น

$$P\left(S|U\right)\ \alpha\ L\left(g,k,\eta_{jtgk}\right)P\left(\eta_{jtgk}\middle|\mu_{gk}\right)P\left(\mu_{gk}\right)P\left(g\middle|\pi_{g|k}\right)\cdot P\left(\pi_{g|k}\right)P\left(k|\pi_k\right)P\left(\pi_k\right)P\left(\sigma_{gk}\right)P\left(\beta_{igk}\right)$$

การแจกแจงภายหลังร่วมกัน สำหรับ multinomial logistic regression model บน $\pi_{g|k}$ และ π_k

$$S=\left\{g,k,\eta_{jtgk},\mu_{gk},\sigma_{gk},\beta_{igk},\pi_k,\pi_{g|k},\gamma_{ogk},\gamma_{pg}.\gamma_{ok},\gamma_{pk}\right\} \quad \text{.....(2.37)}$$

สามารถเขียนเป็น

$$P\left(S|U\right)\ \alpha\ L\left(g,k,\eta_{jtgk}\right)P\left(\eta_{jtgk}\middle|\mu_{gk}\right)P\left(\mu_{gk}\right)P\left(g\middle|\pi_{g|k}\right)P\left(\pi_{g|k}\middle|\gamma_{ogk},\gamma_{pg}\right)P\left(\gamma_{ogk}\right)P\left(\gamma_{pg}\right)$$
$$\cdot P\left(k|\pi_k\right)P\left(\pi_k\middle|\gamma_{ok},\gamma_{pk}\right)P\left(\gamma_{ok}\right)P\left(\gamma_{pk}\right)P\left(\sigma_{gk}\right)P\left(\beta_{igk}\right)$$

การประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ Multilevel Mixture IRT model (MMixIRT)

การประมาณค่าพารามิเตอร์ในการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 1 และ 2 พารามิเตอร์ สามารถประมาณโดยโปรแกรมสำเร็จรูป Latent GOLD 4.0 โดยการใช้วิธีการประมาณค่าเป็นไปได้สูงสุด (maximum likelihood estimation) (Vermunt & Magidson, 2005) และใช้ขั้นตอน EM ซึ่งในการประมาณค่าอาจใช้โปรแกรม MPLUS (L. K. Muthén & Muthén, 2008) นอกจากนี้ยังสามารถประมาณค่าพารามิเตอร์โดยใช้โปรแกรม WINBUGS 1.4 เช่นเดียวกับงานวิจัยของ Cho และ Cohen (2010) ซึ่งใช้วิธีการประมาณค่าแบบเบส์ ด้วยขั้นตอน Markov chain Monte Carlo

สำหรับโมเดลในงานวิจัยนี้เป็นโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ไม่สามารถประมาณค่าด้วยโปรแกรม Latent GOLD 4.0 ดังนั้นในการศึกษาครั้งนี้จึงใช้การเขียนคำสั่งประมวลผลบนโปรแกรม R

1.4 วิธีการเลือกโมเดล (Model Selection Methods)

การเลือกโมเดลการตอบสนองข้อสอบที่เหมาะสม เป็นพื้นฐานบนความสอดคล้องของข้อมูล ถ้าข้อมูลไม่สอดคล้องกับโมเดล จะมีผลให้คุณสมบัติของโมเดลการตอบสนองข้อสอบ เช่น ความไม่แปรเปลี่ยนของพารามิเตอร์สำหรับประชากร ไม่สามารถครอบคลุมได้ (Hambleton, 1991) ถ้าโมเดลไม่สอดคล้องกับข้อมูลอาจจะเป็นประโยชน์สำหรับการศึกษาสมมติฐาน เกี่ยวกับ ความแตกต่างในขั้นตอนการตอบสนอง หน้าที่ของการวิเคราะห์โมเดลการตอบสนองข้อสอบแบบผสมเป็นการเลือกโมเดลที่สอดคล้องที่สุดกับข้อมูล

สำหรับโมเดลการตอบสนองข้อสอบแบบ nested การเลือกโมเดลสามารถใช้การทดสอบ likelihood ratio (LR) ในการเปรียบเทียบความสอดคล้องที่สัมพันธ์กับโมเดล สถิติทดสอบ LR (G^2) เป็นการแจกแจงเหมือนกับสถิติไคว์สแควร์ โดยที่ G^2 เป็นการคำนวณเป็น $-2\times\log$ ของ likelihood สำหรับการให้โมเดล ความแตกต่างระหว่าง G^2s สำหรับ สองโมเดลเป็นการแจกแจงเป็นไคว์สแควร์เหมือนกัน และสามารถถูกทดสอบเพื่อตัดสินใจว่าโมเดลใดสอดคล้องที่สุด (Bock & Aitkin, 1981) และสำหรับ mixture IRT model ไม่เป็น nested ดังนั้น LR ไม่เหมาะสมสำหรับการเลือกโมเดล

ทางเลือกหนึ่งเป็นการเลือกโมเดลที่สอดคล้อง โดยใช้ เช่น statistics-based Akaike information criterion (AIC) หรือ Baysian information criterion (BIC) โดยทั้ง AIC และ BIC เป็นกรณีพิเศษของ information criterion โดยทั่วไป แม้ว่าการทดสอบอย่างมีนัยสำคัญจะเป็นไปไม่ได้ในทางสถิติ การเปรียบเทียบค่าสำหรับโมเดลที่แตกต่างกันทำให้ข้อมูลที่ได้ซึ่งสามารถใช้วัดความ

แตกต่างสัมพัทธ์ระหว่างผลที่ได้ สถิตินี้ยังเหมาะสมเมื่อประมาณค่าแบบ maximum likelihood ของพารามิเตอร์ที่ได้รับ

ในบางครั้งการประมาณค่า asymptotic ของ item parameters อาจจะไม่สามารถทำได้ (Lin & Dayton, 1997) เมื่อสิ่งนี้เกิดขึ้น การใช้ AIC และ BIC จะไม่เหมาะสม เช่นสถานการณ์ การประมาณค่า Bayesian parameter สามารถเป็นทางเลือกที่มีประสิทธิภาพ การใช้การประมาณ Markov chain Monte Carlo (MCMC) ผลโดยทั่วไปในการประมาณ Bayesian ของ โมเดลพารามิเตอร์ การเลือกโมเดลสำหรับผลนี้ประกอบไปด้วย pseudo-Bayes factor (PsBF; Bolt et al., 2001; Geisser & Eddy, 1979; Gelfand & Dey, 1994) the deviance information coefficient (DIC; Spiegelhalter, Best, & Carlin, 1998) และ posterior predictive model checks (PPMC; Gelman, Carlin, Stern, & Rubin, 1996) โดย AIC และ BIC สำหรับการใช้กับการประมาณค่าแบบ MCMC เป็นดัชนีซึ่งรายงานในการศึกษานี้เหมือนกับ AIC และ BIC

จากการศึกษาของ Li และคณะ (2009) พิจารณาการเลือกดัชนีสำหรับใช้กับโมเดลการตอบสนองข้อสอบแบบผสมของข้อสอบที่มีการให้คะแนนแบบสองค่า โดยใช้ดัชนีดังนี้ (1) Akaike's information coefficient (AIC) (2) Bayesian information coefficient (BIC) (3)deviance information coefficient (DIC) (4) pseudo-Bayes factor (PsBF) และ (5) posterior predictive model checks (PPMC) โดยการจำลองสถานการณ์ กับ IRT Model แบบ 3PL ใช้กับแบบสอบจำนวน 6, 15 และ 30 ข้อ เปรียบเทียบกับจำนวนผู้สอบ 600 และ 1,200 คน และเงื่อนไขแบบผสม 4 แบบ (4 simulated mixture conditions) พบว่า Bayesian information coefficient (BIC) เลือกโมเดลที่ถูกต้องดีที่สุด ภายใต้เงื่อนไขการจำลองข้อมูล และเหมาะสมสำหรับการวิเคราะห์กับ 1PL 2PL และ 3PL สำหรับ pseudo-Bayes factor (PsBF) เป็นดัชนีที่มีประสิทธิภาพมาก สำหรับ AIC และ DIC มีแนวโน้มที่จะเลือกโมเดลภายใต้เงื่อนไขที่ซับซ้อน และสำหรับ posterior predictive model checks (PPMC) จะมีประสิทธิภาพน้อยที่สุด

ในการศึกษาครั้งนี้จะศึกษากับดัชนี BIC ซึ่งดัชนี BIC คือ ดัชนีที่ใช้ในการบอกถึงความเหมาะสมของจำนวนกลุ่มที่ถูกจำแนก ค่า BIC มีที่มาจากการเปรียบเทียบค่า Maximum Likelihood ระหว่าง 2 โมเดล ซึ่งหลายโปรแกรมแสดงให้เห็นว่าหากค่า BIC มีค่าน้อยที่สุดจำนวนกลุ่มที่ถูกแบ่งออกมาในกรณีดังกล่าวจะมีความเหมาะสมมากที่สุด โดย

$$BIC = -2D^t + m\log n \qquad \text{.....(2.38)}$$

เมื่อ m เป็นจำนวนพารามิเตอร์ และ n เป็น จำนวนครั้งของการสังเกต

ในที่นี้จะเปรียบเทียบผลสำหรับวิธีการสำหรับการใช้กับข้อมูลที่เป็นข้อสอบแบบให้คะแนน 0, 1 (dichotomous item) ดังนั้นในงานวิจัยนี้ การตรวจสอบการจำแนกกลุ่มความสามารถแฝงโดย

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการโมเดล three-parameter logistic (3PL) พารามิเตอร์ กับ mixture IRT ซึ่งสามารถประมาณโดยใช้ขั้นตอน Bayesian ใช้ดัชนี BIC ในการกำหนดจำนวนกลุ่มแฝง

ตอนที่ 2 การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลนับเป็นสาขาหนึ่งของวิชาคณิตศาสตร์เชิงทดลอง ซึ่งเป็นเทคนิคที่ใช้สำหรับการแก้ปัญหาในการคำนวณทางคณิตศาสตร์สำหรับสถานการณ์ที่มีกระบวนการที่ซับซ้อน ซึ่งการวิเคราะห์ทั่วไปไม่สามารถดำเนินการได้อย่างครบถ้วนและสมบูรณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลสามารถให้ผลสรุปจากสภาพการณ์ที่สร้างขึ้นใน การทดลองที่แสดงเหตุการณ์เกิดขึ้นทั้งทางบวกและทางลบ เนื่องจากเทคนิคนี้สามารถจำลองข้อมูลได้ หลายๆ ครั้งจากข้อมูลลักษณะเดียวกัน ผลการศึกษาจึงสามารถสะท้อนให้เห็นถึงเหตุการณ์ที่น่าจะเกิดขึ้นได้จริง โดยหลักการในการจำลองจะสร้างตัวแบบ (model) โดยการใช้เลขสุ่ม (Random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา โดยการนำเสนอในตอนนี้จะนำเสนอ แบ่งออก เป็น 6 ส่วน ได้แก่ ส่วนแรกเป็นการนำเสนอความเป็นมาของเทคนิคมอนติคาร์โล ส่วนที่สองเกี่ยวกับขั้นตอนในการจำลองข้อมูล ส่วนที่สามเกี่ยวกับการประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบ ส่วนที่สี่ประสิทธิภาพของการจำลองข้อมูล ส่วนที่ห้าข้อดีและข้อจำกัดของการศึกษาด้วยเทคนิคการจำลองข้อมูล และส่วนสุดท้าย บทสรุปจากการศึกษาการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล มีรายละเอียดในแต่ละส่วนดังนี้

2.1 ความเป็นมาของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

ในเริ่มต้นบุฟฟอง ได้ทำการทดลองเพื่อสร้างสมการจากความสัมพันธ์ระหว่างความยาวของเข็มกับระยะทางระหว่างเส้นขนานทั้งสองหลังจากนั้นจึงได้มีความพยายามในการพิสูจน์สูตรคณิตศาสตร์ที่เขาได้คิดค้นขึ้นมา จึงทำให้มีการจำลองข้อมูลเกิดขึ้นในปี ค.ศ.1777

สำหรับการจำลองข้อมูลที่เรียกว่าการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลนั้น เริ่มปรากฏขึ้นในปี 1940 โดยนักวิทยาศาสตร์ในชั้นเรียนทางคณิตศาสตร์ที่ทำงานด้านการพัฒนาอาวุธ นิวเคลียร์ในลอส อาลามอส โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเป็นเทคนิคที่สร้างขึ้นมาจากทฤษฎีความน่าจะเป็น (Probability theory) และทฤษฎีการสุ่ม (Theory of random) ซึ่งเป็นรากฐานที่สำคัญของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

พัฒนาการของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล สามารถแบ่งออกได้เป็น 3 ช่วงหลัก ๆ ได้แก่ ช่วงที่ 1 เป็นช่วงก่อนการใช้เทคนิคมอนติคาร์โล ซึ่งเป็นช่วงที่ใช้ในการศึกษาปัญหาต่างๆ

ทางสถิติ เช่น การสุ่มตัวอย่างทางสถิติ เกี่ยวกับการแจกแจงของค่าสัมประสิทธิ์สหสัมพันธ์ การแจกแจงสถิติที ช่วงที่ 2 เป็นช่วงที่มีศัพท์คำว่า เทคนิคมอนติคาร์โล เกิดขึ้น โดย Metropolis และ Ulam เป็นผู้ใช้ศัพท์คำว่า “Monte Carlo” เป็นครั้งแรก โดยใช้เทคนิคนี้เป็นเครื่องมือวิจัยเกี่ยวกับระเบิดปรมณูในยุคสงครามโลกครั้งที่ 2 และช่วงที่ 3 เริ่มต้นประมาณปี ค.ศ.1970 เป็นช่วงที่มีการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล โดยการนำคอมพิวเตอร์ความเร็วสูงมาใช้ ในการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล จึงทำให้การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเริ่มเป็นที่นิยมในการศึกษาเพื่อแก้ไขปัญหาทางสถิติ (อ้างถึงใน ชนะศึก นิชานนท์, 2553)

2.2 ขั้นตอนการจำลองข้อมูล

ชนะศึก นิชานนท์ (2553) กล่าวถึงการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล เป็นวิธีการจำลองข้อมูลด้วยการสร้างตัวแบบ (model) โดยการใช้เลขสุ่ม (Random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา สามารถสรุปขั้นตอนที่สำคัญในการจำลองข้อมูลออกเป็น 3 ขั้นตอนใหญ่ๆ ดังนี้

2.2.1 การสร้างตัวเลขสุ่ม

ขั้นตอนนี้ถือเป็นขั้นตอนที่สำคัญมากสำหรับเทคนิคมอนติคาร์โล เนื่องจาก การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลนั้นจะใช้เลขสุ่มมาช่วยในการหาคำตอบให้กับปัญหาโดยลักษณะของตัวเลขสุ่มที่นำมาใช้ จะมีการแจกแจงแบบสม่ำเสมอ (uniform distribution) ในช่วง (0, 1) สำหรับวิธีการสร้างตัวเลขสุ่มที่ดีนั้น ลักษณะของเลขสุ่มที่ถูกสร้างขึ้นจะต้องมีการแจกแจงแบบสม่ำเสมอในช่วง (0, 1) ตัวเลขสุ่มแต่ละตัวเป็นอิสระกันและมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ (มีวัฏจักรยาว)

เนื่องด้วยการสร้างตัวแปรสุ่ม (Random Variable) นับว่าเป็นขั้นตอนที่สำคัญของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ดังนั้นในขั้นตอนนี้จึงมีการนำเสนอเกี่ยวกับตัวแปรสุ่มเพื่อให้เข้าใจเกี่ยวกับลักษณะของตัวแปรสุ่มจึงแบ่งการนำเสนอตัวแปรสุ่มออกเป็น 3 ส่วน ส่วนแรกเป็นการนำเสนอความหมายของตัวแปรสุ่ม ส่วนที่สองเป็นการนำเสนอประเภทของตัวแปรสุ่ม และส่วนสุดท้ายเป็นการนำเสนอการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม โดยมีรายละเอียดในแต่ละส่วนของตัวแปรสุ่มดังนี้

(1) ความหมายของตัวแปรสุ่ม

ตัวแปรสุ่ม เป็นฟังก์ชันที่มีค่าเป็นจำนวนจริง (Real-valued function) ซึ่งกำหนดโดยแต่ละสมาชิกในปริภูมิตัวอย่างจากการลองสุ่ม (random trial) ซึ่งหมายถึงการกระทำใดๆ ที่ผู้กระทำไม่สามารถทราบผลลัพธ์ล่วงหน้าจนกว่าจะได้ทำเสร็จสิ้นไปแล้ว จึงจะทราบผลลัพธ์ที่ถูกต้อง (คณาจารย์ภาควิชาคณิตศาสตร์คณะวิทยาศาสตร์จุฬาลงกรณ์มหาวิทยาลัย, 2551)

(2) ประเภทของตัวแปรสุ่ม

ตัวแปรสุ่ม สามารถจำแนกได้เป็น 2 ชนิด ได้แก่ ตัวแปรสุ่มต่อเนื่อง และ ตัวแปรสุ่มไม่ต่อเนื่อง

ตัวแปรสุ่มไม่ต่อเนื่อง (discrete random variable) คือ ตัวแปรสุ่มที่มีค่าเป็นจำนวนที่นับได้ครบ หรือ เป็นตัวแปรที่มีค่าเป็นจำนวนที่จับคู่แบบหนึ่งต่อหนึ่งกับจำนวนเต็มได้ทั้งหมด ปริภูมิตัวอย่างของตัวแปรสุ่มประเภทนี้จะเรียกว่า ปริภูมิตัวอย่างที่ไม่ต่อเนื่อง (discrete sample space)

ตัวแปรสุ่มต่อเนื่อง (continuous random variable) คือ ตัวแปรสุ่มที่มีค่าต่อเนื่องกันได้หลายค่านับไม่ถ้วน ปริภูมิตัวอย่างที่มีจุดตัวอย่างเป็นตัวแปรสุ่มต่อเนื่อง เรียกว่า ปริภูมิตัวอย่างที่ต่อเนื่อง (continuous sample space) ดังนั้นจะเห็นได้ว่าคะแนนมาตรฐาน (Z-score) ที่ได้จากการทดสอบ หรือค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบ เป็นคะแนนสุ่มแบบต่อเนื่อง

(3) การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม มีหลายประเภท แต่ในการนำเสนอครั้งนี้ จะเป็นการนำเสนอที่พบบ่อยสำหรับการวิเคราะห์ข้อสอบ ซึ่งได้แก่ การแจกแจงแบบยูนิฟอร์ม (Uniform distribution) การแจกแจงแบบปกติ (Normal distribution) การแจกแจงแบบปกติหาค่าโดยประมาณของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงทวินาม การแจกแจงแกมมา (Gamma distribution) การแจกแจงแบบเบต้า (Beta distribution) การแจกแจงแบบไคว์สแควร์ (Chi-square distribution) และการแจกแจงแบบที (t-distribution) มีรายละเอียดเกี่ยวกับการแจกแจง แต่ละประเภทดังนี้ (คณาจารย์ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2551)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(3.1) การแจกแจงแบบยูนิฟอร์ม (Uniform distribution)

การแจกแจงแบบยูนิฟอร์มมีลักษณะการแจกแจงที่มีค่าได้ทุกค่าจริงในช่วง $[a,b], -\infty < a,b < \infty$ ด้วยความน่าจะเป็นเท่า ๆ กัน สามารถเขียนแทนด้วยสัญลักษณ์ $X \sim U$ (a,b) โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x) = \begin{cases} \frac{1}{b-a} & ; \quad a < x < b \quad \text{เมื่อ x มีค่าอื่น ๆ} \end{cases} \qquad \text{.....(2.40)}$$

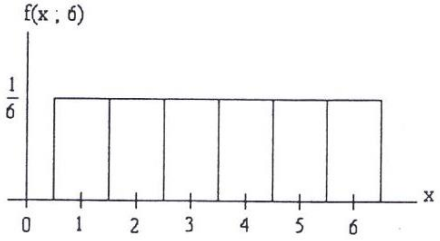
การแจกแจงแบบยูนิฟอร์มจะมีค่าเฉลี่ยของ X เท่ากับ

$$\mu = \frac{a+b}{2} \qquad \text{.....(2.41)}$$

การแจกแจงแบบยูนิฟอร์มจะค่าความแปรปรวนของ X เท่ากับ

$$\sigma^2 = \frac{(b-a)^2}{12} \qquad \text{.....(2.42)}$$

การแจกแจงแบบยูนิฟอร์ม สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจงแบบยูนิฟอร์ม สำหรับค่าพารามิเตอร์บางค่าได้ดังนี้



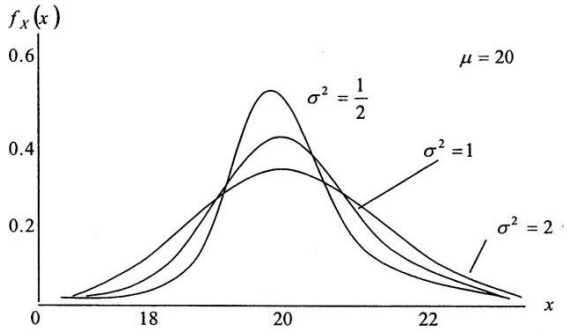
ภาพที่ 2.6 การแจกแจงแบบยูนิฟอร์ม

(3.2) การแจกแจงแบบปกติ (Normal distribution)

การแจกแจงแบบปกติ นับเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มต่อเนื่องที่สำคัญที่สุด (คณาจารย์ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์ มหาวิทยาลัย, 2551; มานพ วราภักดิ์, 2548) เนื่องจากเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่ส่วนมากจะมีค่าใกล้เคียงค่าเฉลี่ยของตัวแปรเหล่านั้น จะมีค่าของตัวแปรที่มากกว่าหรือน้อยกว่าค่าเฉลี่ยเป็นส่วนน้อย การแจกแจงแบบปกตินี้ บางครั้งอาจจะเรียกว่า การแจกแจงแบบเกาส์ (Gaussian distribution) ตัวแปรสุ่มที่มีการแจกแจงแบบปกติ เรียกว่า ตัวแปรสุ่มปกติ (normal random variable) โดยสามารถเขียนแทนด้วยสัญลักษณ์ $X \sim N(\mu, \sigma^2)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} ; -\infty < x < \infty \qquad \text{.....(2.43)}$$

การแจกแจงแบบปกติมีจะมีค่าเฉลี่ยของ X เท่ากับ μ และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ σ สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจงปกติ สำหรับค่าพารามิเตอร์บางค่าได้ดังนี้



ภาพที่ 2.7 การแจกแจงแบบปกติ

สำหรับงานวิจัยนี้ใช้การจำลองข้อมูลโดยกำหนดให้มีการแจกแจงปกติ

(3.3) การแจกแจงแบบปกติหาค่าโดยประมาณของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงทวินาม

การแจกแจงของตัวแปรสุ่ม (x) ที่มีการแจกแจงแบบทวินาม b (x ; n, p) เมื่อมีค่า n น้อยจะหาค่า b (x ; n , p) ได้จากการแจกแจงปัวซอง (Poisson distribution) โดยการประมาณของความน่าจะเป็นที่มีการแจกแจงทวินามเมื่อ n มีค่ามาก และ p มีค่าใกล้ 0 หรือ 1 แต่ในครั้งนี้จะใช้ การแจกแจงแบบปกติหาค่าโดยประมาณของความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงทวินาม หาค่าโดยประมาณของ b (x ; n, p) เมื่อ n มีค่ามากพอ โดยสามารถเขียนแทนด้วยสัญลักษณ์ $Z \sim N(0,1)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$Z = \frac{X - np}{\sqrt{npq}} \qquad \text{.....(2.44)}$$

การแจกแจงจะมีค่าเฉลี่ยของ X เท่ากับ $\mu = np$ และมีค่าความแปรปรวนของ X เท่ากับ $\sigma^2 = npq$

(3.4) การแจกแจงแกมมา (Gamma distribution)

การแจกแจงแบบแกมมามีลักษณะการแจกแจงในลักษณะเบ้ขวา (positive skewness) โดยตัวแปรสุ่มที่มีการแจกแจงแบบแกมมา ค่าพารามิเตอร $\alpha > 0$ และ $\beta > 0$ โดย สามารถเขียนแทนด้วยสัญลักษณ์ $X \sim G(\alpha, \beta)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{B}}; X > 0 \\ 0 \end{cases} \qquad \text{.....(2.45)}$$

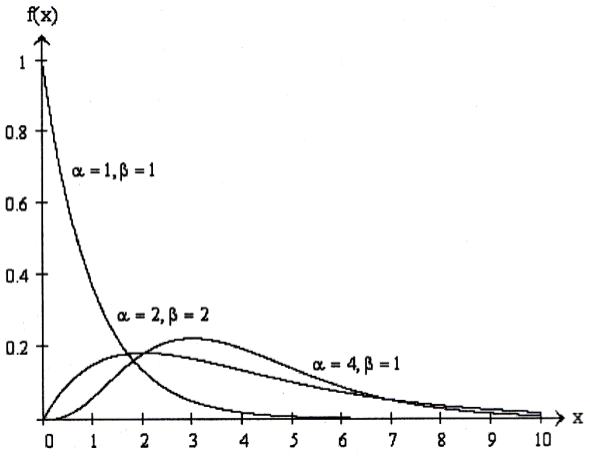
การแจกแจงแบบแกมมาจะมีค่าเฉลี่ยของ X เท่ากับ

$$\mu = \frac{\alpha}{\lambda} = \frac{\beta\gamma(\alpha+1)}{\gamma(\alpha)} = \alpha\beta \qquad \text{.....(2.46)}$$

การแจกแจงแบบแกมมาจะมีค่าความแปรปรวนของ X เท่ากับ

$$\sigma^2 = \frac{\alpha}{\lambda^2} = \alpha\beta^2 \qquad \text{.....(2.47)}$$

การแจกแจงแบบแกมมา สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจง แกมมา สำหรับค่าพารามิเตอร์บางค่าได้ดังนี้



ภาพที่ 2.8 การแจกแจงแบบแกมมา

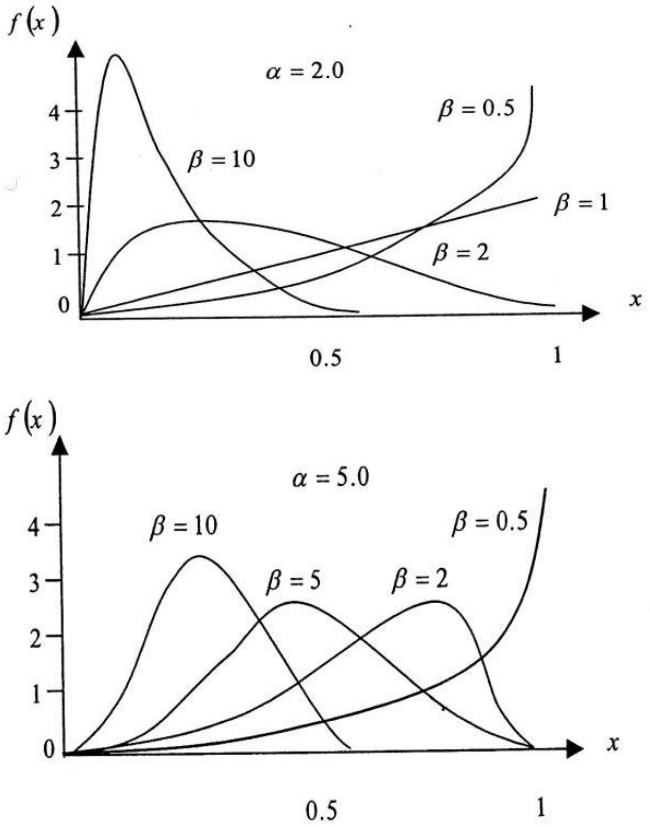
(3.6) การแจกแจงแบบเบต้า (Beta distribution)

การแจกแจงแบบแกมมามีลักษณะการแจกแจงในลักษณะเบ้ซ้าย (negative skewness) โดยตัวแปรสุ่มที่มีการแจกแจงแบบเบต้า ค่าพารามิเตอร $\alpha > 0$ และ $\beta > 0$ โดย สามารถเขียนแทนด้วยสัญลักษณ์ $X \sim Be(\alpha, \beta)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \quad ; \alpha > 0, \beta > 0 \qquad \text{.....(2.48)}$$

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & ; 0 \le x \le 1 \end{cases} \qquad \text{.....(2.49)}$$

การแจกแจงแบบเบต้าจะมีค่าเฉลี่ยของ X เท่ากับ $\mu = \dfrac{\alpha}{\alpha+\beta}$ และมีค่าความแปรปรวนของ X เท่ากับ $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ การแจกแจงแบบเบต้า สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจง แกมมา สำหรับค่าพารามิเตอร์บางค่าได้ดังนี้



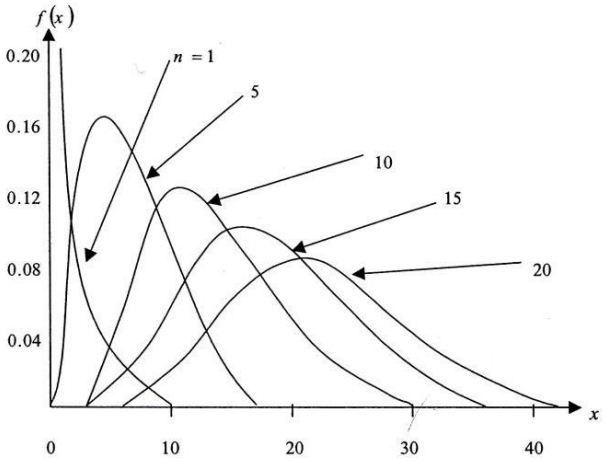
ภาพที่ 2.9 การแจกแจงแบบเบต้า

(3.7) การแจกแจงแบบไคว์สแควร์ (Chi-square distribution)

การแจกแจงแบบไคว์สแควร์ จัดเป็นการแจกแจงกรณีหนึ่งของการแจกแจงแบบแกมมา $G(\alpha,\lambda)$ ในกรณีที่ $\lambda = 0.5$ และ $\alpha = n/2$ โดย n เป็นค่าพารามิเตอร์ที่มีจำนวนเต็มบวก และเรียก n ว่าเป็นระดับขั้นความเสรี หรือ จำนวนองศาความเสรี (Number of degree of freedom; df) การแจกแจงแบบนี้จะให้ประโยชน์มากสำหรับการอนุมานเชิงสถิติ โดยสามารถเขียนแทนด้วยสัญลักษณ์ $X \sim \chi^2(n)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(\chi) = \begin{cases} \dfrac{1}{2^{\frac{V}{2}}\gamma\left(\dfrac{v}{2}\right)} x^{\frac{v}{2}-1} e^{\frac{-x}{2}} \\ 0 \end{cases} \qquad \text{.....(2.50)}$$

การแจกแจงแบบไคว์สแควร์จะมีค่าเฉลี่ยของ X เท่ากับ $\mu = v$ และมีค่าความแปรปรวนของ X เท่ากับ $\sigma^2 = 2v$ การแจกแจงแบบไคว์สแควร์ สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจงไคว์สแควร์สำหรับค่าพารามิเตอร์บางค่าได้ดังนี้



ภาพที่ 2.10 การแจกแจงแบบไคว์สแควร์

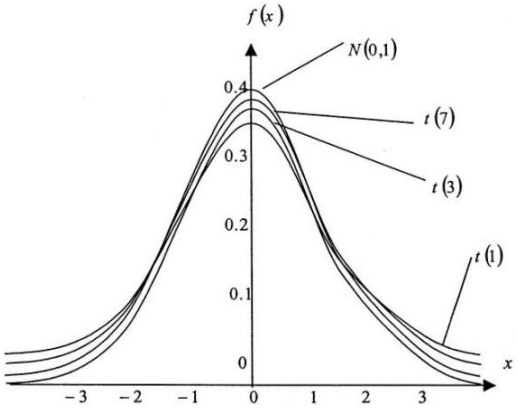
(3.8) การแจกแจงแบบที (t- distribution)

การแจกแจงแบบที บางครั้งอาจเรียกว่าการแจกแจงทีของสติวเดนท์ (Student's t-distribution) ตัวแปรสุ่ม x ที่มีการแจกแจงความน่าจะเป็นที ด้วยระดับขั้นความเสรี n สามารถเขียนแทนด้วยสัญลักษณ์ $X \sim t(n)$ โดยมีฟังก์ชันความน่าจะเป็นดังนี้

$$h(t) = \frac{\gamma\left(\frac{\gamma+2}{2}\right)}{\gamma\left(\frac{v}{2}\right)\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{\frac{-v+1}{2}} ;-\infty < t < \infty \qquad \text{.....(2.51)}$$

การแจกแจงแบบทีจะมีค่าเฉลี่ยของ X เท่ากับ $\mu = 0$ และมีค่าความแปรปรวนของ X เท่ากับ $\sigma^2 = \frac{v}{v-2}$ เมื่อ $v \geq 3$

การแจกแจงแบบแบบที สามารถแสดงตัวอย่างกราฟของฟังก์ชันการแจกแจงแบบที สาหรับค่าพารามิเตอร์บางค่าได้ดังนี้



ภาพที่ 2.11 การแจกแจงแบบที

(4) การประยุกต์ปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม

ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ศึกษา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจจะมีขั้นตอนอื่นอีกหลายๆ ขั้นตอน ซึ่งขั้นตอนเหล่านี้มีบางขั้นตอนที่ต้องใช้ตัวเลขสุ่ม

(5) การทดลองกระทำการสุ่ม

การทดลองกระทำการสุ่มเป็นการทดลองโดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะซ้ำๆกัน (replication) เพื่อประมาณค่าที่แท้จริง โดยจะใช้ความคลาดเคลื่อนมาตรฐาน (MSE) หรือ รากที่สองของความคลาดเคลื่อนมาตรฐาน (RMSD) ในการประเมินประสิทธิภาพของการประมาณค่า

3.3 การประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบ

การวิเคราะห์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ มีการประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์เป็นจำนวนมาก ในหลายด้าน เช่น การประเมินกระบวนการในการประมาณค่าหรือความครอบคลุมของค่าพารามิเตอร์ การประเมินคุณสมบัติทางสถิติของโมเดลในการวิเคราะห์ข้อสอบ การเปรียบเทียบวิธีการต่างๆ ในโมเดลการวิเคราะห์ข้อสอบ ไม่ว่าจะเป็นการทำหน้าที่ต่างกันของข้อสอบหรือการประเมินผลที่ได้จากการวัดหลายมิติ เป็นต้น (M. Harwell, Stone, Hsu, & Kirisci, 1996) จะเห็นว่าการวิเคราะห์ข้อสอบโดยการประยุกต์ใช้เทคนิคมอนติคาร์โลมีความสำคัญ ในที่นี้มีการนำเสนอแบ่งออกเป็น 2 ส่วนคือ ส่วนแรกเกี่ยวกับการประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบ และ ข้อพึงระวังสำหรับการวิเคราะห์ข้อสอบด้วยการใช้เทคนิคมอนติคาร์โล ส่วนที่สองเกี่ยวกับขั้นตอนในการวิเคราะห์ข้อสอบด้วยเทคนิคมอนติคาร์โล

3.3.1 การประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบ และ ข้อพึงระวังสำหรับการวิเคราะห์ข้อสอบด้วยการใช้เทคนิคมอนติคาร์โล

การวิเคราะห์ข้อสอบในปัจจุบันมีการประยุกต์ใช้เทคนิคมอนติคาร์โลมาก โดยเฉพาะการวิเคราะห์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) โดยส่วนใหญ่ จะมีการประยุกต์ใช้เทคนิคมอนติคาร์โลสำหรับการวิเคราะห์ข้อสอบ ซึ่งสามารถสรุปการประยุกต์สำหรับการวิเคราะห์ข้อสอบได้ 3 ประเด็น (ชนะศึก นิชานนท์, 2553) ดังนี้

(1) การประเมินประสิทธิภาพของวิธีในการประมาณค่าหรือความครอบคลุมของค่าพารามิเตอร์ที่ได้จากการประมาณด้วยโมเดลการวิเคราะห์ข้อสอบ

(2) การประเมินคุณสมบัติของสถิติที่ใช้ในการวิเคราะห์ของโมเดลในการวิเคราะห์ข้อสอบว่ามีความเหมาะสมหรือไม่

(3) การเปรียบเทียบวิธีการต่างๆ ในโมเดลการวิเคราะห์ข้อสอบ ไม่ว่าจะเป็นการทำหน้าที่ต่างกันของข้อสอบ หรือการประเมินผลที่ได้จากการวัดหลายมิติ เป็นต้น

ทั้งนี้เนื่องจากการวิเคราะห์การประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์ข้อสอบสามารถวิเคราะห์เพื่อตอบคำถามตามวัตถุประสงค์ที่ซับซ้อน โดยสามารถสรุปลำดับขั้นตอนการวิเคราะห์ของเทคนิคมอนติคาร์โลได้ 8 ขั้นตอนตามลำดับดังนี้

(1) การตอบคำถามวิจัยโดยสามารถแบ่งคำถามวิจัยออกเป็นวัตถุประสงค์ที่เฉพาะเจาะจงได้ เช่น เพื่อศึกษาผลของจำนวนข้อสอบและผู้สอบที่แตกต่างกัน เพื่อลักษณะการแจกแจงค่าพารามิเตอร์ข้อสอบที่ส่งผลต่อการประมาณค่าพารามิเตอร์ของโมเดลการวิเคราะห์ข้อสอบ เป็นต้น

(2) การกำหนดเงื่อนไขประเภทต่างๆ ของตัวแปรต้นที่ส่งผลต่อตัวแปรตามได้ เช่น จำนวนข้อสอบ ส่งผลต่อค่าพารามิเตอร์ของผู้สอบหรือไม่

(3) การออกแบบการทดลองให้มีความสอดคล้องและเหมาะสมกับวัตถุประสงค์ของการวิจัย

(4) การจำลองข้อมูลให้สอดคล้องกับเงื่อนไขของโมเดลการวิเคราะห์ข้อสอบ

(5) การประมาณค่าพารามิเตอร์โดยใช้ข้อมูลจากการจำลอง

(6) การเปรียบเทียบผลที่ได้จากการประมาณค่า โดยใช้ค่าสถิติต่าง ๆได้ เช่น ค่ามัธยฐาน ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า เป็นต้น

(7) การกำหนดจำนวนรอบ (Replicated R time)

(8) การคำนวณค่าสถิติของผลการจำลองข้อมูลจำนวน R รอบทั้งสถิติเชิงบรรยาย เพื่อพิจารณาการแจกแจงและสถิติเชิงอ้างอิงสำหรับการทดสอบความแตกต่าง

นอกจากนี้ในปัจจุบันมีการประยุกต์ใช้เทคนิคมอนติคาร์โลสำหรับการวิเคราะห์ข้อสอบกันอย่างแพร่หลาย Harwell และ คณะ (1996) จึงให้ข้อควรระวัง สำหรับงานวิจัยที่จะใช้การวิเคราะห์ข้อสอบด้วยการใช้เทคนิคมอนติคาร์โลที่สำคัญก่อนตัดสินใจใช้เทคนิคมอนติคาร์โลสำหรับการวิเคราะห์ข้อสอบ สามารถสรุปได้ดังนี้ (M. Harwell et al., 1996)

1. ปัญหาในการวิจัยสามารถแก้ได้ด้วยการวิเคราะห์ด้วยเทคนิคมอนติคาร์โลจริงหรือไม่

2. การศึกษาในประเด็นนั้นสามารถขยายองค์ความรู้เดิมได้หรือไม่

3. ในการออกแบบการทดลองและการวิเคราะห์เหมาะสมที่จะใช้เทคนิคมอนติคาร์โลมาศึกษาหรือไม่

4. การวิเคราะห์ข้อมูลควรจะนำโปรแกรมที่มีอยู่เดิมมาใช้หรือปรับปรุงโปรแกรมใหม่เพื่อให้เหมาะสมกับสภาพปัญหาการวิจัยของเราหรือไม่

5. ผลลัพธ์ในการวิเคราะห์ข้อมูลในงานวิจัยขึ้นอยู่กับค่าเริ่มต้นสำหรับวิธีการประมาณค่าพารามิเตอร์หรือไม่

6. ข้อตกลงเบื้องต้นของการแจกแจงลักษณะตัวแปรอิสระและค่าต่างๆ ที่กำหนดในการจำลองข้อมูลตรงกับสภาพความเป็นจริงหรือไม

3.3.2 ขั้นตอนในการวิเคราะห์ข้อสอบด้วยเทคนิคมอนติคาร์โล

การวิเคราะห์ข้อสอบด้วยเทคนิคมอนติคาร์โลมีขั้นตอนที่สำคัญสามารถสรุปได้ 4 ขั้นตอน สำหรับการนำไปสู่การปฏิบัติการวิเคราะห์ ขั้นตอนแรก ได้แก่ การกำหนดปัญหาใน การศึกษาวิจัย ขั้นตอนต่อมา เป็นขั้นตอนของการออกแบบการทดลอง ซึ่งรวมถึงการระบุตัวแปรอิสระและตัวแปรตาม การออกแบบการวิจัยเชิงทดลอง จำนวนรอบในการคำนวณ และการเลือกโมเดลการวิเคราะห์ข้อสอบ ขั้นตอนที่สามเป็นการเขียนและระบุโปรแกรมในการจำลองข้อมูลและการประมาณค่าพารามิเตอร์ และขั้นตอนสุดท้ายเป็นขั้นตอนเกี่ยวกับการวิเคราะห์ผลจากการจำลองข้อมูล โดยในแต่ละขั้นตอนมีรายละเอียดดังนี้ (Naylor และคณะ, 1968 อ้างถึงใน ชนะศึก นิชานนท์, 2553)

(1) การกำหนดปัญหา ขั้นตอนนี้นับเป็นขั้นตอนที่สำคัญของกระบวนการวิจัย การวิเคราะห์ข้อสอบด้วยเทคนิคมอนติคาร์โลก็เช่นเดียวกัน เริ่มแรกนักวิจัยจะต้องมีการกำหนดปัญหาและข้อคำถามการวิจัยก่อน หลังจากนั้นจึงตั้งสมมติฐานการทดสอบ และมีการวัดผลกระทบจากเงื่อนไขต่างๆ จากการจำลองข้อมูล โดยทั่วไปการกำหนดปัญหาการวิจัยมักจะมาจากการทบทวนเอกสารและรายงานการวิจัยที่เกี่ยวข้องกับเรื่องที่เราสนใจศึกษา

(2) การออกแบบการศึกษาด้วยเทคนิคมอนติคาร์โล การออกแบบการศึกษานั้นจะต้องออกแบบให้สามารถตอบข้อคำถามของการวิจัยและสมมติฐานการวิจัยได้ โดยการออกแบบจะต้องมีการออกแบบทั้งตัวแปรต้นหรือตัวเป็นที่เป็นสาเหตุและตัวแปรตามซึ่งเป็นผลกระทบที่เกิดขึ้นจากตัวแปรต้น นอกจากนี้ยังต้องมีการประเมินผลทั้งในเรื่องของความตรงภายในและความตรงภายนอกซึ่งการออกแบบการศึกษาด้วยด้วยเทคนิคมอนติคาร์โลมีลักษณะคล้ายคลึงกับการศึกษาจากข้อมูลเชิงประจักษ์ ดังนี้

(2.1) การกำหนดและระบุค่าของตัวแปรต้น โดยมีการกำหนดและระบุค่าของตัวแปร รวมทั้งเงื่อนไขในการจำลองข้อมูลให้สอดคล้องกับปัญหาและคำถามของการวิจัย โดยค่าของตัวแปรซึ่งคือว่าเป็นค่าคงที่ ซึ่งได้แก่ ขนาดกลุ่มตัวอย่าง (n) ความยาวของข้อสอบ (L) ความแปรปรวนของตัวแปรอิสระ (S^2) ซึ่งค่าของตัวแปรเหล่านี้ควรจะเกิดจากคำถามวิจัย นอกจากนี้ค่าของพารามิเตอร์ในโมเดลยังควรเป็นตัวแทนของตัวแปรอิสระ เนื่องจากในการศึกษาเทคนิคมอนติคาร์โล ต้องมีการพิจารณาว่าค่าพารามิเตอร์เหล่านี้เป็นค่าของตัวแปรคงที่ (fixed effect) หรือเป็นค่าของตัวแปรสุ่ม (random effect) เนื่องจากถ้ามีการสุ่มค่าอำนาจจำแนกและค่าความยากง่ายมาใช้ในการศึกษา จะทำให้โมเดลในการศึกษาเป็นโมเดลแบบสุ่ม ซึ่งสามารถอ้างอิงไปยังประชากร แต่ถ้าโมเดลไม่ได้มีการสุ่มขึ้นมาใช้ในการศึกษา โมเดลนั้นจะไม่สามารถอ้างอิงกลับไปยังประชากรได้

(2.2) การเลือกแบบการทดลอง โดยทั่วไปตัวแปรอิสระมักจะเป็นตัวกำหนดแบบการทดลองที่เหมาะสม เช่น ถ้าจำนวนตัวแปรอิสระและระดับค่าของตัวแปรน้อย การใช้แบบการทดลองแบบแฟกตอเรียลจะมีความเหมาะสมกว่าแบบการทดลองอื่น ในการศึกษาด้วยเทคนิคมอนติคาร์โลมักจะใช้แบบการทดลองโดยยึดเป้าหมายและวัตถุประสงค์ในการศึกษาเป็นหลัก ซึ่งการเลือกแบบทดลองอย่างระมัดระวังจะช่วยในการวางแผนการวิเคราะห์ผลลัพธ์ได้อย่างถูกต้อง (Lewis และ Ovar, 1989 อ้างถึงใน Harwell และคณะ, 1996) ในงานวิจัยของ Harwell และ Janoskey (1991) มีตัวแปรจัดกระทำหรือตัวแปรต้นเป็นขนาดกลุ่มตัวอย่าง ความยาวของแบบสอบ และความแปรปรวนของการแจกแจงค่าอำนาจจำแนก ใช้การออกแบบการทดลองแบบ แฟกตอเรียลระหว่างกลุ่มตัวอย่างแบบสมบูรณ์ (completely between-subjects factorial design) นอกจากนี้ยังมีงานวิจัยอีกชิ้นหนึ่งซึ่งถือว่าเป็นตัวอย่างที่ดีเกี่ยวกับการออกแบบการทดลอง คือ งานวิจัยของ Yen (1987 อ้างถึงใน Harwel และคณะ, 1996) ซึ่งได้เปรียบเทียบการศึกษาจากการใช้โปรแกรมการวิเคราะห์ข้อสอบระหว่างโปแกรม BILOG และโปรแกรม LOGIST โดยมีเงื่อนไขด้านความยาวของข้อสอบและลักษณะการกระจายของค่าพารามิเตอร์ของผู้สอบ ออกแบบการทดลองแบบแฟกตอเรียล โดยเรื่องของความยาวของข้อสอบและลักษณะการกระจายของค่าพารามิเตอร์ของผู้สอบ ได้ออกแบบ

การทดลองแฟกตอเรียลแบบ between-subjects factor และส่วนในการเปรียบเทียบระหว่างโปรแกรมคอมพิวเตอร์ทั้งสองได้ออกแบบการทดลองแฟกตอเรียลแบบ within-subjects factor

(2.3) การเลือกตัวแปรตาม การเลือกตัวแปรตามนั้นไม่เพียงแต่จะต้องสอดคล้องกับคำถามในการวิจัยแต่จะต้องเลือกตัวแปรตามที่มีความไวต่อตัวแปรจัดกระทำหรือตัวแปรต้น เนื่องจากการเลือกตัวแปรที่มีความไวและควรใช้ประโยชน์ได้ถ้ามีการแปลงข้อมูลเป็นรูปแบบอื่น เช่น การหาค่า RMSD สามารถแปลงค่าเพื่อให้มีการแจกแจงแบบปกติ ทำให้สามารถนำไปสรุปอ้างอิงได้ทั่วไป นอกจากนี้ถ้าเป็นเรื่องเกี่ยวกับการศึกษาเปรียบเทียบวิธีการในการศึกษา IRT นักวิจัยก็สามารถใช้คุณลักษณะของแบบสอบ เช่น ความเป็นเอกมิติ การทำหน้าที่ต่างกันของข้อสอบหรือผู้สอบ เป็นตัวแปรตามในการศึกษาถึงผลกระทบของตัวแปรอิสระได้อีกด้วย สำหรับค่าความสัมพันธ์ของค่าจริงกับค่าประมาณ ก็สามารถใช้ให้เป็นตัวแปรตามในการใช้ เทคนิคมอนติคาร์โล เนื่องจากค่าสัมพันธ์นั้นสามารถใช้เมตริกที่ต่างกันหาความสัมพันธ์ของตัวแปรอิสระและตัวแปรตามได้ เช่น ค่าความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ประมาณกับความคลาดเคลื่อนมาตรฐาน ส่วนข้อเสียก็คือ ความสัมพันธ์เหล่านี้สะท้อนความสัมพันธ์เฉพาะอันดับของตัวแปรและแสดงอิทธิพลของตัวแปรอิสระเท่านั้น เช่น ค่าความสัมพันธ์ระหว่างค่าอำนาจจำแนกที่แท้จริงกับค่าที่ประมาณ มีค่าเท่ากับ 0.9 ซึ่งหมายความว่าโดยค่าเฉลี่ยของค่าอำนาจจำแนกที่แท้จริงนั้น อาจจะสูงกว่าค่าเฉลี่ยของค่าอำนาจจำแนกที่ประมาณได้ แต่ไม่รับรองว่าค่าอำนาจจำแนกที่แท้จริงนั้นกับค่าอำนาจจำแนกที่ ประมาณจะใกล้เคียงกันหรือดีกว่ามากน้อยเพียงใด เช่น 0.8 กับ 0.9

(2.4) การกำหนดจำนวนรอบ การกำหนดจำนวนรอบในการศึกษาด้วยเทคนิคมอนติคาร์โลเปรียบเทียบได้กับการกำหนดขนาดกลุ่มตัวอย่าง โดยมีเกณฑ์ที่ใช้ในการกำหนดประยุกต์ใช้มาจากการกำหนดขนาดกลุ่มตัวอย่างสำหรับในการศึกษาจากข้อมูลเชิงประจักษ์ ในการศึกษาการวิเคราะห์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ จำนวนรอบจะขึ้นอยู่กับวัตถุประสงค์ในการศึกษา ด้วยเทคนิคมอนติคาร์โล โดยพิจารณาได้จากความต้องการในลดค่าความแปรปรวนของการสุ่ม ตัวอย่างในการประมาณค่าพารามิเตอร์ และความต้องของการทดสอบสถิติของผลการจำลองข้อมูล ว่าอำนาจในการตรวจสอบผลกระทบที่สนใจเพียงพอหรือไม่ จำนวนรอบมีอิทธิพลโดยตรงกับความแม่นยำในการประมาณค่าพารามิเตอร์ ถ้ากลุ่มตัวอย่างขนาดใหญ่ (มีจำนวนรอบมาก) จะให้การประมาณค่าพารามิเตอร์ด้วยความแปรปรวนของการสุ่มตัวอย่างน้อย ดังนั้นถ้านักวิจัยไม่กำหนดจำนวนรอบหรือกำหนดจำนวนรอบน้อยจะทำให้ความแปรปรวนของการสุ่มมีมากเพียงพอที่จะให้ทำให้การประมาณค่าพารามิเตอร์มีความลำเอียงมาก ซึ่งจะส่งผลให้ความเที่ยงและความน่าเชื่อถือของผลการวิจัยที่ได้ต่ำมาก

เทคนิคในการลดความแปรปรวนในการประมาณค่า คือ การเพิ่มจำนวนรอบซึ่งจะทำให้ได้ค่าที่คงที่และน่าเชื่อถือได้มากกว่า เนื่องจากสามารถเปรียบเทียบค่าพารามิเตอร์ระหว่าง รอบได้ ข้อดีอีกประการหนึ่ง คือ จำนวนรอบสะท้อนความเบี่ยงเบนของค่าประมาณ ซึ่งถ้าเป็นความเบี่ยงเบนระหว่างเงื่อนไขมีค่าน้อย แสดงว่าตัวแปรอิสระส่งผลต่อตัวแปรตามน้อย โดยมีสมการใน การคำนวณความเบี่ยงเบนของค่าพารามิเตอร์ที่แท้จริงและค่าประมาณได้จาก ค่า RMSD (Root Mean Square Deviation) ซึ่งเป็นค่าการถอดรากของค่าเฉลี่ยกำลังสองของ ความแตกต่างระหว่างค่าที่ ประมาณได้กับค่าพารามิเตอร์ที่แท้จริงสามารถเขียนเป็นสูตรทางคณิตศาสตร์ได้ดังนี้

$$RMSD = \left[\frac{\sum_{i=1}^{n} \left(\hat{a} - a_i \right)^2}{n} \right]^{1/2} \qquad \text{.....(2.52)}$$

เมื่อ $\hat{a}$ เป็นค่าพารามิเตอร์ที่ได้จากการประมาณค่า

a_i เป็นค่าพารามิเตอร์แท้จริง

n เป็น จำนวนพารามิเตอร์

นอกจากนี้เมื่อพิจารณาความแปรปรวนของการประมาณค่าจากจำนวนรอบ (the variance of the estimates across replications) และฟังก์ชันความแปรปรวนความคลาดเคลื่อนของข้อมูลเชิงประจักษ์ (an empirical error variance) โดยสามารถสรุปเป็นสมการได้ดังนี้

$$\frac{\sum_{r=1}^{R} \left(\hat{a}_{ir} - a_i \right)^2}{R} = \left(\bar{\hat{a}}_{ir} - a_i \right) + \frac{\sum_{r=1}^{R} \left(\hat{a}_{ir} - \bar{\hat{a}}_i \right)^2}{R} \qquad \text{.....(2.53)}$$

จากสมการข้างต้น ถ้าค่าดังกล่าวมีค่าน้อย แสดงให้ว่า การประมาณค่ามีความคงที่นั้น ก็คือการประมาณค่ามีความเที่ยง ในทางตรงกันข้ามถ้าค่าดังกล่าวมีค่ามาก นักวิจัยต้อง ระมัดระวังในการนำค่านั้นไปใช้ในการวิเคราะห์เนื่องจากค่าดังกล่าวไม่คงที่หรือการประมาณค่านั้นไม่มี ความเที่ยง

เมื่อพิจารณาในเรื่องของจำนวนรอบกับอำนาจการทดสอบ (power) พบว่า จำนวน รอบนั้นมีความสำคัญกับอำนาจการทดสอบ โดยนักวิจัยต้องเลือกจำนวนรอบให้มีขนาดมากพอกับ อำนาจที่ต้องการสำหรับการใช้สถิติตัวนั้น ๆ ในการทดสอบจากการศึกษาของ Stone (1993 อ้างถึง ใน Harwell และคณะ, 1996) โดยได้ศึกษาความสัมพันธ์ระหว่างจำนวนรอบกับอำนาจการทดสอบเป็น 2 ขั้นตอน โดยขั้นตอนแรกใช้เทคนิคมอนติคาร์โล ในการศึกษาผลของจำนวนผู้สอบ N (N=250, 500, 1000) , จำนวนข้อสอบ (L=10, 20, 30) และการแจกแจงของความสามารถ (θ) (D= Normal, Skewed, Platykurtic) ทดลองโดยใช้แบบแฟคตอเรียล โดยมีค่า RMSD เป็นตัวแปรตาม และใช้

โมเดล 2PL ในการระบุค่าอำนาจจำแนกและค่าความยากในแต่ละข้อ กำหนดให้ R=10 (10 ชุด) ในแต่ละเงื่อนไข แสดง RMSD ด้วยกราฟ และทดสอบด้วย ANOVA เพื่อดูอิทธิพลของตัวแปรอิสระ ค่าอิทธิพลสูงสุดแสดงด้วยค่า η^2 (correlation ratio=η^2) จะเหมาะสมเนื่องจากสอดคล้องกับข้อตกลงเบื้องต้นด้านการแจกแจงเป็นโค้งปกติ

ในขั้นตอนที่ 2 η_s^2 ได้จากการศึกษาเทคนิคมอนติคาร์โล จะใช้ประมาณค่าอำนาจการทดสอบของ ANOVA เพื่อตรวจสอบอิทธิพลอันเกิดจากจำนวนรอบที่แตกต่างกัน (R=10, 25, 50, 100) โดยใช้ η_s^2 เป็นตัวประมาณขนาดอิทธิพล โปรแกรม STAT-POWER จะใช้ประมาณค่าอำนาจการทดสอบที่ระดับ .05 และองศาอิสระเท่ากับ 27 และ η^2 กระบวนการนี้จะทำให้เกิดการเปลี่ยนแปลงอำนาจการทดสอบของ ANOVA F-Test ความสัมพันธ์ระหว่างจำนวนรอบกับอำนาจการทดสอบ โดยอำนาจการทดสอบจะเปลี่ยนแปลง เมื่อจำนวนรอบเปลี่ยนแปลงไป กล่าวคือ จำนวนรอบยิ่งมากอำนาจการทดสอบก็จะมีค่าสูงขึ้น

การเพิ่มจำนวนรอบจนถึง 500 อาจไม่จำเป็น เมื่อจำนวนผู้สอบ (N) และจำนวน ข้อสอบ (L) มีขนาดใหญ่พออาจน้อยกว่า 100 รอบก็ได้ แต่จากการตรวจสอบจำนวนรอบ (R) อาจจะจำเป็นเพื่อความน่าเชื่อถือของการศึกษาในประเด็นดังต่อไปนี้ เช่น การตรวจสอบการแจกแจงการสุ่ม การวิเคราะห์ข้อสอบที่มีความแปรปรวนขนาดใหญ่ การศึกษาอิทธิพลของเงื่อนไขที่มี ความสลับซับซ้อน และการศึกษาค่าพารามิเตอร์ที่เป็นค่าสุดโต่ง

ดังนั้นจะเห็นได้ว่าจำนวนรอบมีความจำเป็นต่อความเที่ยงในการตรวจสอบผลกระทบของผลการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลมากในลักษณะงานวิจัยใน 4 ประเภท ได้แก่ ประเภทแรก งานวิจัยที่ให้ความสำคัญกับลักษณะการกระจายของการสุ่มตัวอย่างจากข้อมูลเชิงประจักษ์ เช่น งานวิจัยที่มีวัตถุประสงค์เพื่อตรวจสอบคุณสมบัติของค่าสถิติหรือการทดสอบนัยสำคัญทางสถิติ ประการที่สอง งานวิจัยที่สนใจศึกษาค่ากลางของการวิจัยระดับข้อสอบที่มีค่าความแปรปรวนของกลุ่มตัวอย่างมาก ประการที่สามงานวิจัยที่มีเป้าหมายในการศึกษาผลกระทบที่เพิ่มขึ้นของบริบทที่ซับซ้อน เช่น ผลกระทบที่เกิดขึ้นจากปฏิสัมพันธ์กับอิทธิพลหลัก ประการสุดท้าย โมเดลที่มีค่าพารามิเตอร์ที่มีค่าสุดโต่งมาก ทั้งนี้จำนวนรอบขั้นต่ำสำหรับการศึกษา การวิเคราะห์ ด้วยทฤษฎีการตอบสนองควรในการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลควรมีค่าไม่ต่ำกว่า 25 รอบ ถือเป็นเงื่อนไขที่อาจนำไปใช้ศึกษาในทฤษฎีการตอบสนองข้อสอบด้วยเทคนิคมอนติคาร์โลได้

(3) การเขียนและการระบุโปรแกรมในการจำลองข้อมูล การเลือกโปรแกรม คอมพิวเตอร์นักวิจัยอาจใช้หลายโปรแกรมในการจำลองข้อมูล รวมทั้งการวิเคราะห์ผลลัพธ์ ในแต่ละขั้นตอนจะต้องมีการประเมินความถูกต้องแม่นยำ ใน 3 ประเด็นดังนี้ ประการแรก ค่าของข้อมูลจาก การจำลองมีคุณภาพเพียงใดเมื่อเทียบกับข้อมูลที่มีการศึกษาในอดีต กล่าวคือ ควรมีการตรวจสอบ ความ

ตรงของค่าที่ได้จากโมเดลหรือกระบวนการจำลองข้อมูล ประการที่สอง โมเดลการจำลอง ข้อมูลสามารถทำนายปรากฏการณ์ในอนาคตได้ถูกต้องแม่นยำเพียงใด ประเด็นสุดท้าย คือ ผลจากการจำลองข้อมูลนั้นเมื่อนำไปศึกษากับข้อมูลเชิงประจักษ์จะสามารถทำนายได้เมื่อเวลาผ่านไปหรือไม่ ซึ่งมีรายละเอียดในการเขียนและระบุโปรแกรมในการจำลองดังนี้

(3.1) การจำลองคำตอบ การจำลองคำตอบเริ่มด้วยการกำหนดค่าเริ่มต้น (Seed) ให้กับตัวเลขสุ่ม ซึ่งจะแปลงเป็นค่าความน่าจะเป็นผู้ตอบในการตอบคำถามได้ถูกต้องและต่อจากนั้น จึงแปลเป็นคำตอบของผู้สอบในการตอบข้อสอบแบบไม่ต่อเนื่อง (0, 1) โดยทั่วไปในขั้นตอนนี้จะเป็นหน้าที่ของคอมพิวเตอร์เป็นผู้ดำเนินการ

(3.2) ตัวเลือกค่าเริ่มต้น ซึ่งผู้จำลองข้อมูลจะเป็นผู้กำหนดค่าเริ่มต้นเอง โดยใช้เติมในช่องว่าง (Prompted) ประโยชน์ก็คือ ง่ายต่อการจำลองคำตอบในข้อต่อ ๆ มา ประเด็นที่สำคัญ คือ เป็นค่าที่สัมพันธ์กับความคลาดเคลื่อนในการสุ่ม ซึ่งหลีกเลี่ยงได้ยากในการจำลองข้อมูล เทคนิคอย่างหนึ่งที่จะลดความแปรผัน คือ ใช้พารามิเตอร์ข้อสอบและค่าเริ่มต้นร่วมกันทุกครั้ง เมื่อมีการจำลองข้อมูล เช่น การจำลองข้อสอบ 20 ข้อและ 30 ข้อ ค่าเริ่มต้นที่ใช้ในการจำลองข้อมูล 20 ข้อ ควรจะเป็นค่าเริ่มต้นเดียวกันกับเมื่อจำลองข้อมูล 30 ข้อ เทคนิคอีกประการหนึ่ง คือ การจำลองประชากรข้อมูลคำตอบจำนวนมาก แล้วสุ่มคำตอบมาจากประชากรที่จำลองขึ้นนั้นหลาย ๆ รอบ มากกว่าการใช้ค่าเริ่มต้นหลายตัว เพื่อจะจำลองชุดข้อมูลให้ได้ตามต้องการ การใช้โมเดลพารามิเตอร์ที่ต่างกันในการจำลองข้อสอบ 20 ข้อ 30 ข้อ รวมทั้งค่าเริ่มต้นที่ต่างกันจะทำให้คำตอบมีความเป็นอิสระแก่กันมากขึ้น แต่อาจจะเกิดความคลาดเคลื่อนมากกว่าเมื่อเทียบกับการใช้โมเดลพารามิเตอร์และค่าเริ่มต้นแบบเดียวกัน วิธีที่ควรใช้คือ การจำลองข้อมูลทุกชุดควรใช้โมเดลพารามิเตอร์ต่างกันแต่ใช้ค่าเริ่มต้นร่วมกัน

(3.3) การจำลองตัวเลขสุ่ม จะใช้ตัวเลขสุ่มที่มีการแจกแจงแบบยูนิฟอร์ม (Uniform distribution) เป็นส่วนใหญ่และใช้วิธีการจำลองข้อมูลแบบ Congruential generators (Cewis และ Ovaw, อ้างถึงใน Harwell และคณะ, 1996) วิธี Congruential generators จะใช้โมเดลพีชคณิตใน การจำลองตัวเลขที่สุ่มขึ้นมากับตัวเลขสุ่มที่ผ่านมา ตัวเลขสุ่มจะเริ่มจำนวนจาก 0....m ซึ่งจะสุ่มโดยโปรแกรม โดยจะวิ่งเป็นวงจรที่เรียกว่า “ความยาวรอบ” (period) เมื่อครบรอบก็จะวนกลับมาใช้เลขเดิมอีก จากการศึกษาในอดีต พบว่า การใช้วิธี Congruential generators มีช่วงความยาวรอบเพียงพอกว่าจะเกิดข้อมูลซ้ำ การแจกแจงปกติมาตรฐานจะใช้มากในการใช้เทคนิคมอนติคาร์โล โดยการแปลงข้อมูลจาก Uniform (0, 1) เป็น $Normal \sim N(0,1)$ วิธีที่ใช้แปลงคือ วิธีของ Box – Muller และ Marsaglia

หลักในการเลือกโปรแกรมเพื่อใช้ในการจำลองข้อมูล ประกอบด้วย 4 อย่าง คือ วิธีการควรง่ายต่อการทำความเข้าใจและเขียนโปรแกรม โปรแกรมควรมีความกะทัดรัด รหัสปลายทางควรมีความสมเหตุสมผล (คำตอบ 0, 1) และขั้นตอนการคำนวณ ควรใช้ตัวเลขสุ่มเทียม (Pseudo-random numbers) (Ripley, 1988)

ความถูกต้องของตัวเลขสุ่มขึ้นอยู่กับตัวจำลองเลขสุ่มและเครื่องคอมพิวเตอร์นักวิจัยควรตรวจสอบตัวจำลองและเอกสารด้านสมรรถนะ แม้บางตัวอาจเป็นที่เชื่อถือได้แต่ก็ควรตรวจสอบว่าสามารถนำไปใช้ตรงกับปัญหาเฉพาะหรือไม่ อาจเป็นการทดสอบและตรวจเอกสารซึ่งจะเกี่ยวข้องกับความเชื่อถือได้ของโปรแกรม การตรวจสอบความพอเพียงของตัวจำลองข้อมูลด้วยกระบวนการทดสอบความเหมาะสมด้วยสถิติ Kolmogorov Smirnov and Coraner – van – Mises goodness - of – fit การทดสอบลำดับขึ้นลง การทดสอบช่วงว่างและค่าสูงสุด และการทดสอบการสุ่ม

(3.4) การแปลงตัวเลขสุ่มเป็นคำตอบ ในเบื้องต้น เวกเตอร์ของ N (จำนวนผู้สอบ), θ (ความสามารถ) ถูกสุ่มจากการแจกแจงที่กำหนดจากกลุ่มผู้สอบ ซึ่งค่าความสามารถมักจะได้จาก การสุ่มกลุ่มตัวอย่างที่มีการแจกแจงแบบปกติ และเป็นโมเดล IRT แบบเอกมิติหรือจากการแจกแจง ปกติหลายตัวแปรที่มีค่าความสัมพันธ์ระหว่างตัวแปรในกรณีนี้ศึกษาการวัดแบบหลายมิติ เช่น การสุ่มความน่าจะเป็นของคำตอบที่ตอบแบบ (0, 1) จะได้จากสมการ ดังนี้

$$P_i(\theta) = c_i + \frac{1-c_i}{1+e^{-Da_i(\theta-b_i)}} \qquad \text{.....(2.54)}$$

$$Q_i = 1 - P_i \qquad \text{.....(2.55)}$$

ค่าความน่าจะเป็นของคำตอบ (P_i) จะถูกแปลงไปเป็นคำตอบ 0, 1 โดยการเปรียบเทียบค่าความน่าจะเป็นจากการสุ่ม การแจกแจงแบบ Uniform ถ้าความน่าจะเป็นในการตอบถูกของผู้สอบสูงกว่าหรือเท่ากับตัวเลขสุ่ม คำตอบในข้อนั้นจะเป็น 1 แต่ถ้าต่ำกว่าตัวเลขสุ่มจะเป็น 0 ในกรณีเป็นคำตอบหลายค่า ถ้าตัวเลขสุ่มตกอยู่ในช่วงใดก็จะเป็นคำตอบ K และ K+1 ในกระบวนการทั้งหมด จะกระทาซ้ำด้วยตัวเลขสุ่มที่แตกต่างกันในแต่ละข้อและผู้สอบทั้งหมด

(3.5) การประมาณค่าพารามิเตอร์ในโมเดล ผู้วิจัยอาจใช้โปรแกรมสำเร็จรูป เช่น BILOG หรือ MULTILOG หรือการสร้างโปรแกรมด้วยตนเองก็ได้ ประเด็นสำคัญ 2 ประเด็นในการเขียนและการเลือกโปรแกรมวิเคราะห์ คือ การกำหนดค่าเริ่มต้น (Starting Value) และ การแก้ปัญหา เมตริกไม่ลู่เข้า (non-convergent solutions) ในกระบวนการทั้งหมดของการประมาณค่าจะเกี่ยวกับขั้นตอนการคำนวณทวนซ้ำและการกำหนดค่าเริ่มต้น โดยปกติคอมพิวเตอร์จะกำหนดให้ แต่นักวิจัยส่วนใหญ่จะเป็นผู้กำหนดเอง

หากนักวิจัยพบปัญหาการไม่ลู่เข้า ในการศึกษาแบบมอนติคาร์โล นักวิจัยสามารถดำเนินการแยกได้เป็น 3 กรณี ได้แก่ ประการแรก นักวิจัยไม่ต้องให้สนใจความไม่ลู่เข้านั้น แต่ใช้ค่าประมาณจากจำนวนครั้งของการคำนวณทวนซ้ำที่มากที่สุด ประการที่สองนักวิจัยควรแยกการประมาณค่าของค่าสถิติสรุปรวม ได้แก่ RMSD และประการที่สามนักวิจัยควรใช้วิธีการคำนวณอื่น ๆ เช่น วิธีของเบส์ (Bayesian) เพื่อบังคับค่าพารามิเตอร์ โดยควรคำนึงถึงค่าพารามิเตอร์ที่ประมาณค่ากับค่าพารามิเตอร์ที่แท้จริงให้มีค่าใกล้เคียงกันด้วย มิฉะนั้นแล้วค่าที่ได้จะเป็นค่าที่ลำเอียง

ปัญหาสำคัญของการกำหนดค่าเริ่มต้นและการแก้ปัญหาการไม่ลู่เข้า มักพบในข้อมูลจำนวนน้อย (N=200, 40 ข้อ) และในโมเดล IRT ที่ซับซ้อนมากขึ้น การแก้ปัญหาคือ การพัฒนาโปรแกรมเพื่อใช้ประมาณค่า วิธีการของเบส์ (Bayesian) สามารถช่วยได้ โดยการพิจารณา อย่างพิถีพิถันกับการแจกแจงของข้อมูลและการแจกแจงของค่าพารามิเตอร์

(4) การวิเคราะห์ผล การวิเคราะห์ผลจะตั้งอยู่บนพื้นฐานของคำถามวิจัย การออกแบบการทดสอบสมมติฐานทางสถิติ กระบวนการวิเคราะห์ โดยปกติการวิเคราะห์ผลจากการใช้เทคนิคมอนติคาร์โล ประกอบด้วย การใช้ตารางสรุปผลรวม สถิติเชิงบรรยายเบื้องต้นหรือการนำเสนอด้วยกราฟ แผนภูมิ การดูผลกระทบจากตัวแปรอิสระ อาจต้องใช้สถิติเชิงอ้างอิงปัญหาของการวิเคราะห์คือการ มีค่าต่าง ๆ กันจำนวนมาก เมื่อต้องการรายงานผล Harwell และ Janosky (1991) เสนอว่าควรจะมีการใช้ทั้งเชิงบรรยายและเชิงอ้างอิง เพื่อเป็นการเพิ่มโอกาสในการตรวจสอบข้อมูล ซึ่งจะทำให้มีความเชื่อมั่นมากขึ้น

การวิเคราะห์ผลด้วยการใช้สถิติเชิงสรุปอ้างอิงสามารถใช้ได้หลายวิธี แต่มักเป็นการวิเคราะห์ถดถอยพหุคูณและการวิเคราะห์ความแปรปรวน ถ้าตัวแปรต้นเป็นตัวแปรระดับนามบัญญัติ การวิเคราะห์ความแปรปรวนจะดีกว่า แต่ถ้าเป็นตัวแปรระดับช่วงชั้น การใช้การวิเคราะห์ถดถอยพหุคูณจะดีกว่า แต่ส่วนใหญ่แล้วในทฤษฎีการตอบสนองข้อสอบจะใช้ การวิเคราะห์ถดถอยพหุคูณ เนื่องจากตัวแปรส่วนใหญ่จะเป็นระดับช่วงชั้น หรืออัตราส่วน ในโมเดลการวิเคราะห์ถดถอยพหุคูณประชากรสามารถเขียนแทนด้วย

$$\tau = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_T X_T \qquad \text{.....(2.56)}$$

และ $$\hat{\tau}_s = \tau_s + \varepsilon_s \qquad \text{.....(2.57)}$$

τ มักจะเป็นผลลัพธ์ เช่น ค่าของ RMSD ซึ่งจะได้รับผลจากตัวแปรทำนาย X_S (s=0,..., T)

β_0 เป็นค่าเฉลี่ย, β_T เป็นสัมประสิทธิ์การถดถอยพหุคูณของประชากร , ε_s เป็นความคลาดเคลื่อน และ $\hat{\tau}_s$ เป็นค่าผลลัพธ์สังเกตได้ โมเดลประมาณสำหรับกลุ่มตัวอย่าง คือ

$$\tau_s = b_0 + b_1 x_1 + b_2 x_2 + ... + b_T x_T \qquad \text{.....(2.58)}$$

ข้อตกลงของสมการถดถอย กล่าวว่า ความคลาดเคลื่อนจะต้องเป็นอิสระจากกัน และในการสร้างข้อมูลก็จะต้องมีการตรวจสอบความเป็นอิสระดังกล่าว ความเป็นอิสระของ τ_s อาจยอมรับได้จากความถูกต้องของตัวจำลองข้อมูล และการตรวจสอบการแจกแจงแบบโค้งปกติ หรือ อาจใช้วิธีการตรวจสอบด้วยความแข็งแกร่งของการแจกแจงปกติ จากการทบทวนเอกสาร เช่น การ แปลงค่า RMSD เป็น Log (RMSD) ถ้านำมาใช้ในการคำนวณต้องมีการแจกแจงเป็นโค้งปกติด้วย แต่ถ้าไม่สามารถแปลงค่าได้ให้มีการแจกแจงเป็นโค้งปกติ อาจต้องนำสถิติ non-parametric มาใช้แทน

Harwell และ Janosky (1991) ได้ศึกษาการวิเคราะห์การถดถอยพหุคูณและการวิเคราะห์ ANOVA เพื่อวิเคราะห์ผลลัพธ์จากการใช้เทคนิค MC โดยการจำลองข้อมูลแบบ 0, 1 ด้วย การวิเคราะห์โมเดล 2 PL ประมาณค่าพารามิเตอร์โดยใช้โปรแกรม BILOG เปรียบเทียบค่าประมาณ ค่า a และ b ในแต่ละข้อกับค่าแท้จริงโดยใช้ RMSD ผลจากการศึกษาสรุปได้ว่า ขนาดกลุ่มตัวอย่าง (N) ของข้อสอบจำนวน 15 และ 25 ข้อ จะมีผลต่อความถูกต้องในการประมาณค่าพารามิเตอร์ และพบว่าการใช้ log (RMSD) สามารถทำให้ตัวแปรตามสอดคล้องกับข้อตกลงเบื้องต้นทางสถิติด้าน การแจกแจงแบบปกติ โดยถ้าตัวแปรต้นเป็นตัวแปรเมทริกซ์ จะมีความเหมาะสมที่จะวิเคราะห์ด้วย การถดถอยพหุคูณ นอกจากนี้ยังทดสอบแฟกตอเรียล ANOVA แบบสุ่มสมบูรณ์ซึ่งให้ผลคล้ายกับ การวิเคราะห์ด้วยการถดถอยพหุคูณ

3.4 ประสิทธิภาพของการจำลองข้อมูล

การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเป็นการจำลองข้อมูลโดยการใช้ตัวเลขสุ่ม (Random Numbers) โดยการพิจารณาว่าวิธีการจำลองข้อมูลนั้นดีหรือไม่ สามารถพิจารณาได้ 3 ประการ ได้แก่ ประการแรก วิธีการสุ่มนั้นจะต้องมีลักษณะของตัวเลขสุ่มที่ถูกสร้างขึ้นมามีการแจกแจงอย่างสม่ำเสมอ (Uniformity) ในช่วง 0 ถึง 1 ประการที่สอง ตัวเลขสุ่มในแต่ละตัวต้องเป็นอิสระซึ่งกันและกัน (Independence) และประการสุดท้าย ตัวเลขสุ่มที่ได้มานั้นต้องมีช่วงยาวก่อนจะเกิดการสุ่มซ้ำใหม่ (มีวัฏจักรยาว) ดังนั้น วิธีการจำลองข้อมูลที่ดี มีประสิทธิภาพ จึงสามารถพิจารณาได้จากตัวเลขสุ่มที่ได้มานั้นต้องมีลักษณะ 3 ประการดังกล่าวข้างต้น จึงจะถือว่าข้อมูลที่ได้ ได้มาจากวิธีการจำลองข้อมูลที่ดี (ศิริรัตน์ วงศ์ประกรณ์กุล, 2539)

3.5 ข้อดีและข้อจำกัดของการศึกษาด้วยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล

การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเป็นเทคนิคที่ใช้สาหรับการแก้ปัญหาในการคำนวณทางคณิตศาสตร์สำหรับสถานการณ์ที่มีกระบวนการที่ซับซ้อนหรือไม่สามารถแก้ไขได้ในเชิงเหตุผล ซึ่งการวิเคราะห์ทั่วไปไม่สามารถดำเนินการได้อย่างครบถ้วนและสมบูรณ์ โดยการจำลองข้อมูลที่แสดงเหตุการณ์เกิดขึ้นทั้งทางบวกและทางลบ โดยการสร้างเลขสุ่ม (random number) มาช่วยในการหา

คำตอบของปัญหาที่ต้องการศึกษา ดังนั้นสถานการณ์ส่วนใหญ่ของงานวิจัยที่เหมาะสมกับการใช้ข้อมูลจำลอง จึงเป็นสถานการณ์ที่ต้องการตรวจสอบการแจกแจงทางสถิติ การเปรียบเทียบตัวประมาณค่าที่อยู่ในสถานการณ์ที่เก็บข้อมูลได้ยาก การศึกษาความแกร่งของสถิติ การเปรียบเทียบขั้นตอนการคำนวณของฟังก์ชัน หรือการประเมินขั้นตอนในการคำนวณอย่างใด อย่างหนึ่ง (Harwell และคณะ, 1996 อ้างถึงใน ชนะศึก นิชานนท์, 2553) สามารถสรุปข้อดีของการศึกษาด้วยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ได้ดังนี้

1) ในงานวิจัยบางเรื่องที่ไม่สามารถเก็บข้อมูลได้โดยตรง การวิจัยประเภทนี้สามารถให้กระบวนการทางคณิตศาสตร์ในการจำลองข้อมูลขึ้นมาสาหรับการวิเคราะห์แทนการใช้ข้อมูลจริงที่ไม่สามารถเก็บรวบรวมข้อมูลในการทดลอง

2) การจำลองข้อมูลสามารถสร้างตัวแบบที่มีความซับซ้อนมากได้ ซึ่งบางครั้งในการเก็บรวมข้อมูลจริงอาจเก็บได้ยากและมีจำนวนข้อมูลน้อย ซึ่งไม่สามารถนำมาวิเคราะห์สถิติในบางประเภทได้

3) การจำลองข้อมูลสามารถกำหนดและจัดกระทำค่าพารามิเตอร์ ซึ่งนำไปใช้ในการศึกษาปัจจัยที่ส่งผลต่อองค์ประกอบต่างๆ ได้ รวมทั้งค่าใช้จ่ายจากการจำลองข้อมูลยังน้อยกว่า การเก็บรวบรวมข้อมูลจริงอีกด้วย

4) การศึกษาวิจัยด้วยการจำลองข้อมูลสามารถกำหนดระยะเวลาที่แน่นอนในการดำเนินการทดลองได้ ซึ่งบางงานวิจัยต้องการใช้ผลงานวิจัยที่เร่งด่วน

5) การจำลองข้อมูลสามารถทำการสุ่มและสร้างตัวแปรเทียมสาหรับใช้ในตัวแบบได้โดยตรง ซึ่งในการเก็บรวบรวมด้วยข้อมูลจริงบางครั้ง บางสถานการณ์ไม่สามารถกระทำได้โดยตรง สำหรับเรื่องข้อจากัดของการศึกษาด้วยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล สามารถสรุปสาระสำคัญได้ดังนี้

(1) ในการศึกษาบางครั้งผลที่ได้จากการศึกษาด้วยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลไม่สามารถครอบคลุมได้ในทุกกรณี เหมือนกับสภาพการณ์ที่เกิดขึ้นจริงได้

(2) สำหรับในงานวิจัยหรือการศึกษาสภาพปัญหาในบางประเภทมีกระบวนการในการสร้างตัวแบบที่สามารถกระทำได้ยาก

(3) การนำผลการวิจัยไปใช้จากการจำลองข้อมูลต้องพิจารณาเงื่อนไขที่ศึกษาว่า เงื่อนไขเหล่านั้นสอดคล้องหรือตรงกับสภาพความเป็นจริงมากน้อยเพียงใด

(4) ผลการจำลองข้อมูลนั้นขึ้นอยู่กับจำนวนรอบและความถูกต้องของตัวเลขที่จำลองได้ ดังนั้นคุณธรรมของนักวิจัยที่จำลองข้อมูลจึงเป็นเรื่องที่สำคัญ ซึ่งนับว่าเป็นเรื่องยากต่อการประเมิน

(5) การเลือกใช้วิธีการที่เหมาะสมสำหรับการจำลองข้อมูล โดยถ้านักวิจัยเลือกวิธีการที่ไม่เหมาะสม อาจจะทำให้ผลสรุปในงานวิจัยผิดพลาดและไม่สามารถนำไปใช้ประโยชน์ได้จริง

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ดังนั้นในการศึกษาด้วยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล จึงต้องพิจารณาด้วยงานวิจัยที่ศึกษามีลักษณะหรือสภาพการณ์อย่างไร เหมาะสมกับการศึกษาด้วยการจำลองข้อมูลเทคนิคมอนติคาร์โลหรือควรจะเก็บจากข้อมูลจริง การวิจัยที่ใช้วิธีการจำลองข้อมูลควรเป็นงานวิจัยที่มีความสลับซับซ้อนไม่สามารถได้จากการเก็บรวบรวมข้อมูลจริง คำถามวิจัยมักเป็นคำถามที่ต้องหาคำตอบได้ด้วยวิธีการทางคณิตศาสตร์ หรือวิธีการทางสถิติ จึงจะทำให้ผลการวิจัยครั้งนั้นได้ผลที่ถูกต้องตามวัตถุประสงค์ของงานวิจัย (Stone, 1993 อ้างถึงใน Harwell, M. และคณะ, 1996)

จากการศึกษาการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล สามารถสรุปได้ว่า การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเป็นเทคนิคที่ใช้สำหรับการแก้ปัญหาในการคำนวณทางคณิตศาสตร์สำหรับสถานการณ์ที่มีกระบวนการที่ซับซ้อนหรือไม่สามารถแก้ไขได้ในเชิงเหตุผล ซึ่งการวิเคราะห์ทั่วไปไม่สามารถดำเนินการได้อย่างครบถ้วนและสมบูรณ์ โดยการจำลองข้อมูลที่แสดงเหตุการณ์เกิดขึ้นทั้งทางบวกและทางลบ โดยการสร้างเลขสุ่ม (random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

การวิเคราะห์ข้อสอบในปัจจุบัน โดยเฉพาะการวิเคราะห์ข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ มีการประยุกต์ใช้เทคนิคมอนติคาร์โลในการวิเคราะห์เป็นจำนวนมาก โดยเฉพาะในประเด็นเกี่ยวกับการประเมินกระบวนการในการประมาณค่าหรือความครอบคลุมของค่าพารามิเตอร์ นอกจากนี้ยังมีการประยุกต์ใช้เทคนิคมอนติคาร์โลในการประเมินคุณสมบัติทางสถิติของโมเดลในการวิเคราะห์ข้อสอบ รวมทั้งยังมีการประยุกต์ใช้เทคนิคมอนติคาร์โลในการเปรียบเทียบวิธีการต่างๆ ในโมเดลการวิเคราะห์ข้อสอบ การวิเคราะห์ข้อสอบด้วยเทคนิคมอนติคาร์โลมีขั้นตอนที่สำคัญสามารถสรุปได้ 4 ขั้นตอนสำหรับการนำไปสู่การปฏิบัติการวิเคราะห์ ขั้นตอนแรก ได้แก่ การกำหนดปัญหาในการศึกษาวิจัย ขั้นตอนต่อมา เป็นขั้นตอนของการออกแบบการทดลอง ซึ่งรวมถึงการระบุตัวแปรอิสระและตัวแปรตาม การออกแบบการวิจัยเชิงทดลอง จำนวนรอบในการคำนวณ และการเลือกโมเดลการวิเคราะห์ข้อสอบ ขั้นตอนที่สามเป็นการเขียนและระบุโปรแกรมในการจำลองข้อมูลและการประมาณค่าพารามิเตอร์ และขั้นตอนสุดท้ายเป็นขั้นตอนเกี่ยวกับการวิเคราะห์ผลจากการจำลองข้อมูล

ตอนที่ 3 การทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O– NET)

หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน 2551 (ศึกษาธิการ, 2552) กำหนดให้มีการวัดผลการเรียนรู้ของผู้เรียนไว้ 4 ระดับ คือ ระดับชั้นเรียน ระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา และระดับชาติ มีรายละเอียดดังนี้

การประเมินระดับชั้นเรียน เป็นการวัดและประเมินผลที่อยู่ในกระบวนการเรียนรู้ ผู้สอนดำเนินการเป็นปกติ และสม่ำเสมอในการจัดการเรียนการสอน ใช้เทคนิคการประเมินที่หลากหลาย เป็นการตรวจสอบว่าผู้เรียนมีพัฒนาการความก้าวหน้าในการเรียนรู้ อันเป็นผลมาจากการจัดกิจกรรมการเรียนการสอนหรือไม่ และมากน้อยเพียงใด มีสิ่งที่จะต้องได้รับการพัฒนาปรับปรุงและส่งเสริมเพียงใด นอกจากนี้ยังเป็นข้อมูลให้ผู้สอนใช้ปรับปรุงการเรียนการสอนของตนด้วย ทั้งนี้โดยสอดคล้องกับมาตรฐานการเรียนรู้และตัวชี้วัด

การประเมินระดับสถานศึกษา เป็นการตรวจสอบผลการเรียนของผู้เรียนเป็นรายปี หรือรายภาค ผลการอ่าน คิดวิเคราะห์ และเขียนคุณลักษณะอันพึงประสงค์ และกิจกรรมพัฒนาผู้เรียน และเป็นการประเมินเกี่ยวกับการจัดการศึกษาของสถานศึกษา ว่าส่งผลต่อการเรียนรู้ของผู้เรียนตามเป้าหมาย หรือไม่ ผู้เรียนต้องการพัฒนาในด้านใด รวมทั้งนำผลการเรียนของผู้เรียน ในสถานศึกษาเปรียบเทียบกับเกณฑ์ระดับชาติ และระดับเขตพื้นที่การศึกษา เพื่อเป็นข้อมูลในการปรับปรุงนโยบายและหลักสูตรสถานศึกษาให้เหมาะสม

การประเมินระดับเขตพื้นที่การศึกษา เป็นการประเมินคุณภาพผู้เรียนในระดับเขตพื้นที่การศึกษาตามมาตรฐานการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน เพื่อใช้เป็นข้อมูลพื้นฐานในการพัฒนาคุณภาพผู้เรียนด้วยวิธีการและเครื่องมือที่เป็นมาตรฐานที่จัดทำและดำเนินการโดยเขตพื้นที่การศึกษา

การประเมินระดับชาติ เป็นการประเมินคุณภาพของผู้เรียนในระดับชาติตามมาตรฐานการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน สถานศึกษาต้องจัดให้ผู้เรียนทุกคนที่เรียนในชั้นประถมศึกษาปีที่ 3 ชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 เข้ารับการประเมิน ผลจากการประเมินใช้เป็นข้อมูลในการเทียบเคียงคุณภาพการศึกษาในระดับต่าง ๆ เพื่อนำไปใช้ในการวางแผน ยกระดับคุณภาพการจัดการศึกษา ตลอดจนเป็นข้อมูลสนับสนุนการตัดสินใจในระดับนโยบายของประเทศ

ในการประเมินระดับชาตินี้ จำเป็นต้องมีสถาบันทดสอบทางการศึกษา เพื่อเป็นองค์กรกลางในการศึกษา วิจัยและพัฒนา และให้บริการประเมินผลทางการศึกษาและทดสอบทางการศึกษา รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและนานาชาติ และเพื่อให้สถาบันที่จัดตั้งขึ้นมีการบริหารและการจัดการที่มีความเป็นอิสระคล่องตัว และมีประสิทธิภาพ

สมควรจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติขึ้นเป็นองค์การมหาชนตามกฎหมายว่าด้วยองค์การมหาชน จึงจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. ขึ้นตามพระราชกฤษฎีกาจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) พ.ศ. 2548 โดย สทศ. มีการจัดสอบระดับชาติหลายประเภท เช่น การทดสอบ GAT : General Aptitude Test (ความถนัดทั่วไป) สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่จะสอบเข้ามหาวิทยาลัย PAT : Professional Aptitude Test (ความถนัดทางวิชาชีพและวิชาการ) สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่จะสอบเข้ามหาวิทยาลัย O-NET : Ordinary National Educational Test (การทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน) เป็นการสอบความรู้รวบยอดปลายช่วงชั้นที่ 1 - 4 (6 ภาคเรียน) ของชั้น ป. 3 ป.6 ม. 3 และ ม. 6 ตามมาตรฐานการเรียนรู้ของหลักสูตร ในงานวิจัยนี้ศึกษากับข้อมูลผลการทดสอบระดับชาติ O-NET เนื่องจาก การทดสอบ O-NET เป็นการทดสอบที่นักเรียนทุกคนต้องสอบ และมีการนำผลการสอบ O-NET ไปใช้ในการประเมินผ่านช่วงชั้น และ เป็นส่วนหนึ่งในการคัดเลือกเข้ามหาวิทยาลัย สำหรับรายละเอียดการสอบ O-NET นำเสนอดังต่อไปนี้

3.1 ความสำคัญของการสอบ O-NET

การทดสอบ O-NET มีจุดหมายสำคัญคือ เพื่อประเมินคุณภาพการศึกษาขั้นพื้นฐานของนักเรียนแต่ละโรงเรียน และนำข้อมูลมาทำแผนพัฒนานักเรียนให้สามารถอ่านออกเขียนได้ รู้จักคิดวิเคราะห์ เป็นการวัดผลสัมฤทธิ์ทางการเรียนของผู้เรียนแต่ละช่วงชั้นเมื่อเรียนครบ 3 ปี โดยจะทดสอบที่ชั้นสูงสุดของช่วงชั้นนั้น ๆ การทดสอบ O – NET จะทำการทดสอบผู้เรียนที่อยู่ในชั้นประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และ มัธยมศึกษาปีที่ 6 ข้อสอบ O-NET จะครอบคลุมเนื้อหาและมาตรฐานการเรียนรู้ตรงตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ซึ่งประกอบด้วย 8 กลุ่มสาระดังนี้ คือ กลุ่มสาระการเรียนรู้ภาษาไทย กลุ่มสาระการเรียนรู้สังคม ศาสนาและวัฒนธรรม กลุ่มสาระการเรียนรู้ภาษาต่างประเทศ (ภาษาอังกฤษ) กลุ่มสาระการเรียนรู้คณิตศาสตร์ กลุ่มสาระการเรียนรู้วิทยาศาสตร์ กลุ่มสาระการเรียนรู้สุขศึกษา กลุ่มสาระการเรียนรู้ศิลปะ และกลุ่มสาระการเรียนรู้การงานอาชีพ เนื่องจากการทดสอบ O-NET เป็นการทดสอบมาตรฐานระดับชาติ เพื่อวัดมาตรฐานการศึกษาของนักเรียนชั้นมัธยมศึกษาทั่วประเทศ จึงจัดสอบด้วยข้อสอบเดียวกัน หรือข้อสอบคู่ขนานกัน และสอบในวันเดียวกัน ใช้มาตรฐานเดียวกันทั้งประเทศ จัดสอบโดยสถาบันทดสอบทางการศึกษาแห่งชาติ การทดสอบ O-NET เป็นการทดสอบที่มีความสำคัญ เนื่องจากนักเรียนจะจบการศึกษาแต่ละช่วงชั้นได้ ต้องสอบ O-NET และใช้ผลการทดสอบ O-NET ในการประเมินการผ่านช่วงชั้น และผลการทดสอบ O-NET สามารถใช้เป็นผลอ้างอิงในการประเมินคุณภาพการศึกษาระดับชาติตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน แสดงได้จาก

เป้าหมายของการปฏิรูปการศึกษาทศวรรษที่สอง (2552-2561) ยุทธศาสตร์ที่ 1 กำหนดให้คนไทยและการศึกษาไทยมีคุณภาพได้มาตรฐานสากล ตัวบ่งชี้ที่ 1.1 ระบุว่า ผลสัมฤทธิ์ทางการเรียนในวิชาหลัก (กระทรวงศึกษาธิการ, 2552) จากการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานมีคะแนนเฉลี่ยมากกว่าร้อยละ 50 ทั้งนี้โรงเรียนจะนำผลคะแนนสอบไปกรอกในแบบ รายงานผลการเรียนนักเรียน (ปพ.1) ในช่องผลการประเมินระดับชาติ

นอกจากนี้ นโยบายรัฐมนตรีว่าการกระทรวงศึกษาธิการกำหนดให้ใช้ผลการสอบ O-NET เป็นส่วนหนึ่งของการสอบแข่งขันคัดเลือกเข้าเรียนต่อชั้นมัธยมศึกษาปีที่ 1 และมัธยมศึกษาปีที่ 4 ในโรงเรียนที่มีอัตราการแข่งขันสูง ใช้ในการประกันคุณภาพการศึกษา (พระราชบัญญัติการศึกษา มาตรา 9 (3), 48, 49) ทั้งการประกันคุณภาพภายใน และการประเมินคุณภาพภายนอก ใช้ผลการทดสอบ O-NET เป็นเกณฑ์ในการปรับปรุงและพัฒนาการเรียน การสอน การบริหาร และใช้เป็นส่วนหนึ่งของการคัดเลือกเข้าเรียนต่อในระดับที่สูงขึ้น เป็นส่วนหนึ่งของ EXIT EXAM ให้ครูนำไปใช้ปรับปรุงการเรียนการสอน นอกจากนี้ จากที่ประชุมอธิการบดีแห่งประเทศไทย มีนโยบายให้ใช้ ผลการสอบ O-NET เป็นองค์ประกอบหนึ่งในการคัดเลือก บุคคลเข้าศึกษาในระบบ Admissions

จากที่กล่าวข้างต้นจะเห็นได้ว่า การทดสอบ O-NET เป็นการทดสอบที่มีความสำคัญมากเป็นการทดสอบระดับชาติ เป็นการทดสอบที่มีมาตรฐาน สามารถจำแนกผู้เรียนเป็นรายบุคคลได้ว่ามีความสามารถระดับใด ผ่านมาตรฐานใดบ้าง ผู้วิจัยจึงสนใจนำผลการทดสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มาใช้ในการจำแนกผู้เรียนตามความสามารถแฝง ทั้งนี้จะศึกษากับเฉพาะวิชาคณิตศาสตร์ เนื่องจากวิชาคณิตศาสตร์เป็นวิชาที่มีความสำคัญ และมีบทบาทต่อการพัฒนาความคิด ทำให้มีความคิดสร้างสรรค์ คิดอย่างมีเหตุผล เป็นระบบ มีแบบแผน สามารถวิเคราะห์ปัญหาและสถานการณ์ได้ถี่ถ้วนรอบคอบ ทำให้สามารถคาดการณ์ วางแผน ตัดสินใจ และแก้ปัญหาได้อย่างถูกต้องและเหมาะสม

3.2 ลักษณะทั่วไปของข้อสอบ O – NET

1. แบบทดสอบจะมีทั้งปรนัย และอัตนัย ในอัตราส่วนระหว่าง 80% ถึง 90% ต่อ 10% ถึง 20% ข้อสอบแบบปรนัยจะเป็นข้อสอบแบบ 4 ตัวเลือก สำหรับข้อสอบอัตนัยจะเป็นข้อสอบแบบเขียนคำตอบสั้นๆ (Short Answer)

2. เวลาในการทำข้อสอบวิชาละ 2 ชั่วโมง

3. ข้อสอบแต่ละข้อ คะแนนอาจจะไม่เท่ากันขึ้นอยู่กับความยากง่ายของข้อสอบ

4. ข้อสอบครอบคลุมสาระและทักษะสำคัญของ 8 กลุ่มสาระการเรียนรู้

3.3 วิธีการคิดคะแนน

1. เมื่อมีการตรวจกระดาษคำตอบ มีวิธีการวิเคราะห์ข้อสอบทุกข้อเพื่อหาคุณภาพของข้อสอบ ถ้าข้อสอบข้อใดไม่มีคุณภาพตามเกณฑ์ ไม่สามารถวัดหรือจำแนกได้ ข้อสอบข้อนั้นจะไม่นำมาคิดคะแนน ดังนั้นคะแนนที่ได้จะเป็นคะแนนสอบที่ได้มาจากข้อสอบที่มีคุณภาพทุกข้อ

2. คะแนนผลการสอบจะแปลงคะแนนเป็นคะแนนมาตรฐานรายวิชา

3.4 ลักษณะข้อสอบ O-NET วิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6

เนื้อหาที่ใช้สอบในวิชาคณิตศาสตร์ ประกอบด้วย 6 สาระ ได้แก่ จำนวนและการดำเนินการ การวัด เรขาคณิต พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น ทักษะ / กระบวนการทางคณิตศาสตร์

จำนวนข้อสอบ O-NET ในวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2555 มีทั้งหมด 40 ข้อ ข้อละ 2.5 คะแนน รวม 100 คะแนน แยกตามรูปแบบของข้อสอบ และแยกตามมาตรฐานการเรียนรู้ แสดงดังตารางต่อไปนี้

ตารางที่ 2.5 จำนวนข้อสอบ O-NET วิชาคณิตศาสตร์ แยกตามรูปแบบของข้อสอบ ปีการศึกษา 2555 (สถาบันทดสอบทางการศึกษา, 2555)

รูปแบบข้อสอบ	จำนวนข้อ	คะแนน
4 ตัวเลือก 1 คำตอบ	32	80
ระบายคำตอบที่เป็นค่า/ตัวเลข	8	20
รวม	40	100

การวิจัยในครั้งนี้ ผู้วิจัยเลือกข้อสอบแบบเลือกตอบ โดยใช้การสุ่มอย่างง่าย จำนวน 15 ข้อ ให้ครอบคลุม 4 สาระการเรียนรู้ ซึ่งได้แก่ สาระที่ 1 จำนวนและการดำเนินการ สาระที่ 2 การวัด สาระที่ 3 พีชคณิต และสาระที่ 4 การวิเคราะห์ข้อมูลและความน่าจะเป็น

ตารางที่ 2.6 จำนวนข้อสอบ O-NET วิชาคณิตศาสตร์ แยกตามสาระ และมาตรฐานการเรียนรู้ ปีการศึกษา 2555 (สถาบันทดสอบทางการศึกษา, 2555)

สาระ/ มาตรฐานการเรียนรู้	ช่วงชั้นที่ 4	
	ข้อ	คะแนน
สาระที่ 1 จำนวนและการดำเนินการ		
มาตรฐาน ค 1.1 เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง	3	7.5
มาตรฐาน ค 1.2 เข้าใจถึงผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่างๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้	1	2.5
มาตรฐาน ค 1.3 ใช้การประมาณค่าในการคำนวณและแก้ปัญหาได้	2	5.0
มาตรฐาน ค 1.4 เข้าใจในระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้	2	5.0
สาระที่ 2 การวัด		
มาตรฐาน ค 2.1 เข้าใจพื้นฐานเกี่ยวกับการวัด	3	7.5
มาตรฐาน ค 2.2 วัดและคาดคะเนขนาดของสิ่งที่ต้องการวัดได้	1	2.5
มาตรฐาน ค 2.3 แก้ปัญหาเกี่ยวกับการวัดได้		
สาระที่ 3 เรขาคณิต		
มาตรฐาน ค 3.1 อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติได้	-	-
มาตรฐาน ค 3.2 ใช้การนึกภาพ (visualization) ใช้เหตุผลเกี่ยวกับปริภูมิ (spatial reasoning) และใช้แบบจำลองทางเรขาคณิต (geometric model) ในการแก้ปัญหาได้	-	-
สาระที่ 4 พีชคณิต		
มาตรฐาน ค 4.1 อธิบายและวิเคราะห์แบบรูป (pattern) ความสัมพันธ์ และฟังก์ชันต่าง ๆ ได้	9	22.5
มาตรฐาน ค 4.2 ใช้นิพจน์ สมการ อสมการ กราฟ และแบบจำลองทางคณิตศาสตร์อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมายและนำไปใช้แก้ปัญหาได้	10	25.0
สาระที่ 5 การวิเคราะห์ข้อมูลและความน่าจะเป็น		
มาตรฐาน ค 5.1 เข้าใจและใช้วิธีการทางสถิติในการวิเคราะห์ข้อมูลได้	6	15.0
มาตรฐาน ค 5.2 ใช้วิธีการทางสถิติและความรู้เกี่ยวกับความน่าจะเป็นในการคาดการณ์ได้อย่างสมเหตุสมผล	1	2.5
มาตรฐาน ค 5.3 ใช้ความรู้เกี่ยวกับสถิติและความน่าจะเป็นช่วยในการตัดสินใจ และแก้ปัญหา ได้	2	5.0
สาระที่ 6 ทักษะกระบวนการทางคณิตศาสตร์		
มาตรฐาน ค 6.1 มีความสามารถในการแก้ปัญหา		
มาตรฐาน ค 6.2 มีความสามารถในการให้เหตุผล	-	-
มาตรฐาน ค 6.3 มีความสามารถในการสื่อสาร การสื่อความหมายทางคณิตศาสตร์และการนำเสนอ	-	-
มาตรฐาน ค 6.4 มีความสามารถในการเชื่อมโยงความรู้ต่างๆ ทางคณิตศาสตร์และเชื่อมโยงคณิตศาสตร์กับศาสตร์อื่นๆ ได้	-	-
มาตรฐาน ค 6.5 มีความคิดริเริ่มสร้างสรรค์	-	-
รวม	40	100

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับเป็นโมเดลใหม่ที่ถูกนำเสนอโดย Cho และ Cohen (2010) และพัฒนาต่อในมุมมองที่แตกต่างกันโดย Tay และคณะ (2011) ซึ่งเป็นโมเดลที่ขยายจากโมเดลการตอบสนองข้อสอบแบบผสมดังนั้น ในการเสนองานวิจัยที่เกี่ยวข้องจะนำเสนอทั้งโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ และโมเดลการตอบสนองข้อสอบแบบผสม

4.1 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

จากการศึกษางานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ พบว่ามีการเสนอโมเดลในลักษณะการรวมกันของโมเดลที่แตกต่างกัน และมีการประยุกต์ใช้ในมุมมองที่แตกต่างกันคือใช้ในการจัดกลุ่มซึ่งเป็นบริบทการวิจัยในองค์กร (De Jong & Steenkamp, 2010; Tay, Diener, et al., 2011) ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบซึ่งเป็นบริบทการวิจัยทางการศึกษา (Cho & Cohen, 2010)

Tay, Diener, Drasgow และ Vermunt (2011) เสนอโมเดล Multilevel Mixed-Measurement IRT ซึ่งเป็นการรวมกันของ Multilevel model Latent Class model และ IRT model (2 PL model) ซึ่งเป็น Multilevel Mixture IRT แบบ multidimensional เป็นการประยุกต์ทางทฤษฎี และวิธีการเกี่ยวข้องกับ Cross – cultural, Multilevel และ Measure equivalence โมเดลนี้ยอมให้กลุ่มของแต่ละบุคคลสามารถข้ามโมเดลการวัดในอีกลำดับขั้น หน่วยลำดับขั้นถูกจำแนกในขอบเขตซึ่งมีสัดส่วนที่คล้ายกันของระดับบุคคล ในการศึกษานี้ประยุกต์ใช้โมเดลในการจำแนกกลุ่มบุคคลตามการรายงานอารมณ์ของตนเอง โดยมีการศึกษาจำลองข้อมูลใน Monte Carlo และวิเคราะห์ข้อมูลจริงจากการ Gallop World Poll ซึ่งใช้กลุ่มตัวอย่างจำนวน 121,740 คน ใน 116 ประเทศ เป็นข้อมูลในปี 2005-2006 ซึ่งแบบวัดที่ใช้จะเป็นการให้คะแนนแบบสองค่า (ตอบ No ได้, Yes ได้ 1) มีการประมาณค่าพารามิเตอร์โดยใช้ Maximum Likelihood จากโปรแกรม Latent GOLD 4.0 และใช้เกณฑ์ Baysian Information Criterion (BIC) ซึ่งใช้ทั่วไปในการประเมิน Latent class ในการจำแนกกลุ่ม ซึ่งผลในการจำแนกกลุ่มในระดับบุคคล จำแนกเป็น 4 กลุ่ม และในการจำแนกกลุ่มระดับข้ามประเทศจำแนกเป็น 5 กลุ่ม

Cho และ Cohen (2010) เสนอโมเดล Multilevel Mixture IRT ซึ่งเป็นการรวมกันของ Multilevel model Latent Class model และ IRT model (Rasch model) ซึ่งเป็น Multilevel Mixture IRT แบบ one-dimensional ซึ่งโมเดลนี้ถูกแนะนำว่าเป็นประโยชน์ในการระบุกลุ่มแฝงซึ่ง

ยอมให้มิติรบกวนซึ่งอาจจะแตกต่างกันที่ระดับที่แตกต่างกัน นอกจากนี้ Cho และ Cohen ได้เสนอวิธีการประยุกต์ใช้โมเดลในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ศึกษาโดยจำลองข้อมูลและใช้ข้อมูลจริงจากการทดสอบมาตรฐานคณิตศาสตร์ ใช้การประมาณค่าแบบ Baysian ซึ่งถูกอธิบายโดยการวิเคราะห์ Label Switching การใช้ Prior และขั้นตอนการเลือก model

De Jong และ Steenkamp (2010) ศึกษาในขอบเขตของการผสมกันของโมเดล Mixture IRT ที่เป็นแบบพหุระดับ (Multilevel) พหุมิติ (Multidimensional) และลำดับ (Ordinal) สำหรับข้อมูลขนาดใหญ่ในการศึกษาข้ามวัฒนธรรม ซึ่ง Mixture model มีความเกี่ยวข้อง สัมพันธ์กัน และซ้อนกัน กับ multilevel multidimensional IRT model. ในการศึกษานี้ซึ่งเสนอโมเดลสำหรับการยืนยันในการวิจัย ความสำคัญสำหรับพารามิเตอร์ข้อสอบเป็นการแจกแจงแบบผสมในการจัดสถานการณ์ที่กลุ่มที่แตกต่างกันของหลายประเทศที่มีกระบวนการวัดที่แตกต่างกัน เมื่อประเทศอยู่ภายในกลุ่มที่มีความเป็นวิวิธพันธ์ การศึกษาจำลองข้อมูลถูกรวมให้แสดงพารามิเตอร์ และประยุกต์โมเดลในข้อมูลจริงใน 2 องค์ประกอบของ subjective well- being คือ ผลทางบวกและทางลบ การศึกษาครั้งนี้ศึกษาโดยใช้กลุ่มตัวอย่าง 13,112 คน ใช้แบบวัด five-point ordinal scaleใช้การประมาณค่าแบบ maximum likelihood ใช้โปรแกรม MATLAB ไม่ใช้ WinBUGs เนื่องจาก ใน WinBugs มีความยากในการประมาณ thresholds และ structural models ที่เกิดขึ้นพร้อมกัน โมเดลถูกระบุโดยการให้คะแนนผ่านพารามิเตอร์ข้อสอบในแต่ละการทำซ้ำซึ่งทำไม่ได้ใน WinBUGs และสุดท้ายวิเคราะห์ได้ช้ากับข้อมูลขนาดใหญ่ ผลการวิเคราะห์ในการศึกษานี้แสดงถึง 4 กลุ่มที่แตกต่างกัน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2 งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม

จากการศึกษางานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม พบว่ามีการนำโมเดลนี้ไปประยุกต์ในหลายด้าน ได้แก่ การจัดกลุ่ม (Bolt et al., 2001; Carter, Dalal, Lake, Lin, & Zickar, 2011; Finch & Pierson, 2011; Van Nijlen & Janssen, 2008) ความไม่แปรเปลี่ยนของการวัด (Tay, Newman, & Vermunt, 2011) การศึกษาเกี่ยวกับความตรงที่วัดได้ และความตรงเชิงทำนาย (Egberink, Meijer, & Veldkamp, 2010; Maij-de Meij et al., 2008) การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (Maij-de Meij, Kelderman, & van der Flier, 2010) การเปรียบเทียบดัชนีการเลือกโมเดลสำหรับใช้กับการตอบสนองข้อสอบแบบผสมของข้อสอบที่เป็นแบบให้คะแนนสองค่า (Li et al., 2009)

Bolt, Cohen และ Wollack (2001) ประยุกต์ใช้โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมในการตรวจสอบความแตกต่างระหว่างบุคคล โดยจำแนกกลุ่มบุคคลจากประเภทที่ตอบสนองในข้อสอบแบบเลือกตอบ โมเดลนี้อธิบายสำหรับความสัมพันธ์กันทั่วไประหว่างประเภทการตอบสนองโดยสมมติว่าผู้สอบอยู่ใช้กลุ่มแฝงต่อเนื่อง ซึ่งมีความแตกต่างกันของการชอบหรือจิตใจที่เอนเอียงกับการตอบสนองนั้น มีการจำแนกกลุ่มโดยพารามิเตอร์ใน Nominal response model และมีความแตกต่างกันข้ามกลุ่ม ในการศึกษานี้ใช้การประมาณค่าโมเดลและจำแนกกลุ่มโดยขั้นตอนของ Markov Chain Monte Carlo โดยใช้โปรแกรม WinBUG และใช้เกณฑ์ cross-validation log-likelihood และ local dependent criteria มีการใช้ข้อมูลจากการสอบวิชาภาษาอังกฤษ ในการแสดงว่าโมเดลสามารถจำแนกผู้สอบที่มีความชอบแตกต่างกัน แล้วจัดกลุ่ม และมีการจำลองข้อมูลในการประเมินพารามิเตอร์ข้อสอบ และการจำแนกที่ถูกต้องในการสมมติการออกแบบการทดสอบแบบเลือกตอบเพื่อวินิจฉัย ซึ่งผลการศึกษาจำแนกกลุ่มแบ่งออกเป็น 2 กลุ่มคือ กลุ่มที่ 1 เป็นกลุ่มที่การใช้คำที่ยากเป็นสาเหตุของการตอบผิด ส่วนกลุ่มที่ 2 การใช้เครื่องหมายวรรคตอน เป็นสาเหตุของการตอบผิด สามารถนำผลการศึกษา โครงสร้างการทดสอบ และการใช้แบบทดสอบแบบเลือกตอบไปประยุกต์ใช้ในการวินิจฉัยพุทธิปัญญาของรูปแบบการตอบสนองข้อสอบต่อไป

Finch และ Pierson (2011) ประยุกต์ Mixture IRT ไปใช้กับข้อมูลจาก Youth Risk Behavior Survey (2009) ในการระบุกลุ่มย่อยของวัยรุ่นที่เสี่ยงกับการติดในพฤติกรรมที่ไม่ดีต่อสุขภาพ และหาข้อคำถามสำรวจแต่ละบุคคลที่มีประสิทธิภาพในการระบุกลุ่มย่อย สำหรับ Mixture IRT ที่ใช้เป็นการรวมกันของ 2 PL model กับ LC model และแบบวัดที่ใช้เป็นแบบให้คะแนนสองค่า มีการประมาณค่าพารามิเตอร์โดยใช้ Maximum Likelihood และใช้เกณฑ์ Sample size adjusted BIC (SBIC) โดยใช้โปรแกรม MPLUS ซึ่งจุดมุ่งหมายในการศึกษานี้คือ (1) แสดงประโยชน์ของ MixIRT สำหรับการระบุกลุ่มย่อยในประชากร และเน้นการใช้พารามิเตอร์ที่เฉพาะเจาะจง ตามทฤษฎีการตอบสนองข้อสอบ (2) ระบุเกี่ยวกับรูปแบบของวัยรุ่นบนพื้นฐานของความชอบ นิสัยสำหรับการติดพฤติกรรมเสี่ยงเกี่ยวกับเพศ และการใช้สาร ซึ่งผลแสดง 4 กลุ่มของประชากรวัยรุ่นกับความแตกต่างในพฤติกรรมเสี่ยงของเพศกับการใช้สาร กลุ่มแรกมีนิสัยชอบในการเพิ่มความเสี่ยงพฤติกรรมทางเพศ กลุ่มที่สอง มีความชอบในการสูบบุหรี่ กลุ่มที่สามมีความชอบในการใช้สาร เช่น Marijuana, Methamphetamine และกลุ่มที่สี่ มีความชอบน้อยที่สุดสำหรับการชอบในพฤติกรรมเกี่ยวกับเพศ หรือการใช้สาร และสุดท้ายในแต่ละข้อคำถามจะถูกระบุสำหรับแต่ละกลุ่มซึ่งสามารถระบุความเสี่ยงที่มีประสิทธิภาพ

Egberink, Meijer และ Veldkamp (2010) ได้ศึกษาเกี่ยวกับการตรวจสอบความตรงที่วัดได้ และความตรงเชิงทำนายในการศึกษาความตระหนักรู้ในสถานที่ทำงาน โดยการประยุกต์ใช้ Mixture IRT ในการประเมินโครงสร้างที่เป็นพหุมิติ (multidimensional) และตรวจจับรูปแบบการตอบสนองที่แตกต่างกันสำหรับกลุ่มที่แตกต่างกัน ในการศึกษา MixIRT นี้เป็นการรวมกันของ Graded response model กับ LC model กลุ่มตัวอย่างที่ศึกษาจำนวน 9,283 คน แบบวัดที่ใช้เป็นแบบมาตรประมาณค่า 5 ระดับ จำนวน 144 ข้อ วิเคราะห์ข้อมูลโดย Latent GOLD และมีการใช้เกณฑ์ Baysian Information Criterion (BIC) ในการจำแนกกลุ่ม ผลการศึกษาแบ่งออกเป็น 4 กลุ่ม ผลแสดงให้เห็นถึงความตระหนักมีความแตกต่างกันเชิงคุณภาพสำหรับกลุ่มที่แตกต่างกันของบุคคล และความตรงเชิงทำนายของคะแนน ถูกปรับปรุงสำหรับบุคคลที่ในกลุ่มที่แตกต่างกัน เป็นการเปรียบเทียบความสอดคล้องกับโครงสร้างที่เป็นเอกมิติ (unidimensional)

Tay , Newman และ Vermunt (2011) ได้เสนอโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมการวัดกับตัวแปรร่วม และใช้โมเดลในการสืบหาความเท่าเทียมกันของการวัดสิ่งที่สังเกตได้ และไม่ได้ ซึ่งจากแนวคิดเกี่ยวกับการวัดความไม่แปรเปลี่ยนการตอบสนองข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิม ตรวจสอบความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มที่สังเกตได้ (เช่น เพศ เชื้อชาติ) ในทางตรงข้าม ทฤษฎีการทดสอบแบบผสมการวัด ใช้ในการค้นคว้าความ ไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มที่ไม่สามารถสังเกตได้ (เช่น กลุ่มแฝงของการตอบ ถูกจำแนกโดยความแตกต่างกันของคะแนน) ทั้งสองแนวคิด สามารถถูกรวมโดยการใช้โมเดลการตอบสนองข้อสอบแบบผสมการวัดกับตัวแปรร่วม (Mixed-Measurement Item Response Theory with Covariates model: MMIRT-C) ซึ่งในตัวแปรร่วม (คุณลักษณะที่สังเกตได้) ถูกออกแบบในการรวมกันกับ LCs สำหรับการอธิบาย ถ้าความไม่แปรเปลี่ยนในการวัดเป็นคุณลักษณะที่เป็นกลุ่มที่สามารถสังเกตได้ และสังเกตไม่ได้ (จำแนกหรือต่อเนื่อง) เกิดขึ้นพร้อมกัน ดังนั้นความก้าวหน้าเกี่ยวกับแนวคิดทั่ว ๆ ไป ถูกสังเกตในความไม่แปรเปลี่ยนของการวัด MMIRT-C สามารถใช้ในการอ้างถึงการวัดกลุ่มแฝง และตัดสินความสัมพันธ์ของสมาชิกในกลุ่มแฝงกับคุณลักษณะที่สังเกตได้ และ ตัดสินว่า มีการวัดที่สังเกตได้ หรือมีความไม่เท่าเทียมกันเกิดขึ้น อย่างเหนือกว่า ภายในการวัดกลุ่มแฝงปกติ วิธีการนี้จะแสดงถึงการใช้การวัดของ พฤติกรรมของคน กับประสบการณ์ในการทำงาน และเพศเป็นตัวแปรร่วม แนวคิดที่เสนอนี้เป็นการขยายการวิจัยเกี่ยวกับความไม่แปรเปลี่ยนของการวัดในบริบทขององค์กรโดยการตัดสินจาก 1 คำถาม (เช่นมีความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มจำแนกที่สังเกตได้หรือไม่) หรืออาจกล่าวได้ว่า แนวคิดนี้ขยายจากโมเดล IRT-DIF ซึ่ง MMIRT-C สามารถตัดสินได้ว่ามีความแตกต่างที่ไม่สามารถสังเกตได้ในการใช้มาตรวัด ประเมินระดับซึ่งกลุ่มย่อยแฝงแสดงโมเดลการวัดที่แตกต่างกัน

ตรวจสอบการเชื่อมโยงระหว่างคุณลักษณะที่สังเกตได้และสมาชิกกลุ่มที่ไม่สามารถสังเกตได้ ตรวจสอบความสัมพันธ์ระหว่างคุณลักษณะที่สังเกตได้ กับคุณลักษณะแฝง ที่อยู่ภายในแต่ละกลุ่ม ที่สังเกตไม่ได้ ในการศึกษานี้ มีการจำลองข้อมูลก่อนการศึกษาเชิงประจักษ์ และในการศึกษาครั้งนี้เป็นการแสดงให้เห็นว่า MMIRT-C สามารถใช้ในการตัดสิน overall DIF เราแสดง โมเดลที่เข้มงวดของ MMIRT-C หรือ IRT-C model ซึ่งสามารถใช้ระบุ uniform และ non-uniform observe DIF ข้ามอายุ และเพศพร้อมกัน กระบวนการนี้เปรียบเทียบแนวคิด IRT DIF โดยมาตรฐาน ทั้ง Lord's χ^2 และ DFIT methodology เมื่อเกิด DIF กับประสบการณ์การทำงาน(ปี) และเพศ กับกระบวนการ IRT-DIF และสุดท้ายแสดงเงื่อนไขที่เฉพาะของ LC ใน MMIRT-C (unobserved DIF) เราสามารถใช้สำหรับบางส่วนของ observed DIF ต่อ ตัวแปรร่วม จุดประสงค์นี้ไม่ใช่การจัดการเพียงแต่กระบวนการการตรวจสอบความแตกต่างจาก DIF แต่ยังให้รายละเอียดการตัดสินทางการปฏิบัติที่พัฒนาในโมเดลที่เฉพาะเจาะจง ใช้เกณฑ์ BIC ในการแบ่งกลุ่ม ในการศึกษานี้ใช้โปรแกรม Latent Gold ใช้การประมาณค่าแบบ Maximum Likelihood และมีการเปรียบเทียบความแตกต่าง โดยเน้นที่ observe DIF กับ overall DIF สำหรับการศึกษาเชิงประจักษ์ ได้รวบรวมข้อมูลจากการทดสอบขนาดใหญ่ของสหภาพระหว่างบุคลากรในโรงเรียนรัฐบาล (ครู เจ้าหน้าที่ห้องสมุด พยาบาล) ซึ่งเป็นกลุ่มตัวอย่างจากโรงเรียน 466 แห่ง จากรัฐแถบตะวันออกเฉียงเหนือ

Annette Meij Kelderman และ Flier (2008) โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมที่เหมาะสมกับข้อมูลแบบสอบถามบุคลิกภาพ การอธิบายคุณลักษณะของกลุ่มแฝง และการตรวจสอบความเป็นไปได้สำหรับการพัฒนาการทำนาย โมเดลการตอบสนองข้อสอบแบบผสมมีจุดมุ่งหมายคือแสดงการตอบสนองของพฤติกรรมที่มีต่อการทดสอบบุคลิกภาพและอาจจะจัดการความเป็นไปได้สำหรับการทำนาย ความแตกต่างกันในประชากรถูกออกแบบโดยระบุกลุ่มย่อยที่เหมือนกัน ในการศึกษานี้โมเดลการตอบสนองข้อสอบแบบผสมถูกประยุกต์ใช้กับแบบวัด Extroversion and Neuroticism scales ของการสำรวจ Amsterdam Biographical Questionnaire และเป็นการรวมกับโมเดล nominal response model ซึ่งถูกระบุว่าเป็นโมเดลที่มีความเหมาะสมที่สุด กับโมเดล LC (3 กลุ่ม) สำหรับกลุ่มแฝงมีความแตกต่างกันเกี่ยวกับการพึงปรารถนาทางสังคมและภูมิหลังเกี่ยวกับชาติพันธุ์ ซึ่งมีแนวโน้มการตอบสนองในการจำแนกที่แตกต่างกัน ประเด็นที่สำคัญ คือการประยุกต์ผลของโมเดล Mixture IRT ในการทำนายที่ดีกว่าของความสัมพันธ์กับเกณฑ์ภายในที่เปรียบเทียบกับโมเดล 1 กลุ่มสำหรับ Neuroticism scale การทำนายถูกพัฒนาขึ้น แต่สำหรับ Extraversion scale ไม่สามารถทำนายได้ ผลแสดงให้เห็นถึง

ประโยชน์ที่เป็นไปได้ของการประยุกต์ โมเดลทฤษฎีการตอบสนองข้อสอบแบบผสม กับแบบวัดบุคลิกภาพ

Annette Meij Kelderman,และ Flier (2010) ได้พัฒนาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยการใช้ Mixture IRT model ซึ่งโดยปกติการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จะเปรียบเทียบข้อคำถามข้ามกลุ่มที่เด่นชัด หรือกลุ่มที่สังเกตได้ (manifest group) อย่างไรก็ตาม กลุ่มที่สังเกตได้ (manifest group) กับการทำหน้าที่ต่างกันของข้อสอบอาจจะไม่เหมือนกันกับแหล่งที่ถูกต้องของความลำเอียง ดังนั้นในการศึกษานี้จึงมีสมมติฐานว่าวิธีการในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบภายใต้โมเดลที่รวมตัวแปรแฝงจะมีความไวต่อแหล่งของความลำเอียงมากกว่าวิธีการอื่น ๆ การศึกษานี้จึงมีการเปรียบเทียบการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบที่มีการใช้ตัวแปรที่สังเกตได้โดยใช้สถิติ χ^2 กับแบบที่ใช้ตัวแปรกลุ่มแฝงโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมกับ Rasch model ในการศึกษานี้มีการจำลองข้อมูลซึ่งแสดงให้เห็นว่าโมเดลการตอบสนองข้อสอบแบบผสม ประกอบด้วยตัวแปรกลุ่มแฝง ซึ่งระบุการทำหน้าที่ต่างกันของข้อสอบได้ดีกว่าวิธีตรวจสอบแบบอื่นที่มีการใช้ตัวแปรที่สังเกตได้เพียงอย่างเดียว ความแตกต่างระหว่างการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่รวมตัวแปรแฝง กับสังเกตได้นี้เป็นความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้ และแหล่งที่ถูกต้องของการเกิด DIF ในกลุ่มที่เล็กกว่า นอกจากนี้มีการศึกษาถึงความแตกต่างของขนาดกลุ่มตัวอย่าง และระดับนัยสำคัญ มีการศึกษาเปรียบเทียบกับข้อมูลเชิงประจักษ์ในการทดสอบศัพท์ของชุดทดสอบสมรรถนะทั่วไป (the General Aptitude Test Battery: GATB) ในการประมาณพารามิเตอร์ข้อสอบใช้โปรแกรม *l*EM (Vermunt, 1997) และใช้ BIC เป็นเกณฑ์ตัดสินจำนวนกลุ่มแฝง

Carter และคณะ (2011) ใช้โมเดลการตอบสนองข้อสอบแบบผสมในการวิเคราะห์เชิงสำรวจในการตอบสนองในองค์กร โดยประยุกต์ใช้กับดัชนีรายละเอียดของงาน ซึ่งในบทความนี้ผู้วิจัยแสดงถึงการใช้โมเดลการตอบสนองข้อสอบแบบผสมและอธิบายถึงประโยชน์สำหรับการวิเคราะห์การสำรวจเกี่ยวกับองค์กร ผู้วิจัยเริ่มโดยการให้รายละเอียดทั่วไปเกี่ยวกับโมเดลการตอบสนองข้อสอบแบบผสมโดยเน้นที่มุมมองทางเทคนิคและการประยุกต์ใช้ในองค์กร ซึ่งโมเดลการตอบสนองข้อสอบแบบผสมในที่นี้เป็นการรวมกับ Partial Credit Model ซึ่งเป็น polytomous Rasch model และใช้โปรแกรม WINMIRA พารามิเตอร์ของข้อคำถามและสัดส่วนถูกประมาณโดยการใช้ expectation-maximization (EM) algorithm และใช้เงื่อนไข maximum likelihood การแนะนำนี้นักวิจัยจะสามารถใช้โมเดลในการตรวจสอบข้อตกลงเกี่ยวกับคะแนน ซึ่งระบุอิทธิพลของการตอบสนองอย่าง

เป็นระบบซึ่งไม่สัมพันธ์กับเนื้อหาข้อสอบ และประเมินตัวแปรความแตกต่างระหว่างกลุ่มและแต่ละบุคคลซึ่งเป็นตัวทำนายของสมาชิกในกลุ่ม หลังจากการสรุปงานวิจัยโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมจะระบุถึงปัญหาที่เกี่ยวข้องกับนักวิจัยองค์กร ผู้วิจัยนำเสนอการใช้โมเดลนี้กับดัชนีรายละเอียดเกี่ยวกับงาน และเน้นที่การใช้ ‘‘?’’ ในการตอบสนอง (ความไม่แน่ใจ) ผลการศึกษาจำแนกเป็น 3 กลุ่ม กลุ่มหนึ่งมีการตอบสนองในทิศทางบวก อีกกลุ่มหนึ่งในทิศทางลบ และอีกกลุ่มหนึ่งใช้ ‘‘?’’ ในการตอบสนอง ความเป็นจริง ในการจัดการ ช่วงเวลาในการดำรงตำแหน่งงาน อายุ การแข่งขัน และเพศถูกตัดสินเป็นความสัมพันธ์ของสมาชิกในกลุ่ม ซึ่งผลนี้ถูกถกเถียงในรูปแบบของการประยุกต์ใช้โมเดล และ การวิจัยในอนาคตต่อไป

Li และคณะ (2009) ได้พิจารณาดัชนีการเลือกโมเดลสำหรับใช้กับการตอบสนองข้อสอบแบบผสมของข้อสอบที่เป็นแบบให้คะแนนสองค่า โดยเปรียบเทียบดัชนี 5 แบบ ได้แก่ Akaike’s information coefficient (AIC), Bayesian information coefficient (BIC), deviance information coefficient (DIC), pseudo-Bayes factor (PsBF), and posterior predictive model checks (PPMC) โดยการจำลองสถานการณ์ กับ IRT Model แบบ 1PL, 2PL และ 3PL ใช้กับแบบสอบจำนวน 6, 15 และ 30 ข้อ เปรียบเทียบกับจำนวนผู้สอบ 600 และ 1,200 คน และเงื่อนไขแบบผสม 4 แบบ (4 simulated mixture conditions) การผสมสำหรับกลุ่มแฝง 4 กลุ่ม ถูกจำลองโดยการจัดการกับความยากของข้อสอบ อำนาจจำแนก พารามิเตอร์การเดา สำหรับแต่ละกลุ่มแฝง ซึ่งจำลองความสามารถภายใต้โค้งปกติสำหรับแต่ละกลุ่มแฝง ใช้การประมาณค่าแบบ Bayesian สำหรับโมเดลการตอบสนองข้อสอบแบบผสม ใช้ขั้นตอนการสุ่มแบบ Gibbs และใช้โปรแกรม WinBUGS 1.4 (Spiegelhalter et al., 2003) การประมาณค่าแบบ MCMC ของ Gibbs จำลอง Markov chain ซึ่งมีค่าสำหรับพารามิเตอร์ขงโมเดลที่ถูกสุ่มจากการแจกแจงก่อนหน้าอย่างเป็นเงื่อนไขที่สมบูรณ์ โดยการทำซ้ำจำนวนมาก ผลจากการจำลองข้อมูลชี้ให้เห็นว่าดัชนี BIC เลือกโมเดลที่ถูกต้องมากที่สุดภายใต้เงื่อนไขการจำลองที่มากมายนี้ และ PsBF มีประสิทธิภาพมากรองลงมา สำหรับ AIC และ PPMC มีแนวโน้มเลือกโมเดลที่ซับซ้อนภายใต้เงื่อนไขบางอย่าง และ DIC มีประสิทธิภาพน้อยที่สุด

Nijlen และ Janssen (2008) ได้ประยุกต์ใช้ MixIRT ในการตรวจสอบความชำนาญของกฎการสะกดคำในภาษา Dutch กับนักเรียน Grade 3- 6 จำนวน 1078 คน ซึ่งการตรวจสอบความชำนาญในการสะกดคำสามารถถูกพิจารณาเป็น purely dimension-like หรือมี category-like elements ที่แตกต่างกัน ถ้าความแตกต่างเชิงปริมาณในความชำนาญปรากฏ จะสามารถพิจารณา

เป็น dimension-like อย่างไรก็ตามถ้าความแตกต่างเชิงคุณภาพความชำนาญจะมีรูปแบบ category-like มากไปกว่านี้ ประเด็นจะถูกระบุว่าถ้าความแตกต่างเชิงคุณภาพปรากฏ จะสามารถเชื่อมโยงกับรูปแบบที่เห็นได้ชัดของประชากร หรือการใช้กลุ่มแฝงสามารถให้ข้อมูลที่มากกว่า ในการศึกษานี้ได้ประยุกต์ใช้ mixture Rasch model มีการใช้โปรแกรม BILOG-MG (Zimowski, Muraki, Mislevy & Bock, 2003) ในการประมาณค่าใช้ EM-algorithm และเกณฑ์ AIC และ the BIC เป็นเกณฑ์พิจารณา ซึ่งผลชี้ให้เห็นว่าความชำนาญของกฎการสะกดคำไม่สามารถพิจารณา dimension-like อย่างแท้จริง กลุ่มแฝง 3 กลุ่มของนักเรียนสามารถถูกทำให้แตกต่างกัน ซึ่งพบว่านักเรียน grade 3 จะถูกจัดอยู่ใน class เดียวกัน นักเรียน grade 4 ถูกจัดอยู่ใน class เดียวกัน ส่วน grade 5-6 อยู่ class เดียวกัน ซึ่งแสดงถึงความแตกต่างเชิงคุณภาพในความชำนาญของกฎการสะกดคำและ กฎที่สัมพันธ์กัน

จากการศึกษางานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม และโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ สามารถแสดงรายละเอียดได้ดังตารางที่ 2.7

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 2.7 งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับและทฤษฎีการตอบสนองข้อสอบแบบผสม

ผู้วิจัย	สิ่งที่ศึกษา	โมเดล	แบบสอบ /แบบวัด	software	วิธีประมาณค่า	ปัจจัยเงื่อนไข	เกณฑ์การพิจารณา
งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ							
Cho & Cohen (2010)	เสนอโมเดล MMixIRT และการประยุกต์ใช้โมเดลกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (การรวมกันของ Multilevel model กับ Latent class model และ Rasch model) โดยมีการจำลองข้อมูลก่อนการศึกษากับข้อมูลจริง	โมเดล 1 พารามิเตอร์ (1PL)	แบบสอบคณิตศาสตร์	WinBUG 1.4 (Spiegelhalter , Thomas & Best, 2003)	MCMC algorithm โดยใช้ Label switching	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล และในแต่ละระดับแบบข้ามระดับ นักเรียน และโรงเรียน โดยการศึกษา 1 พารามิเตอร์ (b)	Akaike Information Criterion (AIC) และ Baysian Information Criterion (BIC)
Tay.et.al (2011)	เสนอโมเดล MMixIRT และประยุกต์ใช้กับการรายงานอารมณ์ของแต่ละคน ข้ามระดับประเทศ (การรวมกันของ Multilevel model Latent class model และ 2PL model)	โมเดล 2 พารามิเตอร์ (2PL)	ใช้ข้อมูลจาก Gallup World Poll และเพิ่ม 3 ข้อ ในการประเมิน global life ซึ่งวัดความพึงพอใจใน อดีต ปัจจุบัน และอนาคต เป็น มาตรประมาณค่า 10 ระดับ	Latent GOLD 4.0 (Vermunt & magidson, 2005)	Maximum likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล และในแต่ละระดับแบบข้ามระดับ โดยการศึกษา 2 พารามิเตอร์ (a, b)	Baysian Information Criterion (BIC)
De Jong และ Steenkamp (2010)	เสนอโมเดลสำหรับการยืนยันโดยศึกษาในขอบเขตของการผสมกันของโมเดล Mixture IRT ที่เป็นแบบพหุระดับ (Multilevel) พหุมิติ (Multidimensional) และลำดับ (Ordinal) สำหรับข้อมูลขนาดใหญ่ในการศึกษาข้ามวัฒนธรรม โดยมีการจำลองข้อมูลก่อนการศึกษากับข้อมูลจริง	Multilevel IRT multi dimensional IRT	แบบวัด subjective well being แบบมาตรประมาณค่า 5 ระดับ	MATHLAB	Maximum likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล และในแต่ละระดับแบบข้ามระดับ	ไม่ระบุ

ตารางที่ 2.7 (ต่อ) งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับและทฤษฎีการตอบสนองข้อสอบแบบผสม

ผู้วิจัย	สิ่งที่ศึกษา	โมเดล	แบบสอบ /แบบวัด	software	วิธีประมาณค่า	ปัจจัยเงื่อนไข	เกณฑ์การพิจารณา
งานวิจัยที่เกี่ยวข้องกับโมเดลการตอบสนองข้อสอบแบบผสม							
Bolt, Cohen and Wollack (2001)	เสนอโมเดล MixIRT และประยุกต์ใช้กับการจำแนกกลุ่มบุคคลในการทดสอบภาษาอังกฤษเพื่อใช้ในการวินิจฉัย (การรวมกันของ Latent class model Nominal Response model) โดยศึกษากับข้อมูลจำลองก่อนศึกษากับข้อมูลจริง	Nominal response model		WinBUG	MCMC algorithm	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล ในการตอบข้อสอบ	cross-validation log-likelihood และ local dependence criteria
Finch and Pierson (2011)	ประยุกต์ Mixture IRT ไปใช้กับข้อมูลจาก Youth Risk Behavior Survey (2009) ในการระบุกลุ่มย่อยของวัยรุ่นที่เสี่ยงกับการติดในพฤติกรรมที่ไม่ดีต่อสุขภาพ และหาข้อคำถามสำรวจแต่ละบุคคลที่มีประสิทธิภาพในการระบุกลุ่มย่อย	โมเดล 2 พารามิเตอร์ (2PL)	แบบสำรวจ Youth Risk Behavior Survey (2009)	MPLUS	Maximum Likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบข้อสอบ	Sample size adjusted BIC (SBIC)
Egberink, Meijer และ Veldkamp (2010)	ศึกษาเกี่ยวกับการตรวจสอบความตรงที่วัดได้และความตรงเชิงทำนายในการศึกษาความตระหนักรู้ในสถานที่ทำงาน โดยการประยุกต์ใช้ Mixture IRT ในการประเมินโครงสร้างที่เป็นพหุมิติ และตรวจจับรูปแบบการตอบสนองที่แตกต่างกันสำหรับกลุ่มที่แตกต่างกัน ในการศึกษา MixIRT นี้เป็นการรวมกันของ Graded response model กับ LC model	Graded response model	แบบวัดความตระหนักรู้ในสถานที่ทำงานแบบมาตรประมาณค่า 5 ระดับ	Latent GOLD 4.0	Maximum Likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบข้อสอบ	Baysian Information Criterion (BIC)

ตารางที่ 2.7 (ต่อ) งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับและทฤษฎีการตอบสนองข้อสอบแบบผสม

ผู้วิจัย	สิ่งที่ศึกษา	โมเดล	แบบสอบ /แบบวัด	software	วิธีประมาณค่า	ปัจจัยเงื่อนไข	เกณฑ์การพิจารณา
Tay, Newman และ Vermunt (2011)	เสนอโมเดล Mixed-measurement IRT-C ซึ่งเป็นโมเดลขยายจาก IRT-DIF (การรวมกันของ Latent class model และ 2 PL model) โดยมีการจำลองข้อมูลก่อนการศึกษากับข้อมูลจริง	โมเดล 2 พารามิเตอร์ (2PL)	union citizenship /participation scale	Latent GOLD 4.5	Maximum Likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบ โดยพฤติกรรมกับประสบการณ์ ในการทำงานและเพศเป็นตัวแปรร่วม	BIC
Annette Meij Kelderman,and Flier (2008)	ประยุกต์ใช้โมเดลในการตรวจสอบความเป็นไปได้สำหรับการพัฒนาการทำนาย และการอธิบายคุณลักษณะของกลุ่มแฝงโดยใช้ในการวัดบุคลิกภาพ	Nominal Response model และ Partial Credits elmod	แบบวัด Extroversion and Neuroticism scales	lEM (Vermunt, 1997)	Likelihood ratio test	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบ กลุ่มแฝงมีความแตกต่างกันเกี่ยวกับการพึงปรารถนาทางสังคมและภูมิหลังเกี่ยวกับชาติพันธุ์	BIC
Annette Meij Kelderman,และ Flier (2010)	พัฒนาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยการใช้ Mixture IRT model	โมเดล 1 พารามิเตอร์ (Rasch model)	แบบทดสอบศัพท์ของชุดทดสอบสมรรถนะทั่วไป (the General Aptitude Test Battery: GATB)	lEM (Vermunt, 1997) Multilog (Thissen, Chen, & Bock, 2003)	Maximum Likelihood	การตรวจสอบ DIF โดยศึกษาเปรียบเทียบ MixIRT ซึ่งระบุความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล แบบ 1 พารามิเตอร์ (b) และศึกษาเปรียบเทียบกับการตรวจสอบ DIF ที่สังเกตได้ โดยใช้ χ^2 และเปรียบเทียบ ขนาดกลุ่มตัวอย่าง ความยาวแบบสอบ	BIC

ตารางที่ 2.7 (ต่อ) งานวิจัยที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับและทฤษฎีการตอบสนองข้อสอบแบบผสม

ผู้วิจัย	สิ่งที่ศึกษา	โมเดล	แบบสอบ /แบบวัด	software	วิธีประมาณค่า	ปัจจัยเงื่อนไข	เกณฑ์การพิจารณา
Carter และคณะ (2011)	ประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบผสมในการวิเคราะห์เชิงสำรวจในการตอบสนองในองค์กร โดยประยุกต์ใช้กับดัชนีรายละเอียดของงาน (การรวมกันของ Latent class model และ PCM)	Partial Credit Model	แบบวัด 3 ระดับ Job Descriptive Index (JDI)	WINMIRA	Maximum Likelihood	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบเพื่อการจำแนกกลุ่ม	Consistent Aikake's Information Criteria (CAIC)
Li และคณะ (2009)	ได้พิจารณาดัชนีการเลือกโมเดลสำหรับใช้กับการตอบสนองข้อสอบแบบผสมของข้อสอบที่เป็นแบบให้คะแนนสองค่า โดยเปรียบเทียบ 5 แบบ ได้แก่ AIC, BIC, DIC, PsBF และ PPMC มีการจำลองสถานการณ์ภายใต้เงื่อนไขที่หลากหลาย และเปรียบเทียบกับข้อมูลเชิงประจักษ์	IRT model 1PL, 2PL และ 3PL	แบบทดสอบคณิตศาสตร์ (Statewide Mathematics Test)	WinBUGS 1.4 (Spiegelhalter et al., 2003)	Bayesian, MCMC	เปรียบเทียบ IRT model 3 โมเดล กับ ความยาวแบบสอบ 3 แบบ กับจำนวนจำนวนผู้สอบ 2 แบบ และเงื่อนไขแบบผสม 4 แบบ	AIC, BIC, DIC, PsBF และ PPMC
Nijlen และ Janssen (2008)	ประยุกต์ใช้ MixIRT ในการจำแนกกลุ่มบุคคลเรื่องความชำนาญของการสะกดคำในภาษา Ducth	โมเดล 1 พารามิเตอร์ (Rasch model)	แบบทดสอบความชำนาญในการสะกดคำ	BILOG-MG (Zimowski, Muraki, Mislevy & Bock, 2003)	EM-algorithm	ความสัมพันธ์กันของคุณลักษณะภายในของแต่ละบุคคล กับรูปแบบการตอบเพื่อการจำแนกกลุ่ม	AIC และ BIC

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง แม้ว่าจะมีการศึกษาเกี่ยวกับโมเดลการตอบสนองข้อสอบแบบผสมมากขึ้น และในการศึกษาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับยังมีไม่มาก ทั้งในการวิเคราะห์เพื่อการประมาณค่าพารามิเตอร์ และการนำโมเดลไปประยุกต์ใช้ อย่างไรก็ตาม ข้อมูลในทางการศึกษาส่วนใหญ่เป็นข้อมูลที่เป็นพหุระดับ และเป็นข้อมูลคุณลักษณะแฝงที่ไม่สามารถสังเกตได้ ดังนั้นในการศึกษาเกี่ยวกับโมเดลนี้จึงเป็นประโยชน์ในวงการศึกษา

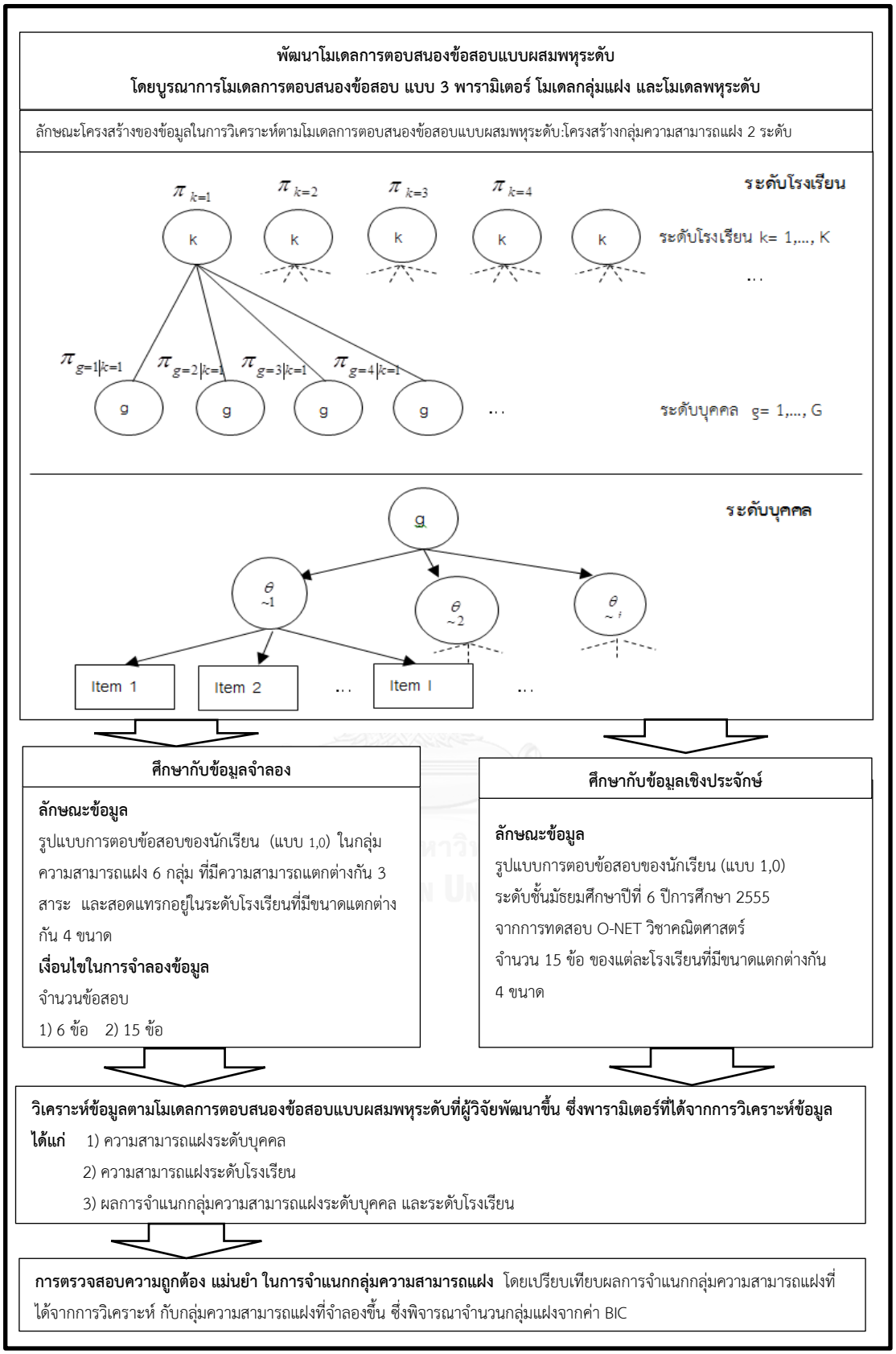
ตอนที่ 5 กรอบแนวคิดในการวิจัย

จากการศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้อง ทำให้ผู้วิจัยพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับซึ่งบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ และนำโมเดลที่พัฒนาขึ้นไปทดลองใช้กับข้อมูลจำลองเพื่อตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝง จากนั้นนำโมเดลที่พัฒนาขึ้นไปใช้กับข้อมูลเชิงประจักษ์ ซึ่งสามารถเขียนกรอบแนวคิดการวิจัยและลักษณะโครงสร้างของข้อมูลสำหรับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ได้ดังภาพที่ 2.12

สำหรับกรอบลักษณะโครงสร้างของข้อมูลในการวิเคราะห์ตามโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ผู้วิจัยปรับจาก Tay, Diener, et al. (2011) ซึ่งโครงสร้างของข้อมูลในการวิเคราะห์กลุ่มความสามารถแฝงมีลักษณะที่สอดแทรกกันในระดับลดหลั่น และระดับบุคคล โดยกรอบด้านบนแสดงโมเดลของความแตกต่างแฝงในระดับลดหลั่น และระดับบุคคล กรอบด้านล่างแสดงถึงความน่าจะเป็นของการตอบสนองข้อสอบซึ่งเป็นไปตามกลุ่มบุคคล และตำแหน่งคุณลักษณะ ซึ่งกลุ่มแฝงของโรงเรียน g ไม่เพียงแต่เป็นความสัมพันธ์กำกับของคุณลักษณะแฝงไปยังตัวชี้วัด แต่ยังเป็นผลของตัวชี้วัดโดยตรง

สำหรับความสามารถของบุคคล (θ) ถูกกำหนดจากผลการตอบข้อสอบในแต่ละข้อ และความสามารถของโรงเรียน ถูกกำหนดจากสัดส่วนของความสามารถแฝงในระดับบุคคล

สำหรับการจำแนกกลุ่มความสามารถแฝงในระดับบุคคล (g) จำแนกจากความน่าจะเป็นในการตอบข้อสอบถูกแต่ละข้อของแต่ละบุคคล และการจำแนกกลุ่มความสามารถแฝงในระดับโรงเรียน (k) จำแนกจากความสามารถของโรงเรียน



ภาพที่ 2.12 กรอบแนวคิดในการวิจัย

บทที่ 3
วิธีดำเนินการวิจัย

การศึกษาในครั้งนี้มีวัตถุประสงค์ 3 ประการ คือ (1) เพื่อพัฒนาโมเดลการประมาณค่ากลุ่มความสามารถแฝงของบุคคลตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ (2) เพื่อตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคลโดยใช้ โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลจำลอง และ (3) เพื่อวิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ในการวิจัยครั้งนี้ผู้วิจัยได้พัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ โดยศึกษาแนวคิด และงานวิจัยต่าง ๆ จนพัฒนาเป็นโมเดลการวิเคราะห์ข้อมูล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ข้อมูลที่ใช้ในการศึกษาครั้งนี้เป็นข้อมูลจำลอง ที่จำลองขึ้นจากโปรแกรม R และข้อมูลเชิงประจักษ์ ผลการตอบข้อสอบวิชาคณิตศาสตร์ ของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยในการวิจัยครั้งนี้ ผู้วิจัยใช้การวิเคราะห์ด้วยโปรแกรม R และมีวิธีการดำเนินการวิจัยโดยสรุปดังนี้

1. การศึกษาเอกสารและงานวิจัยทั้งในและต่างประเทศที่เกี่ยวกับแนวคิด วิธีการ และจุดเด่น ข้อจำกัดของการวิเคราะห์กลุ่มแฝง แนวคิดหลักการ วิธีการใช้งานรวมทั้งวิธีการเขียนคำสั่งสำหรับการวิเคราะห์ด้วยโปรแกรม R และพัฒนาโมเดลการประมาณค่าของทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ

2. จำลองข้อมูลด้วยโปรแกรม R แพคเกจ IRTOYS จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่าตัวแปรที่ส่งผลต่อการประมาณค่าพารามิเตอร์ของข้อสอบ ได้แก่ จำนวนข้อสอบที่ใช้ในการวิเคราะห์ข้อมูล ดังนั้นในการจำลองข้อมูล จึงมีเงื่อนไขในการจำลองข้อมูล ได้แก่ จำนวนข้อสอบที่ใช้ในการวิเคราะห์ ในงานวิจัยนี้จำลองกลุ่มแฝงระดับโรงเรียน 6 กลุ่มแฝง ซึ่งมีความสามารถในการตอบข้อสอบแตกต่างกัน 3 สาระ และอยู่ในโรงเรียนที่มีขนาดแตกต่างกัน 4 ขนาด ได้แก่ ขนาด

เล็ก ขนาดกลาง ขนาดใหญ่ และขนาดใหญ่พิเศษ โดยแต่ละกลุ่มแฝงจำลองผลการตอบข้อสอบสองสถานการณ์ คือ จำนวน 6 ข้อ และ 15 ข้อ

3. วิเคราะห์จำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น โดยศึกษากับข้อมูลจำลอง จากนั้นตรวจสอบความถูกต้องของการจำแนกกลุ่มความสามารถแฝงที่ได้กับข้อมูลจำลอง

4. วิเคราะห์จำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น โดยศึกษากับข้อมูลเชิงประจักษ์ ซึ่งเป็นผลการตอบข้อสอบจากการทดสอบทางการศึกษาขั้นพื้นฐาน (O-NET) วิชาคณิตศาสตร์ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับรายละเอียดของวิธีการดำเนินงานวิจัยแบ่งออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ซึ่งเป็นการบูรณาการโมเดลการวิเคราะห์กลุ่มแฝง โมเดลการตอบสนองข้อสอบ 3 พารามิเตอร์ และโมเดลพหุระดับ

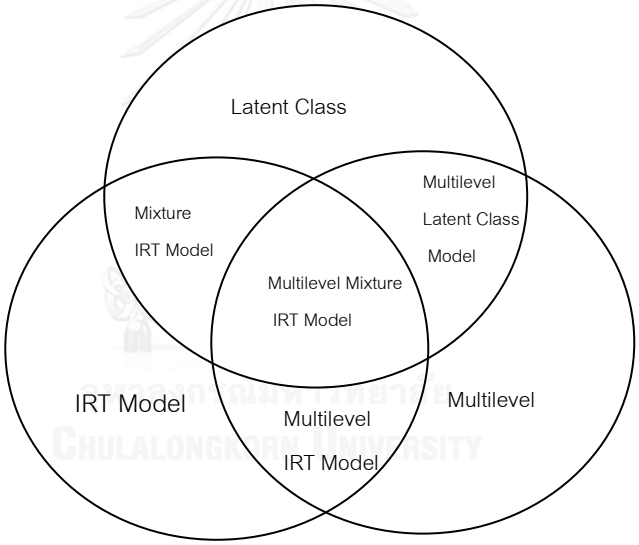
ตอนที่ 2 การศึกษากับข้อมูลจำลอง ในตอนนี้จะเป็นการวิจัยเชิงทดลอง ศึกษาโดยใช้สถานการณ์จำลอง เพื่อตรวจสอบและศึกษาความสามารถในการแบ่งกลุ่มโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น

ตอนที่ 3 การศึกษากับข้อมูลเชิงประจักษ์ ในตอนนี้เป็นการทดลองใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับกับข้อมูลจริง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตอนที่ 1 การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

โมเดลการประมาณค่าความสามารถตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ มีการบูรณาการโมเดลพหุระดับ กับโมเดลการตอบสนองข้อสอบ และโมเดลกลุ่มแฝง ซึ่งเป็นการบูรณาการโมเดลพหุระดับเข้ากับโมเดลการตอบสนองข้อสอบแบบผสม โดยการเพิ่มการประมาณค่าที่ระดับบุคคล โดยคำนึงถึงกลุ่มแฝงระดับบุคคล (individual classes) คุณสมบัติระดับบุคคล (individual traits) และ กลุ่มแฝงระดับลดหลั่น (hierarchical classes) ทั้งนี้ Cho และ Cohen (2010) และ Tay และคณะ (2011) ได้เสนอโมเดลที่คล้ายกัน แต่มีแนวคิดที่แตกต่างกันโดย Cho และ Cohen จำกัดขอบเขตที่โมเดล 1 พารามิเตอร์ (one-dimensional Rasch models) และใช้การประมาณค่าแบบเบส์เซียน (Bayesian) ส่วน Tay และคณะ จะจำกัดขอบเขตเป็นโมเดล 2 พารามิเตอร์ (multidimensional 2PL model) และใช้การประมาณค่าความน่าจะเป็นสูงสุด (maximum likelihood) สำหรับการบูรณาการทั้งสามโมเดลแสดงภาพที่ 2.3 ซึ่งแสดงอีกครั้งดังภาพต่อไปนี้



ภาพที่ 2.3 แผนภาพเวนน์แสดงการบูรณาการกันเป็นโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Cho, 2007)

สำหรับ Cho และ Cohen (2010) เสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับแบบ 1 พารามิเตอร์ (Rasch model) ดังนี้

$$P\left(y_{ijt}=1\middle|gk,\theta_{jtgk}\right)=\frac{1}{1+\exp\left[-\left(\theta_{jtgk}-\beta_{igk}\right)\right]} \qquad \text{.....(3.1)}$$

เมื่อ g = 1,... , G เป็นดัชนีสำหรับกลุ่มแฝงระดับนักเรียน

k = 1,... K เป็นดัชนีสำหรับกลุ่มแฝงระดับโรงเรียน

j = 1,...,J เป็นดัชนีสำหรับผู้สอบ

t = 1,..., T เป็นดัชนีสำหรับโรงเรียน

i = 1,...,I เป็นข้อในแบบสอบ

θ_{jtgk} เป็นความสามารถแฝงของผู้สอบ j ในโรงเรียน t และในกลุ่มแฝง g และ k

β_{igk} เป็นพารามิเตอร์ความยากของข้อสอบ i สำหรับกลุ่มแฝง g และ k

Tay และคณะ (2011) เสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบ 2 พารามิเตอร์ ซึ่งเป็นการบูรณาการโมเดลการตอบสนองข้อสอบแบบผสม (Mixture IRT model) กับข้อมูลระดับลดหลั่น แสดงดังสมการด้านล่าง อย่างไรก็ตาม ความน่าจะเป็นแบบมีเงื่อนไขขึ้นอยู่กับกลุ่มแฝงและลักษณะแฝง (latent trait) $\theta_j : P\left(y_{ji}=1\middle|k,\theta_j\right)$ ดังนั้นสำหรับกลุ่มข้อสอบ ความน่าจะเป็น $P\left(\underset{\sim}{y_j}\middle|k\right)$ ของเวกเตอร์การตอบสนองสำหรับบุคคลที่ j สามารถเขียนแทนด้วย $\int\prod_{i=1}^{I}P\left(y_{ji}\middle|k,\theta_j\right)f\left(\theta_j\right)d\theta_j$ เมื่อ $f\left(\theta_j\right)$ แทนความหนาแน่นปกติมาตรฐาน (standard normal density)

สำหรับความน่าจะเป็นในการตอบสนองที่ถูก ในโมเดลการตอบสนองข้อสอบแบบผสมแบบพหุมิติ (2PL model) แสดงดังนี้

$$P\left(y_{ji}=1\middle|k,\theta_j\right)=\frac{1}{1+\exp\left(-\eta_{jik}\right)} \qquad \text{.....(3.2)}$$

ในที่นี้ $\eta_{jik}=\beta_{ik}+\lambda_{ik}\theta_j$ ถูกยอมให้แตกต่างกันข้ามกลุ่มแฝง (Latent Class) โดยยอมให้ item intercept และ loading เป็นกลุ่มเฉพาะ (k)

และ $\theta_{jk}\sim N\left(\mu_k,\sigma_k^2\right)$ เมื่อ j หมายถึง นักเรียนคนที่ j ซึ่ง j = 1, ..., J และ k หมายถึงกลุ่มแฝงในระดับนักเรียน μ_k แทนค่าเฉลี่ยของความสามารถของชั้นที่ k และ σ_k^2 หมายถึงความแปรปรวนของความสามารถสำหรับชั้นที่ k

จากทั้งสองโมเดลข้างต้นสามารถดูรายละเอียดเพิ่มเติมได้ในบทที่ 2 ในงานวิจัยนี้ผู้วิจัยได้พัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบ 3 พารามิเตอร์ โดยพัฒนาจาก โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ แบบ 2 พารามิเตอร์ ของ Tay และคณะ (2011) เพื่อนำมาประยุกต์ใช้ในทางการศึกษา เนื่องจากการทดสอบทางการศึกษาโดยส่วนใหญ่เป็นข้อสอบแบบเลือกตอบ ทำให้การเดาข้อสอบส่งผลให้เกิดความแตกต่างของโมเดลทางการวัดของแต่ละบุคคล ผู้วิจัยจึงพัฒนาโมเดลโดยบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และเสนอโมเดลดังรายละเอียดในบทที่ 4

สำหรับการประมาณค่าในที่นี้ผู้วิจัยใช้การประมาณค่าแบบเอ็มไพริคอล เบส์ (Empirical Bayes) และพิจารณาความเหมาะสมของการประมาณค่าความน่าจะเป็นสูงสุด (maximum likelihood) และการวิเคราะห์ข้อมูลโดยใช้แพ็คเกจ ltm และ MCLUST ในโปรแกรม R ซึ่งมีขั้นตอนในการวิเคราะห์ข้อมูล 5 ขั้นตอน ดังนี้

1. การประมาณค่าพารามิเตอร์ ได้แก่ discriminant (a), difficulty(b), guessing(c) และ ability(θ) ของรูปแบบการตอบข้อสอบ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ของแต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ ตามโมเดลการตอบสนองข้อสอบ

2. จำแนกกลุ่มความสามารถแฝงระดับบุคคล (Individual Class) โดยนำความน่าจะเป็นในขั้นตอนที่ 1 มาวิเคราะห์กลุ่มแฝง ซึ่งเป็นการจำแนกกลุ่มโดยให้บุคคลที่มีความน่าจะเป็นในการตอบข้อสอบถูกในแต่ละข้อที่มีค่าใกล้เคียงกันอยู่กลุ่มเดียวกัน ซึ่งจะสามารถจำแนกได้ว่าแต่ละบุคคลอยู่ในกลุ่มแฝงใด

3. คำนวณหาความสามารถของโรงเรียนโดยใช้สัดส่วนความน่าจะเป็นของจำนวนกลุ่มแฝงระดับบุคคลในแต่ละโรงเรียน ซึ่งสัดส่วนความน่าจะเป็นนี้คำนวณจากจำนวนบุคคลในแต่ละกลุ่มแฝงที่จำแนกในระดับบุคคลเทียบกับจำนวนคนทั้งหมดในโรงเรียน

4. จำแนกกลุ่มความสามารถแฝงระดับโรงเรียน (School Class) โดยนำความสามารถของโรงเรียนในขั้นตอนที่ 3 มาวิเคราะห์กลุ่มแฝง ซึ่งจำแนกกลุ่มความสามารถแฝงโดยให้โรงเรียนที่มีสัดส่วนความสามารถที่มีค่าใกล้เคียงกันอยู่กลุ่มเดียวกัน

5. คำนวณหาความน่าจะเป็นในการตอบถูกของแต่ละโรงเรียน

Chulalongkorn University

ตอนที่ 2 การศึกษากับข้อมูลจำลอง

งานวิจัยในส่วนนี้ผู้วิจัยได้ทำการตรวจสอบความถูกต้องของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น กับข้อมูลที่จำลองขึ้น

สำหรับการจำลองข้อมูลมีรายละเอียดดังต่อไปนี้

การออกแบบการจำลองข้อมูล

ผู้วิจัยจำลองข้อมูลด้วยโปรแกรม R Version 3.2.3 แพคเกจ IRTOYS Version 0.1.7 ในงานวิจัยนี้ได้จำลองผลการตอบข้อสอบของนักเรียนเป็นแบบให้คะแนนสองค่า คือ 0 และ 1 ตามทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3-PL Item Response Theory) โดยจำลองกลุ่มของโรงเรียนและกำหนดค่าเริ่มต้นของพารามิเตอร์ในการจำลองรูปแบบการตอบ ดังนี้

การจำลองกลุ่มแฝงของโรงเรียน ผู้วิจัยจำลองรูปแบบการตอบของแต่ละโรงเรียนตามความสามารถของนักเรียนในแต่ละโรงเรียนที่มีความเชี่ยวชาญความรู้ในสาระที่ต่างกัน รวมทั้งขนาดของโรงเรียนที่ต่างกัน ซึ่งในที่นี้ผู้วิจัยจำลองกลุ่มแฝงของโรงเรียนออกเป็น 6 กลุ่มแฝงตามความสามารถในการตอบข้อสอบ 3 สาระ เพื่อให้สอดคล้องกับสภาพจริงและมีความครอบคลุม ผู้วิจัยจึงจำลองให้มีกลุ่มแฝงที่มีความสามารถแบบคละความสามารถ 3 กลุ่มแฝง และมีความสามารถปานกลางในทุกสาระ 1 กลุ่มแฝง ความสามารถสูงในทุกสาระ 1 กลุ่มแฝง และความสามารถอ่อนในทุกสาระ 1 กลุ่มแฝง ดังนี้ กลุ่มแฝงที่ 1 มีความสามารถสูงในสาระที่ 1 มีความสามารถปานกลางในสาระที่ 2 มีความสามารถระดับต่ำในสาระที่ 3 กลุ่มแฝงที่ 2 มีความสามารถปานกลางในสาระที่ 1 มีความสามารถต่ำในสาระที่ 2 มีความสามารถระดับสูงในสาระที่ 3 กลุ่มแฝงที่ 3 มีความสามารถต่ำในสาระที่ 1 มีความสามารถสูงในสาระที่ 2 มีความสามารถระดับปานกลางในสาระที่ 3 กลุ่มแฝงที่ 4 มีความสามารถปานกลางในทั้ง 3 สาระ กลุ่มแฝงที่ 5 มีความสามารถสูงในทั้ง 3 สาระ และกลุ่มแฝงที่ 6 มีความสามารถต่ำในทั้ง 3 สาระ ในการจำลองข้อมูลนี้แต่ละกลุ่มแฝงจะจำลองโรงเรียนเป็น 4 ขนาด คือ ขนาดเล็ก ขนาดกลาง ขนาดใหญ่ และขนาดใหญ่พิเศษ โดยจำลองให้แต่ละกลุ่มแฝงมีจำนวนนักเรียนในแต่ละขนาดโรงเรียนไม่เท่ากัน เพื่อให้เกิดความผันแปรภายในกลุ่ม รวมจำลองข้อมูลทั้งสิ้น 32,370 คน ดังนั้นในงานวิจัยนี้จำลองข้อมูลทั้งหมด 24 โรงเรียน นอกจากนี้ผู้วิจัยจำลองผลการตอบข้อสอบ 2 สถานการณ์ ได้แก่ ผลการตอบข้อสอบจำนวน 6 ข้อ ซึ่งแต่ละสาระถูกวัดด้วยข้อสอบ 2 ข้อ และผลการตอบข้อสอบจำนวน 15 ข้อ แต่ละสาระถูกวัดด้วยข้อสอบ 5 ข้อ

ผู้วิจัยกำหนดเงื่อนไขการจำลองข้อมูลเป็นผลการตอบข้อสอบจำนวนข้อ 6 ข้อ และ 15 ข้อ แทนข้อสอบขนาดสั้น และปานกลาง โดยในเริ่มต้นผู้วิจัยนำโมเดลที่พัฒนาขึ้นไปใช้กับผลการตอบข้อสอบที่มีจำนวนข้อน้อย เนื่องจากโมเดลที่พัฒนาขึ้นเป็นโมเดลที่มีความซับซ้อนในการประมาณค่าทำให้ใช้เวลาในการวิเคราะห์แต่ละเงื่อนไขค่อนข้างนาน จึงได้ทดลองใช้โมเดลกับจำนวน 6 ข้อ เมื่อโมเดลมีความสมบูรณ์สามารถใช้ในการวิเคราะห์จำแนกกลุ่มความสามารถแฝงได้ถูกต้อง ผู้วิจัยจึงได้นำโมเดลไปใช้กับผลการตอบข้อสอบ 15 ข้อ เพื่อเป็นการยืนยันว่าโมเดลที่พัฒนาขึ้นสามารถใช้จำแนกกลุ่มความสามารถแฝงกับสถานการณ์อื่นได้ถูกต้องเช่นกัน ในงานวิจัยนี้ผู้วิจัยไม่ได้ทดสอบโมเดลกับสถานการณ์ที่มีข้อสอบจำนวนที่มากกว่านี้ เนื่องจากมีข้อจำกัดเกี่ยวกับประสิทธิภาพของเครื่องคอมพิวเตอร์ ดังนั้นถ้าต้องการทดสอบกับสถานการณ์จำนวนข้อสอบที่มากกว่า 15 ข้อ ควรใช้คอมพิวเตอร์ที่มีประสิทธิภาพสูงขึ้น

ดังนั้นในที่นี้จำลองผลการตอบข้อสอบสำหรับความสามารถแบบกลุ่มที่ 1 สำหรับโรงเรียนที่ 1 ถึง 4 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มาก ความสามารถแบบกลุ่มที่ 2 สำหรับโรงเรียนที่ 5 ถึง 8 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มาก ความสามารถแบบกลุ่มที่ 3 สำหรับโรงเรียนที่ 9 ถึง 12 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มาก ความสามารถแบบ

กลุ่มที่ 4 สำหรับโรงเรียนที่ 13 ถึง 16 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มาก ความสามารถแบบกลุ่มที่ 5 สำหรับโรงเรียนที่ 17 ถึง 20 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มาก ความสามารถแบบกลุ่มที่ 6 สำหรับโรงเรียนที่ 21 ถึง 24 ซึ่งมีขนาดโรงเรียนเป็น เล็ก กลาง ใหญ่ และใหญ่มากดังตารางที่ 3.1 และ 3.2

ตารางที่ 3.1 กำหนดการจำลองกลุ่มแฝงของโรงเรียนจำนวน 24 โรงเรียน

กลุ่มความสามารถแฝงที่จำลอง	โรงเรียน	ขนาดโรงเรียน
กลุ่มแฝงที่ 1	1	เล็ก
	2	กลาง
	3	ใหญ่
	4	ใหญ่มาก
กลุ่มแฝงที่ 2	5	เล็ก
	6	กลาง
	7	ใหญ่
	8	ใหญ่มาก
กลุ่มแฝงที่ 3	9	เล็ก
	10	กลาง
	11	ใหญ่
	12	ใหญ่มาก
กลุ่มแฝงที่ 4	13	เล็ก
	14	กลาง
	15	ใหญ่
	16	ใหญ่มาก
กลุ่มแฝงที่ 5	17	เล็ก
	18	กลาง
	19	ใหญ่
	20	ใหญ่มาก
กลุ่มแฝงที่ 6	21	เล็ก
	22	กลาง
	23	ใหญ่
	24	ใหญ่มาก

สำหรับการกำหนดกลุ่มแฝง และค่าเริ่มต้นสำหรับนักเรียนที่มีความสามารถแตกต่างกัน ในการจำลองข้อมูล มีรายละเอียดดังนี้ การกำหนดกลุ่มแฝงของโรงเรียนเพื่อให้มีความครอบคลุม และสอดคล้องกับสภาพจริงในบริบทของโรงเรียนมากที่สุด ในที่นี้จึงกำหนดกลุ่มแฝงจำนวน 6 กลุ่มแฝง ประกอบด้วย กลุ่มที่คละความสามารถ 3 กลุ่ม กลุ่มที่มีความสามารถปานกลางในทุกสาระ 1 กลุ่ม

กลุ่มที่มีความสามารถเก่งทุกสาระ 1 กลุ่ม และกลุ่มที่มีความสามารถอ่อนทุกสาระ 1 กลุ่ม และในที่นี้ผู้วิจัยกำหนดค่าเริ่มต้น ค่าความยากของข้อสอบ ค่าอำนาจจำแนก พารามิเตอร์การเดา สำหรับแต่ละกลุ่มแฝง และจำลองความสามารถภายใต้โค้งปกติ N(0,1) ซึ่งการกำหนดค่าเริ่มต้นสอดคล้องกับ Li et al. (2009) แสดงดังตารางที่ 3.2 3.3 และ 3.4

ตารางที่ 3.2 การจำลองกลุ่มแฝงของโรงเรียน 4 กลุ่มจำแนกตามความสามารถในการตอบคำถาม

สาระ	ข้อ	ความสามารถที่จำลอง					
		กลุ่มแฝงที่ 1	กลุ่มแฝงที่ 2	กลุ่มแฝงที่ 3	กลุ่มแฝงที่ 4	กลุ่มแฝงที่ 5	กลุ่มแฝงที่ 6
1	1	เก่ง	ปานกลาง	อ่อน	ปานกลาง	เก่ง	อ่อน
	2	เก่ง	ปานกลาง	อ่อน	ปานกลาง	เก่ง	อ่อน
2	3	ปานกลาง	อ่อน	เก่ง	ปานกลาง	เก่ง	อ่อน
	4	ปานกลาง	อ่อน	เก่ง	ปานกลาง	เก่ง	อ่อน
3	5	อ่อน	เก่ง	ปานกลาง	ปานกลาง	เก่ง	อ่อน
	6	อ่อน	เก่ง	ปานกลาง	ปานกลาง	เก่ง	อ่อน

สำหรับการจำลองกลุ่มแฝงของโรงเรียนออกเป็น 6 กลุ่มแฝงตามความสามารถในการตอบข้อสอบ กำหนดเป็น 3 สาระ สำหรับผลการตอบข้อสอบ 6 ข้อ แต่ละสาระถูกวัดด้วยข้อสอบจำนวน 2 ข้อ และผลการตอบข้อสอบ 15 ข้อ แต่ละสาระถูกวัดด้วยข้อสอบจำนวน 5 ข้อ ซึ่งจำลองกลุ่มแฝงในลักษณะเดียวกับการจำลอง 6 ข้อ

สำหรับการจำลอง 6 ข้อ กลุ่มที่ 1 เป็นกลุ่มคละความสามารถโดยมีนักเรียนที่มีความสามารถเก่งในสาระที่ 1 มีความสามารถปานกลางในสาระที่ 2 และมีความสามารถอ่อนในสาระที่ 3 นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 และ 2 สูง ข้อที่ 3 และ 4 ปานกลาง ข้อที่ 5 และ 6 ต่ำ

กลุ่มที่ 2 เป็นกลุ่มคละความสามารถโดยมีนักเรียนที่มีความสามารถปานกลางในสาระที่ 1 มีความสามารถอ่อนในสาระที่ 2 และมีความสามารถเก่งในสาระที่ 3 นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 และ 2 ปานกลาง ข้อที่ 3 และ 4 ต่ำ ข้อที่ 5 และ 6 สูง

กลุ่มที่ 3 เป็นกลุ่มคละความสามารถโดยมีนักเรียนที่มีความสามารถอ่อนในสาระที่ 1 มีความสามารถเก่งในสาระที่ 2 และมีความสามารถปานกลางในสาระที่ 3 นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 และ 2 ต่ำ ข้อที่ 3 และ 4 สูง ข้อที่ 5 และ 6 ปานกลาง

กลุ่มที่ 4 เป็นกลุ่มความสามารถปานกลางทุกสาระ นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 ถึง 6 ปานกลาง

กลุ่มที่ 5 เป็นกลุ่มความสามารถเก่งทุกสาระ นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 ถึง 6 สูง

กลุ่มที่ 6 เป็นกลุ่มความสามารถอ่อนทุกสาระ นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในข้อที่ 1 ถึง 6 ต่ำ

ตารางที่ 3.3 การกำหนดค่าพารามิเตอร์ในการจำลองรูปแบบการตอบ จำนวน 6 ข้อ

ข้อ	กลุ่มที่ 1			กลุ่มที่ 2			กลุ่มที่ 3			กลุ่มที่ 4			กลุ่มที่ 5			กลุ่มที่ 6		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
1	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2.5	0.1	1	2	0.25
2	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2	0.1	1	2.5	0.25
3	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2.5	0.1	1	2	0.25
4	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2	0.1	1	2.5	0.25
5	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2.5	0.1	1	2	0.25
6	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2	0.1	1	2.5	0.25

จากตารางที่ 3.3 การกำหนดค่าพารามิเตอร์ในการจำลองรูปแบบการตอบข้อสอบ จำนวน 6 ข้อ ค่าพารามิเตอร์อำนาจจำแนก (a) ถูกกำหนดเป็น 1 สำหรับกลุ่มความสามารถปานกลางและอ่อน และ 2 สำหรับกลุ่มความสามารถเก่ง ค่าพารามิเตอร์การเดา (c) ถูกกำหนดเป็น 0.25 0.2 และ 0.1 สำหรับข้อที่มีค่าความยากเป็น ยากมาก ปานกลาง และ ง่าย ตามลำดับ ซึ่งแสดงให้เห็นว่าถ้าข้อสอบยากจะมีการเดามากกว่าข้อสอบที่ง่าย ค่าพารามิเตอร์ความยาก (b) ถูกกำหนดเป็นค่าลบ สำหรับข้อสอบที่ง่าย และเป็นจำนวนบวกสำหรับข้อสอบที่ยาก

เมื่อพิจารณาเป็นรายกลุ่มแฝง กลุ่มแฝงที่ 1, 2 และ 3 เป็นกลุ่มที่คละความสามารถ กลุ่มที่เก่งในการตอบข้อสอบแต่ละข้อ ผู้วิจัยกำหนดค่าความยากเท่ากับ -1.5 มีค่าอำนาจจำแนกเท่ากับ 2 มีการเดาเท่ากับ 0.1 ซึ่งจำแนกได้ดีมากสำหรับนักเรียนกลุ่มเก่ง และมีค่าการเดาที่ต่ำ กลุ่มที่มีความสามารถอ่อนในการตอบข้อสอบแต่ละข้อ ผู้วิจัยกำหนดค่าความยากเท่ากับ 1.5 ค่าอำนาจจำแนกเท่ากับ 1 และมีการเดาเท่ากับ 0.25 ซึ่งจำแนกได้ดีกับนักเรียนที่อ่อน และมีค่าการเดาที่มากขึ้น กลุ่มที่มีความสามารถปานกลางในการตอบข้อสอบแต่ละข้อ ผู้วิจัยกำหนดค่าความยากเท่ากับ 0 ค่าอำนาจจำแนกเท่ากับ 1 และมีการเดาเท่ากับ 0.2 ซึ่งจำแนกได้ดีกับนักเรียนที่มีความสามารถปานกลางอ่อน และมีค่าการเดาปานกลาง

เมื่อพิจารณากลุ่มแฝงที่ 4 เป็นกลุ่มที่มีความสามารถระดับปานกลาง ผู้วิจัยกำหนดค่าความยากเป็น 0.25, -0.25 และ 0 ซึ่งเป็นค่าที่ใกล้ 0 แสดงถึงข้อสอบที่มีความยากปานกลาง โดยกำหนดค่าแตกต่างกันเพื่อให้มีความแตกต่างของความยากของข้อสอบ ซึ่งจะเห็นได้ว่าข้อสอบ ข้อที่ 1 และ 2 เป็นข้อที่ยากที่สุด ข้อที่ 3 และ 4 เป็นข้อที่ง่ายที่สุด

เมื่อพิจารณากลุ่มแฝงที่ 5 เป็นกลุ่มที่มีความสามารถเก่ง ผู้วิจัยจึงกำหนดให้ข้อสอบมีความยากเท่ากับ -2.5 และ -2 และ ค่าอำนาจจำแนกเท่ากับ 2 การเดาเท่ากับ 0.1 ซึ่งแสดงให้เห็นว่าข้อสอบทั้ง 6 ข้อ เป็นข้อสอบที่ง่ายสำหรับนักเรียนกลุ่มนี้

เมื่อพิจารณากลุ่มแฝงที่ 6 เป็นกลุ่มที่มีความสามารถอ่อน ผู้วิจัยจึงกำหนดให้ข้อสอบมีความยากเท่ากับ 2.5 และ 2 และ ค่าอำนาจจำแนกเท่ากับ 1 การเดาเท่ากับ 0.25 ซึ่งแสดงให้เห็นว่าข้อสอบทั้ง 6 ข้อ เป็นข้อสอบที่ยากสำหรับนักเรียนกลุ่มนี้

ตารางที่ 3.4 การกำหนดค่าพารามิเตอร์ในการจำลองรูปแบบการตอบ จำนวน 15 ข้อ

ข้อ	กลุ่มแฝงที่ 1			กลุ่มแฝงที่ 2			กลุ่มแฝงที่ 3			กลุ่มแฝงที่ 4			กลุ่มแฝงที่ 5			กลุ่มแฝงที่ 6		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
1	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2	0.1	1	2.5	0.25
2	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2.5	0.1	1	2	0.25
3	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2	0.1	1	2.5	0.25
4	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2.5	0.1	1	2	0.25
5	2	-1.5	0.1	1	0	0.2	1	1.5	0.25	1	0.25	.2	2	-2	0.1	1	2.5	0.25
6	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2.5	0.1	1	2	0.25
7	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2	0.1	1	2.5	0.25
8	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2.5	0.1	1	2	0.25
9	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2	0.1	1	2.5	0.25
10	1	0	0.2	1	1.5	0.25	2	-1.5	0.1	1	-0.25	.2	2	-2.5	0.1	1	2	0.25
11	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2	0.1	1	2.5	0.25
12	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2.5	0.1	1	2	0.25
13	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2	0.1	1	2.5	0.25
14	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2.5	0.1	1	2	0.25
15	1	1.5	0.25	2	-1.5	0.1	1	0	0.2	1	0	.2	2	-2	0.1	1	2.5	0.25

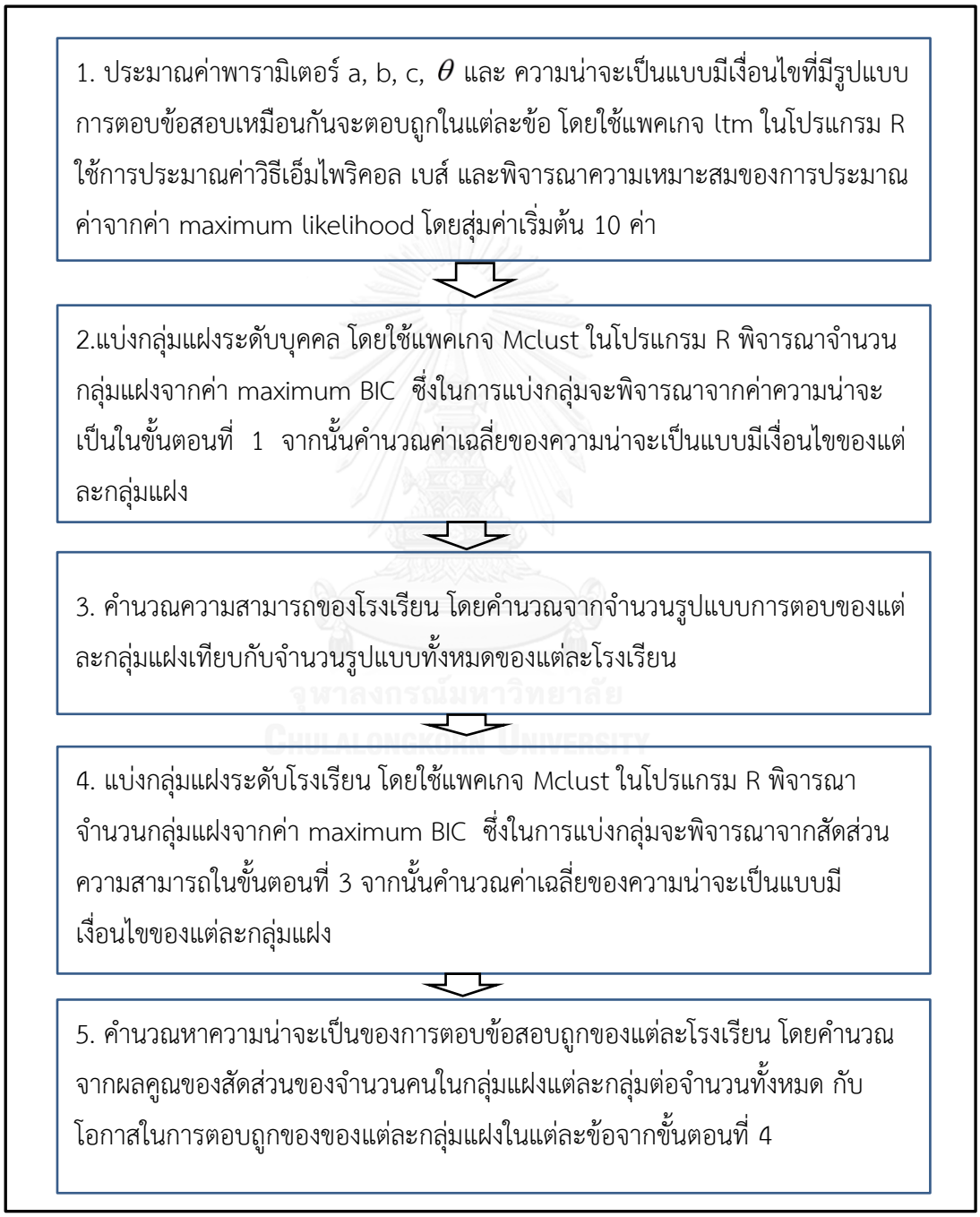
ในการจำลองจำนวนนักเรียนในแต่ละโรงเรียนตามขนาดของโรงเรียนผู้วิจัยจำลองให้แต่ละกลุ่มแฝงมีจำนวนนักเรียนในโรงเรียนแต่ละขนาดไม่เท่ากัน 4 ขนาด โดยในแต่ละกลุ่มแฝงจะประกอบด้วยโรงเรียนที่มีขนาดเล็ก กลาง ใหญ่ และใหญ่พิเศษ ในที่นี้ผู้วิจัยกำหนดจำนวนคนในกลุ่มแฝงที่ 3 ให้มีจำนวนมากกว่ากลุ่มแฝงที่ 1 และ 2 ซึ่งเป็นกลุ่มแฝงที่คละความสามารถเช่นเดียวกัน เพื่อให้มีความแตกต่างของจำนวนผู้ตอบข้อสอบถูกในแต่ละข้อ ดังตารางที่ 3.5

ตารางที่ 3.5 จำนวนนักเรียนที่จำลองขึ้นสำหรับแต่ละกลุ่มแฝงของแต่ละขนาดโรงเรียน

ขนาดโรงเรียน	กลุ่มแฝงที่ 1	กลุ่มแฝงที่ 2	กลุ่มแฝงที่ 3	กลุ่มแฝงที่ 4	กลุ่มแฝงที่ 5	กลุ่มแฝงที่ 6
ขนาดที่ 1	120	120	360	150	120	300
ขนาดที่ 2	400	400	1,200	500	400	1,000
ขนาดที่ 3	800	800	2,400	1,000	800	2,000
ขนาดที่ 4	2,000	2,000	6,000	2,500	2,000	5,000

การวิเคราะห์ข้อมูลจำลอง

การตรวจสอบความถูกต้องของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้นกับข้อมูลที่จำลอง ใช้การประมาณค่าพารามิเตอร์ และการวิเคราะห์จำแนกกลุ่มความสามารถแฝงด้วยโปรแกรม R version 3.2.3 โดยใช้แพคเกจ ltm Version 1.0-0 และ Mclust Version 5.0.1 สำหรับขั้นตอนการวิเคราะห์ข้อมูลโดยละเอียดแสดงในบทที่ 4 ขั้นตอนโดยสรุปแสดงดังภาพต่อไปนี้



ภาพที่ 3.1 ขั้นตอนการวิเคราะห์ข้อมูล

การตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝง

ผู้วิจัยตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝงเทียบกับข้อมูลที่จำลองในเบื้องต้น โดยตรวจสอบว่า ในการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน โรงเรียนที่ถูกจำแนกกลุ่มจากการใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น ตรงกับกับกลุ่มแฝงที่ผู้วิจัยจำลองขึ้นตามเงื่อนไขการจำลองข้อมูลเป็น 6 กลุ่มแฝง ซึ่งแต่ละกลุ่มแฝงประกอบด้วยโรงเรียน 4 โรงเรียน หรือไม่ โดยตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝงของการจำลองสถานการณ์การตอบข้อสอบ 2 สถานการณ์ คือ ผลการตอบข้อสอบ 6 ข้อ และ 15 ข้อ โดยการวิเคราะห์ซ้ำแต่ละสถานการณ์ จำนวน 25 ครั้ง สำหรับการพิจารณาจำนวนกลุ่มแฝง ผู้วิจัยพิจารณาจากค่า BIC

ตอนที่ 3 การศึกษากับข้อมูลเชิงประจักษ์

การศึกษากับข้อมูลเชิงประจักษ์จะใช้ข้อมูลทุติยภูมิ ผลการตอบข้อสอบวิชาคณิตศาสตร์ของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.) มีการกำหนดประชากรและกลุ่มตัวอย่างดังนี้

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในงานวิจัยเป็นนักเรียนระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 4 ระดับชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2555 จำนวน 341,691 คน ใน 3,995 โรงเรียน 4 สังกัด ดังตาราง

ตารางที่ 3.6 จำนวนประชากรจำแนกตามสังกัดและขนาดของโรงเรียน

สังกัด	จำนวนโรงเรียนในสังกัด				จำนวนนักเรียน
	เล็ก	กลาง	ใหญ่	ใหญ่มาก	
กรมส่งเสริมการปกครองท้องถิ่น	70	164	49	19	16,568
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน	516	1,161	428	392	279,979
สำนักกรรมการอุดมศึกษา	2	9	10	4	3,712
สำนักบริหารงานคณะกรรมการส่งเสริมการศึกษาเอกชน	121	216	109	125	41,432
รวม	709	1,550	596	540	341,691

กลุ่มตัวอย่างในการวิจัย

เป็นข้อมูลทุติยภูมิจากฐานข้อมูล (data base) ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.) ปี 2555 ในวิชาคณิตศาสตร์ ซึ่งกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ที่ครอบคลุมสถานศึกษาทั้ง 4 สังกัด กลุ่มตัวอย่างที่ใช้ในการวิจัยได้มาจากการสุ่มแบบแบ่งชั้น

(stratified random sampling) โดยสังกัดของโรงเรียนเป็นระดับชั้น และนักเรียนเป็นหน่วยของการสุ่ม (sampling unit) โดยแบ่ง โรงเรียนเป็น 4 กลุ่ม คือ สังกัดกรมการปกครองส่วนท้องถิ่น (เทศบาล/ ท้องถิ่น) สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ) สังกัดสำนักงานคณะกรรมการการอุดมศึกษา (สาธิต) และสังกัดสำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)

การกำหนดขนาดกลุ่มตัวอย่าง

เนื่องจากการศึกษาวิจัยนี้เป็นการศึกษากับข้อมูลที่มีลักษณะสอดแทรกเป็นระดับลดหลั่น ซึ่งแต่เดิมไม่มีการกำหนดขนาดกลุ่มตัวอย่างที่แน่นอนไว้ นักวิจัยมักใช้การกำหนดขนาดกลุ่มตัวอย่างเช่นเดียวกับการวิเคราะห์ข้อมูลแบบปกติ ระดับเดียว (single level) แต่จากการศึกษาของ Mass และ Hox (2004, 2005) Cohen (2002) Afshartous และ Leeuw (2005) Snijders และ Bosker (1999) และ ศิริชัย กาญจนวาสี (2548) ได้ข้อค้นพบที่สอดคล้องกันว่าควรให้ความสนใจต่อขนาดกลุ่มตัวอย่างในกลุ่ม หรือระดับสูงในการวิเคราะห์ (group-level sample size) มากกว่าระดับบุคคลหรือระดับที่ต่ำสุด (level-1) เพราะจะเป็นการลดความคลาดเคลื่อนในการประมาณค่า เพิ่มความแม่นยำในการประมาณค่าพารามิเตอร์ นอกจากนี้จากผลการศึกษาของ Snijders และ Bosker (1999) พบว่าได้มีข้อเสนอเกี่ยวกับการกำหนดขนาดกลุ่มตัวอย่างโดยกลุ่มระดับสูงที่สุดของการวิเคราะห์ควรมีจำนวนมากกว่า 10 กลุ่มขึ้นไป ส่วนจากการศึกษาของ Mass และ Hox (2004, 2005) มีความเห็นว่าจำนวนกลุ่มตัวอย่างระดับกลุ่มที่สูงสุดของการวิเคราะห์ควรจะมีขนาดที่มากกว่าหรือเท่ากับ 30 กลุ่มขึ้นไป และสมาชิกในแต่ละกลุ่มย่อยควรจะมีอย่างน้อย 30 คน ขึ้นไปด้วยเช่นกัน จึงจะทำให้การประมาณค่ามีความแม่นยำ ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงกำหนดขนาดกลุ่มตัวอย่างตามแนวคิดของ Mass และ Hox (2004, 2005) ที่ให้มีกลุ่มหน่วยการวิเคราะห์มากกว่าหรือเท่ากับ 30 หน่วยขึ้นไป และดำเนินการสุ่มตัวอย่างระดับโรงเรียน ใช้วิธีการสุ่มโรงเรียนแบบแบ่งชั้น (stratified random sampling) โดยในทุกสังกัดของโรงเรียน จะใช้ขนาดของโรงเรียนเป็นชั้นของการสุ่ม 4 ขนาด คือ ขนาดเล็ก (มีจำนวนนักเรียนน้อยกว่า 300 คน) ขนาดกลาง (มีจำนวนนักเรียนตั้งแต่ 300 คนขึ้นไปแต่ไม่ถึง 1,000 คน) ขนาดใหญ่ (มีจำนวนนักเรียนตั้งแต่ 1,000 คนขึ้นไป แต่ไม่ถึง 2,000 คน) และขนาดใหญ่มาก (มีจำนวนตั้งแต่ 2,000 คนขึ้นไป) ดำเนินการสุ่มโรงเรียนของแต่ละขนาดด้วยวิธีการสุ่มตัวอย่างอย่างง่าย (simple random sampling) ตามสัดส่วนโดยมีเงื่อนไขในการกำหนดคุณลักษณะของโรงเรียนจนได้กลุ่มตัวอย่างดังนี้

ตารางที่ 3.7 จำนวนกลุ่มตัวอย่างจำแนกตามสังกัดและขนาดของโรงเรียน

สังกัด	จำนวนโรงเรียนที่สุ่มจำแนกตามขนาด				จำนวน โรงเรียนที่สุ่ม	จำนวน นักเรียนที่สุ่ม
	เล็ก	กลาง	ใหญ่	ใหญ่มาก		
กรมส่งเสริมการปกครองท้องถิ่น	9	20	6	3	38	4,127
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน	51	141	52	48	291	56,582
สำนักกรรมการอุดมศึกษา	1	2	2	1	6	1,030
สำนักบริหารงานคณะกรรมการส่งเสริมการศึกษาเอกชน	9	26	14	15	64	10,556
รวม	70	189	74	67	400	72,295

เครื่องมือที่ใช้ในการวิจัย

แบบสอบที่ใช้ในการศึกษา เป็นแบบสอบจากการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ซึ่งเป็นการทดสอบเพื่อวัดความรู้และความคิดของนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 6 (ปี 2555) ประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ซึ่งในการวิจัยครั้งนี้ศึกษาเฉพาะวิชาคณิตศาสตร์ ซึ่งเป็นข้อสอบแบบเลือกตอบ 5 ตัวเลือก มีจำนวนข้อสอบ 32 ข้อ จัดเป็นข้อสอบที่ตรวจให้คะแนนสองค่า ข้อสอบจะครอบคลุม 4 สาระการเรียนรู้ ได้แก่ สาระที่ 1 จำนวนและการดำเนินการ ครอบคลุม 4 มาตรฐานการเรียนรู้ สาระที่ 2 การวัด ครอบคลุม 2 มาตรฐานการเรียนรู้ สาระที่ 3 พีชคณิต ครอบคลุม 2 มาตรฐานการเรียนรู้ และสาระที่ 4 การวิเคราะห์ข้อมูลและความน่าจะเป็น ครอบคลุม 3 มาตรฐานการเรียนรู้ โดยในงานวิจัยนี้ผู้วิจัยได้ใช้การสุ่มอย่างง่ายในการเลือกข้อสอบแบบเลือกตอบจำนวน 15 ข้อ ซึ่งครอบคลุมทั้ง 4 สาระการเรียนรู้ มาใช้ในหากลุ่มแฝงในระดับโรงเรียน ดังตารางต่อไปนี้

CHULALONGKORN UNIVERSITY

ตารางที่ 3.8 ลำดับข้อสอบของข้อมูลเชิงประจักษ์ และและข้อสอบที่เลือกใช้ในการจำแนกกลุ่มความสามารถแฝง

สาระ/ มาตรฐานการเรียนรู้	ช่วงชั้นที่ 4		ลำดับข้อ	ข้อสอบที่เลือก
	ข้อ	คะแนน		
สาระที่ 1 จำนวนและการดำเนินการ				
มาตรฐาน ค 1.1 เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง	3	7.5	1, 2, 3, 4, 5, 8, 9, 34	2, 4, 8, 9
มาตรฐาน ค 1.2 เข้าใจถึงผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่างๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้	1	2.5		
มาตรฐาน ค 1.3 ใช้การประมาณค่าในการคำนวณและแก้ปัญหาได้	2	5.0		
มาตรฐาน ค 1.4 เข้าใจในระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้	2	5.0		
สาระที่ 2 การวัด				
มาตรฐาน ค 2.1 เข้าใจพื้นฐานเกี่ยวกับการวัด	3	7.5	16, 17, 18. 19	16, 18
มาตรฐาน ค 2.2 วัดและคาดคะเนขนาดของสิ่งที่ต้องการวัดได้	1	2.5		
มาตรฐาน ค 2.3 แก้ปัญหาเกี่ยวกับการวัดได้				
สาระที่ 3 เรขาคณิต				
มาตรฐาน ค 3.1 อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติได้	-	-	-	-
มาตรฐาน ค 3.2 ใช้การนึกภาพ (visualization) ใช้เหตุผลเกี่ยวกับปริภูมิ (spatial reasoning)และใช้แบบจำลองทางเรขาคณิต(geometric model) ในการแก้ปัญหาได้	-	-		
สาระที่ 4 พีชคณิต				
มาตรฐาน ค 4.1 อธิบายและวิเคราะห์แบบรูป (pattern) ความสัมพันธ์ และฟังก์ชันต่าง ๆ ได้	9	22.5	6, 7, 10, 11, 12, 13, 14, 15, 20, 21, 22, 23, 24, 25, 26, 33, 35, 36, 38	6, 10, 13, 15, 20, 22
มาตรฐาน ค 4.2 ใช้นิพจน์ สมการ อสมการ กราฟ และแบบจำลองทางคณิตศาสตร์อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมายและนำไปใช้แก้ปัญหาได้	10	25.0		
สาระที่ 5 การวิเคราะห์ข้อมูลและความน่าจะเป็น				
มาตรฐาน ค 5.1 เข้าใจและใช้วิธีการทางสถิติในการวิเคราะห์ข้อมูลได้	6	15.0	27, 28, 29, 30, 31, 32, 39, 40, 37	27, 29, 31
มาตรฐาน ค 5.2 ใช้วิธีการทางสถิติและความรู้เกี่ยวกับความน่าจะเป็นในการคาดการณ์ได้อย่างสมเหตุสมผล	1	2.5		
มาตรฐาน ค 5.3 ใช้ความรู้เกี่ยวกับสถิติและความน่าจะเป็นช่วยในการตัดสินใจ และแก้ปัญหาได้	2	5.0		
สาระที่ 6 ทักษะกระบวนการทางคณิตศาสตร์				
มาตรฐาน ค 6.1 มีความสามารถในการแก้ปัญหา			-	-
มาตรฐาน ค 6.2 มีความสามารถในการให้เหตุผล	-	-		
มาตรฐาน ค 6.3 มีความสามารถในการสื่อสาร การสื่อความหมายทางคณิตศาสตร์ และการนำเสนอ	-	-		
มาตรฐาน ค 6.4 มีความสามารถในการเชื่อมโยงความรู้ต่างๆ ทางคณิตศาสตร์และเชื่อมโยงคณิตศาสตร์กับศาสตร์อื่นๆ ได้	-	-		
มาตรฐาน ค 6.5 มีความคิดริเริ่มสร้างสรรค์	-	-		
รวม	40	100	40	15

ตัวแปรที่ศึกษา

ผลการตอบข้อสอบที่มีการให้คะแนนแบบสองค่า (0,1) โดยศึกษา 2 สถานการณ์ คือ ผลการตอบจำนวน 6 ข้อ และ 15 ข้อ

การวิเคราะห์ข้อมูลเชิงประจักษ์

เนื่องจากจุดประสงค์หนึ่งของงานวิจัยนี้คือการจำแนกกลุ่มความสามารถแฝงในระดับโรงเรียนโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ งานวิจัยนี้ได้พัฒนาโมเดลในการวิเคราะห์กลุ่มความสามารถแฝงบนโปรแกรม R version 3.2.3 สำหรับการประมาณค่าพารามิเตอร์ด้วยวิธีเอ็มไพริคอล เบส์ โดยประยุกต์ใช้แพคเกจ ltm version 1.0-0 และ วิเคราะห์กลุ่มความสามารถแฝงของระดับบุคคล และระดับโรงเรียน โดยประยุกต์ใช้แพคเกจ Mclust version 5.0.1 ทั้งนี้ผู้วิจัยนำรูปแบบการตอบข้อสอบที่ได้จากข้อมูลเชิงประจักษ์ ซึ่งผู้วิจัยสุ่มผลการตอบโดยใช้โปรแกรม Microsoft access มาใช้ในการวิเคราะห์ข้อมูล

สำหรับขั้นตอนการวิเคราะห์ข้อมูลเชิงประจักษ์มีขั้นตอนเช่นเดียวกับการวิเคราะห์ข้อมูลจำลอง แสดงขั้นตอนการวิเคราะห์ข้อมูลดังภาพที่ 3.2 ซึ่งมีขั้นตอนโดยสรุปดังนี้

1. ประมาณค่าพารามิเตอร์ a, b, c, θ และ ความน่าจะเป็นแบบมีเงื่อนไขที่มีรูปแบบการตอบข้อสอบเหมือนกันจะตอบถูกในแต่ละข้อ โดยใช้แพคเกจ ltm ในโปรแกรม R

2.แบ่งกลุ่มแฝงระดับบุคคล โดยใช้แพคเกจ Mclust ในโปรแกรม R ซึ่งในการแบ่งกลุ่มจะพิจารณาจากค่าความน่าจะเป็นในขั้นตอนที่ 1 จากนั้นคำนวณค่าเฉลี่ยของความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูกของแต่ละกลุ่มแฝง

3. คำนวณความสามารถของโรงเรียน โดยคำนวณจากจำนวนรูปแบบการตอบของแต่ละกลุ่มแฝงเทียบกับจำนวนรูปแบบทั้งหมดของแต่ละโรงเรียน

4. แบ่งกลุ่มแฝงระดับโรงเรียน โดยใช้แพคเกจ Mclust ในโปรแกรม R ซึ่งในการแบ่งกลุ่มจะพิจารณาจากสัดส่วนความสามารถในขั้นตอนที่ 3 จากนั้นคำนวณค่าเฉลี่ยของความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูกของแต่ละกลุ่มแฝง

5. คำนวณหาความน่าจะเป็นของการตอบถูกในแต่ละข้อของแต่ละโรงเรียน โดยคำนวณจากผลคูณของสัดส่วนของจำนวนคนในกลุ่มแฝงแต่ละกลุ่มต่อจำนวนทั้งหมด กับโอกาสในการตอบถูกของของแต่ละกลุ่มแฝงในแต่ละข้อจากขั้นตอนที่ 4

บทที่ 4
ผลการวิเคราะห์ข้อมูล

งานวิจัยนี้มีวัตถุประสงค์หลักคือ เพื่อพัฒนาโมเดลการประมาณค่ากลุ่มความสามารถของบุคคลตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ และตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคลโดยใช้ โมเดลที่พัฒนาขึ้น โดยใช้ข้อมูลจำลอง และนำโมเดลที่พัฒนาขึ้นไปใช้วิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล ซึ่งจะแสดงผลการวิเคราะห์ข้อมูลออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

ตอนที่ 2 ผลการวิเคราะห์ข้อมูลจำลอง เพื่อตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น

ตอนที่ 3 ผลการวิเคราะห์ข้อมูลเชิงประจักษ์ เพื่อจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น

ตอนที่ 1 ผลการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ในที่นี้เป็นการบูรณาการระหว่าง โมเดลการตอบสนองข้อสอบแบบผสมแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ ซึ่งผู้วิจัยได้พัฒนาจากโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับของ Tay และคณะ (2011) ซึ่งเป็นการบูรณาการระหว่าง โมเดลการตอบสนองข้อสอบแบบผสมแบบ 2 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ

สำหรับข้อตกลงเบื้องต้นทางสถิติของโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จะเป็นข้อตกลงเดียวกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับของ Tay และคณะ (2011) ซึ่งอ้างถึงในบทที่ 2

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

การบูรณาการโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้ เป็นการพัฒนาโมเดลโดยบูรณาการโมเดลพหุระดับ กับโมเดลการตอบสนองข้อสอบ (Fox & Glas, 2001) และโมเดลกลุ่มแฝง (Vermunt & Magidson, 2003) การพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้

สามารถแบ่งออกเป็น 3 ลักษณะ ได้แก่ (a) MM-IRT ถูกพัฒนาโดย การสร้างแบบจำลองแบบเป็นลำดับขั้น (hierarchical classes) (b) multilevel latent class model (MLC) ของ Vermunt (2003, 2008a) ซึ่งถูกพัฒนาโดยการเพิ่มคุณลักษณะ ที่ระดับล่าง และ (c) multilevel mixture IRT Model ของ (Vermunt, 2008b) ถูกพัฒนาโดยการเพิ่มการประมาณค่าที่ระดับบุคคล (individual classes) จากที่ 3 โมเดลข้างต้นจะมีเพียงส่วนประกอบหลัก 2 ส่วน ของโมเดล คือ individual classes individual traits และ hierarchical classes แต่ในโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับจะมีทั้ง 3 ส่วน ทั้งนี้ Cho และ Cohen (2010) และ Tay และคณะ (2011) ได้เสนอโมเดลที่คล้ายกัน แต่มีแนวคิดที่แตกต่างกันโดย Cho และ Cohen จำกัดขอบเขตที่โมเดล 1 พารามิเตอร์ (one-dimensional Rasch models) และใช้การประมาณค่าแบบเบส์เซียน (Bayesian) ส่วน Tay และคณะ จะจำกัดขอบเขตเป็นโมเดล 2 พารามิเตอร์ (multidimensional 2PL model) และใช้การประมาณค่าความน่าจะเป็นสูงสุด (maximum likelihood :ML) ซึ่งในการศึกษานี้จะใช้กับ โมเดล 3 พารามิเตอร์ (multidimensional 3PL model) และใช้การประมาณค่าแบบเอ็มไพริคอล เบส์ (Empirical Bayes) โดยพิจารณาความเหมาะสมของการประมาณค่าความน่าจะเป็นสูงสุด ซึ่งสามารถกล่าวได้ว่าโมเดลนี้เป็นโมเดลที่ใหม่ และเหมาะสมกับข้อมูลที่มีลักษณะสอดแทรกกัน (nested data) และมีการให้คะแนนแบบสองค่า และใช้ในการสำรวจกลุ่มแฝงของแต่ละบุคคล ซึ่งมีโมเดลการวัดเดียวกันข้ามกลุ่มสังกัด

สำหรับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (Multilevel Mixed-measurement IRT) แบบพหุมิติ (3PL model) ความน่าจะเป็นในการตอบสนองที่ถูก $P\left(\underset{\sim}{y_t} \right)$ ของการสังเกต เมื่อกลุ่มของการตอบสนองในระดับลดหลั่น t (ในกรณีนี้เป็นโรงเรียน) แสดงดังนี้

$$P\left(\underset{\sim}{y_t} \right) = \sum_{k=1}^{K} \pi_k P\left(\underset{\sim}{y_t} | k \right) \quad \text{.....(4.1)}$$

เมื่อ k = 1,... , K เป็นระดับกลุ่ม (เช่น โรงเรียน t) ที่เป็นกลุ่มแฝง คล้ายกับในโมเดลกลุ่มแฝง

π_k เป็นความน่าจะเป็นของหน่วยลำดับขั้น (เช่น โรงเรียน) ในกลุ่มแฝง k

ความน่าจะเป็นแบบมีเงื่อนไข $P\left(\underset{\sim}{y_t} |\ k \right)$ ของการสังเกต กลุ่มของการตอบสนอง n_t ภายในแต่ละโรงเรียน k ซึ่งสามารถแสดงดังนี้

$$P\left(\underset{\sim}{y_t} | k \right) = \prod_{j=1}^{n_t} P\left(y_{tj} | k \right) \quad \text{.....(4.2)}$$

ความน่าจะเป็นของการสังเกต เวคเตอร์การตอบสนองแต่ละบุคคลในโรงเรียน t เป็นอิสระทางสถิติของการตอบสนองของแต่ละบุคคลในกลุ่มโรงเรียน k

ความน่าจะเป็นแบบมีเงื่อนไขนี้ถูกแสดงดังนี้

$$P(y_{tj}|k)=\sum_{g=1}^{G}\pi(g|k)P\left(y_{tj}\middle|\underset{\sim}{\theta}_j,g\right) \quad \text{.....(4.3)}$$

ในที่นี้จะเห็นว่าความน่าจะเป็นของกลุ่มบุคคล $\pi(g|k)$ มีความไม่แน่นอนบนกลุ่มโรงเรียน k อย่างไรก็ตาม ความน่าจะเป็นในการตอบสนองข้อสอบของแต่ละบุคคล j แสดงดังนี้

$$P\left(y_{tj}\middle|\underset{\sim}{\theta}_j,g\right)=\prod_{i=1}^{I}P\left(y_{tji}\middle|\underset{\sim}{\theta}_j,g\right) \quad \text{.....(4.4)}$$

j = 1,...,n_t ในโรงเรียน t ในกลุ่มของข้อ i ซึ่งขึ้นอยู่กับกลุ่มบุคคล k และคุณลักษณะของแต่ละบุคคล $\underset{\sim}{\theta}_j$ เท่านั้น ซึ่งไม่เป็นฟังก์ชันของกลุ่มโรงเรียน k ซึ่งบ่งบอกว่าแต่ละกลุ่มบุคคล g เป็นหนึ่งเดียวที่ถูกนิยามโดยโมเดลการวัด $P\left(y_{ti}\middle|\underset{\sim}{\theta}_j,g\right)$ และความไม่แปรเปลี่ยนข้ามโรงเรียน และกลุ่มโรงเรียน

นอกจากนี้ มีข้อตกลงเบื้องต้นว่ามีความเป็นอิสระของการตอบข้อสอบ (local independence) ภายในแต่ละกลุ่มบุคคล และให้การตอบสนอง $\underset{\sim}{\theta}_j$ เป็นอิสระทางสถิติ

ความน่าจะเป็นแสดงดังนี้

$$\mathrm{P}\left(\mathrm{y_{tji}}\middle|\underset{\sim}{\theta}_j,\mathrm{g}\right)=\mathrm{c_i}+\frac{1-\mathrm{c_i}}{1+\exp(-\eta_{ji})} \quad \text{.....(4.5)}$$

เมื่อ η_{ji} มีค่าเท่ากับ $\beta_i+\lambda_i\theta_j$ โดยที่ β_i , λ_i , c_i แทนพารามิเตอร์ความยาก อำนาจจำแนก และการเดา ตามลำดับ

จากโมเดลที่เสนอข้างต้นนี้ผู้วิจัยสามารถสรุปเป็นขั้นตอนในการวิเคราะห์ ดังรายละเอียดขั้นตอนการวิเคราะห์ข้อมูลต่อไปนี้

การวิเคราะห์ข้อมูล

สำหรับการวิเคราะห์ข้อมูลตามโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ สามารถใช้โปรแกรมสำเร็จรูป Latent gold (Vermunt & Magidson, 2008) ซึ่งสามารถใช้ได้กับโมเดลที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ และ 2 พารามิเตอร์ แต่งานวิจัยนี้ไม่สามารถใช้โปรแกรมสำเร็จรูปในการวิเคราะห์ข้อมูลได้ เนื่องจากเป็นการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ดังนั้นผู้วิจัยจึงใช้การเขียนคำสั่งในโปรแกรม R รุ่น 3.2.3 และประยุกต์ใช้แพคเกจ ltm และ MCLUST ในการวิเคราะห์ข้อมูล

สำหรับแพคเกจ ltm นำมาประยุกต์ใช้ในการประมาณค่าพารามิเตอร์ เนื่องจากสามารถใช้กับการประมาณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบทั้งข้อสอบที่มีการให้คะแนนแบบสอง

ค่า และข้อสอบที่มีการให้คะแนนหลายค่า นอกจากนี้ในการตรวจสอบความถูกต้องของการประมาณค่า ในงานวิจัยนี้ตรวจสอบจากค่า Hessian matrix โดยประยุกต์ใช้แพ็คเกจ matrixcalc ซึ่งในการตรวจสอบพิจารณาดังเงื่อนไขต่อไปนี้ ถ้า Hessian matrix มีค่าเป็นบวกแสดงว่าการประมาณค่ามีความเสถียร ในการกำหนดค่าเริ่มต้นในการประมาณค่าผู้วิจัยใช้การสุ่มค่าเริ่มต้นจำนวน 10 ค่า ซึ่งสอดคล้องกับการกำหนดค่าเริ่มต้นของการประมาณค่าในโปรแกรม Latent gold (Vermunt & Magidson, 2008)

สำหรับแพคเกจ MCLUST นำมาประยุกต์ใช้ในการจำแนกกลุ่มความสามารถแฝง เนื่องจากสามารถใช้ได้กับข้อมูลหลายรูปแบบซึ่งเหมาะสมกับงานวิจัยนี้ ซึ่งข้อมูลที่ใช้ในการจำแนกกลุ่มมี 2 ลักษณะ คือ ความน่าจะเป็นในการตอบข้อสอบถูกรายข้อ และคะแนนความสามารถ นอกจากนี้แพคเกจ MCLUST ใช้ดัชนี BIC ในการกำหนดจำนวนกลุ่มแฝง ซึ่งโดยปกติดัชนีที่ใช้ในการกำหนดจำนวนกลุ่มแฝง สำหรับการวิเคราะห์กลุ่มแฝงพหุระดับ จะใช้ดัชนี Akaike information criteria (AIC; Akaike, 1974) ดัชนี BIC และ consistent AIC (CAIC; Bozdogan, 1987) และ ดัชนี AIC3 (Bozdogan, 1993) ซึ่งในงานวิจัยนี้เลือกใช้ ดัชนี BIC เป็นค่าดัชนีในการกำหนดจำนวนกลุ่ม เนื่องจากเป็นการวิเคราะห์กลุ่มแฝงพหุระดับ ดัชนี BIC เหมาะสมกับการวิเคราะห์ที่มีจำนวนกลุ่มตัวอย่างขนาดใหญ่ (Tay et. al. 2012) แต่สำหรับ ดัชนี AIC เหมาะสมกับการวิเคราะห์ข้อมูลขนาดเล็ก (Lukociene & Vermunt, 2010) จากการศึกษาพบว่าในการเลือกจำนวนกลุ่มแฝงของแพคเกจ MCLUST จะใช้ค่า maximum BIC (Haughton, Legrand, & Woolford, 2012) ซึ่งแตกต่างจากแพคเกจอื่น ๆ ของโปรแกรม R และโปรแกรมสำเร็จรูปอื่น ๆ แต่ให้ผลการจำแนกกลุ่มความสามารถแฝงไม่แตกต่างกัน (Haughton et al., 2012) เนื่องจากในการเขียนชุดคำสั่งสำหรับการคำนวณค่า BIC ในแพคเกจ MCLUST จะคำนวณค่า BIC เป็นจำนวนลบ จากสูตร $2\times\log lik - m\log n$ เมื่อ m แทนจำนวนพารามิเตอร์ และ n แทนจำนวนครั้งของการสังเกต ในขณะที่โดยปกติ BIC คำนวณจาก $-2\times\log lik + m\log n$

จากที่กล่าวข้างต้น สรุปเป็นขั้นตอนการวิเคราะห์ข้อมูลได้ดังขั้นตอนต่อไปนี้

(1) การประมาณค่าพารามิเตอร์ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อของข้อมูลจำลอง

ผู้วิจัยนำรูปแบบการตอบทั้งหมดที่ได้จากการจำลองมาประมาณค่าพารามิเตอร์ของ IRT ประกอบไปด้วย อำนาจจำแนก (a), ความยาก (b), การเดา(c) และ ความสามารถ (θ) ในที่นี้ใช้การประมาณค่าแบบ Empirical Bayes พิจารณาความเหมาะสมของการประมาณค่าพารามิเตอร์จากค่า Maximum likelihood และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ $P\left(y_{tji} \middle| \underset{\sim}{\theta_j}, g \right)$ โดยใช้แพคเกจ ltm

ในโปรแกรม R โดยในการประมาณค่าพารามิเตอร์จะมีการสุ่มค่าเริ่มต้นจำนวน 10 ค่า ซึ่งสอดคล้องกับการกำหนดจำนวนค่าเริ่มต้นของการประมาณค่าในโปรแกรม Latent gold (Vermunt & Magidson, 2008) และตรวจสอบว่า ค่า hessian matrix มีค่าเป็นบวก ซึ่งจะแสดงว่าค่าประมาณที่ได้เสถียร และกำหนดให้โปรแกรมทำซ้ำไม่เกิน 100,000 รอบ จากนั้นเลือกค่าพารามิเตอร์ a, b, c, และ θ ที่ทำให้เกิดค่า log likelihood สูงสุดมาใช้ในการคำนวณต่อไป

(2) การจำแนกกลุ่มความสามารถแฝงระดับบุคคล g (Individual Class)

สำหรับการจำแนกกลุ่มความสามารถแฝงระดับบุคคล จะนำความน่าจะเป็นของแต่ละรูปแบบการตอบ ในขั้นตอนที่ (1) มาวิเคราะห์กลุ่มแฝง โดยในการจำแนกกลุ่มความสามารถแฝงนี้จะใช้การจำแนกจากความน่าจะเป็นในการตอบข้อสอบถูกในแต่ละข้อ โดยจัดให้คนที่มีความน่าจะเป็นในการตอบข้อสอบถูกในแต่ละข้อใกล้เคียงกันอยู่กลุ่มแฝงเดียวกัน การวิเคราะห์ในขั้นนี้ประยุกต์ใช้แพคเกจ Mclust ในโปรแกรม R ซึ่งการพิจารณาจำนวนกลุ่มแฝงจะพิจารณาจากค่า maximum BIC (Haughton et al., 2012) สำหรับการจำแนกกลุ่มโดยใช้แพคเกจ Mclust จะใช้หลักการแบ่งกลุ่มของเกาส์ ซึ่งจะใช้วิธีการประมาณค่ามากที่สุดด้วยเทคนิควิธีทางคณิตศาสตร์ที่เรียกว่า Expectation Maximization (EM) วิธีการดังกล่าวเป็นกระบวนการทำซ้ำ (iterative) แบ่งออกเป็น 2 ขั้นตอน ขั้นตอนที่ 1 เป็นขั้นตอนการประมาณการณ์ (คาดคะเน) ซึ่งเป็นขั้นตอนการคาดคะเนค่าขั้นต้นก่อน ขั้นตอนที่ 2 เป็นขั้นตอนการปรับปรุงค่า แล้วส่งค่ากลับไปให้ขั้นตอนที่ 1 ทำงานกระบวนการทั้งหมดจะหยุดทำก็ต่อเมื่อขั้นตอนในรอบที่ผ่านมากับรอบปัจจุบันมีค่าใกล้เคียงกันมาก ๆ โดยการประมาณค่าในงานวิจัยนี้ก็คือค่าความน่าจะเป็นในการตอบข้อสอบซึ่งจะใช้ฟังก์ชันเกาส์ในการประมาณค่าความน่าจะเป็น เนื่องจากต้องการการจัดกลุ่มของแต่ละกลุ่มตามความน่าจะเป็นว่าแต่ละรูปแบบการตอบควรจัดให้อยู่กลุ่มใด

(3) การคำนวณความสามารถของโรงเรียน

สำหรับการคำนวณหาความสามารถของโรงเรียนคำนวณจากสัดส่วนของจำนวนกลุ่มแฝงระดับบุคคลในแต่ละโรงเรียน

(4) การจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน k (School Class)

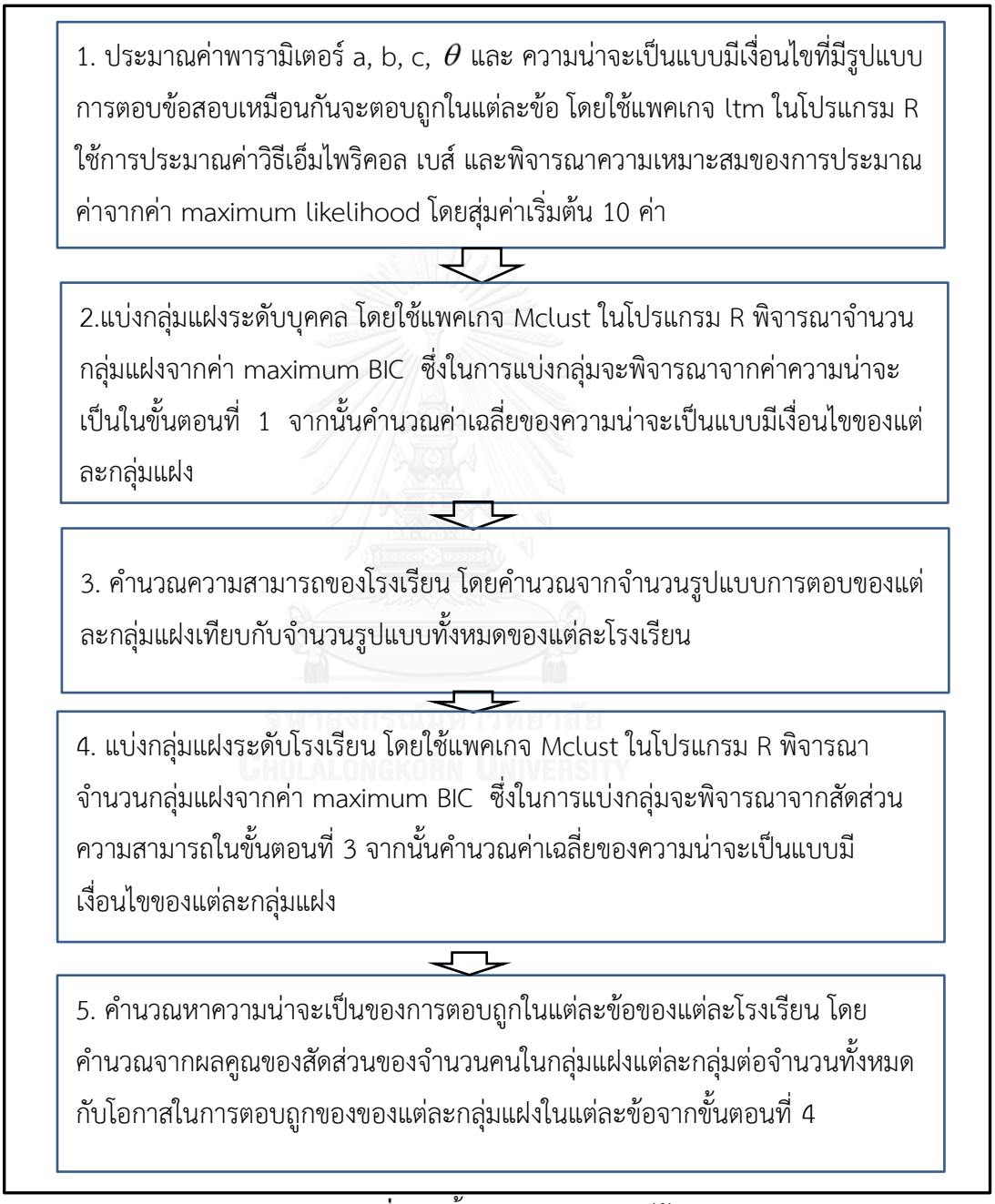
การจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนจะนำความสามารถของโรงเรียนในขั้นตอนที่ (3) มาวิเคราะห์กลุ่มแฝง โดยใช้แพคเกจ Mclust ในโปรแกรม R

(5) การคำนวณหาความน่าจะเป็นในการตอบถูกของแต่ละโรงเรียน

สำหรับการคำนวณหาความน่าจะเป็นในการตอบถูกของแต่ละโรงเรียน คำนวณจากผลคูณของสัดส่วนของจำนวนคนในกลุ่มแฝงต่อจำนวนคนทั้งหมดกับ โอกาสในการตอบถูกของแต่ละกลุ่ม

แฝงในแต่ละข้อ โดยใช้สูตร $\sum_{k=1}^{K}\pi_k P(\theta_k)$ เมื่อ k = 1,... , K เป็นระดับกลุ่ม (เช่น โรงเรียน t) ที่เป็นกลุ่มแฝง และ π_k เป็นความน่าจะเป็นของหน่วยระดับโรงเรียนในกลุ่มแฝง k

จากขั้นตอนการวิเคราะห์ข้างต้นสามารถสรุปได้ดังภาพที่ 3.2 ในบทที่ 3 แสดงอีกครั้งดังภาพต่อไปนี้



ภาพที่ 3.2 ขั้นตอนการวิเคราะห์ข้อมูล

ตอนที่ 2 ผลการตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล จากการใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับด้วยข้อมูลจำลอง

ผลการจำลองข้อมูลและการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล เพื่อตรวจสอบความถูกต้องแม่นยำ โดยใช้โมเดลที่พัฒนาขึ้น สำหรับการจำลองรูปแบบการตอบข้อสอบของแต่ละโรงเรียนตามความสามารถของนักเรียนในแต่ละโรงเรียน ที่มีความเชี่ยวชาญความรู้ในด้านที่ต่างกัน รวมทั้งขนาดของโรงเรียนที่ต่างกัน ซึ่งผู้วิจัยจำลองข้อมูล 2 สถานการณ์ คือ (1) จำลองกลุ่มแฝงของโรงเรียนออกเป็น 6 กลุ่มแฝงตามความสามารถในการตอบข้อสอบ ในที่นี้กำหนดเป็น 3 สาระ ซึ่งแต่ละด้านถูกวัดด้วยข้อสอบ 2 ข้อรวม 6 ข้อ (2) จำลองกลุ่มแฝงของโรงเรียนออกเป็น 6 กลุ่มแฝงตามความสามารถในการตอบข้อสอบ ซึ่งกำหนดเป็น 3 ด้าน โดยในสถานการณ์นี้แต่ละด้านถูกวัดด้วยข้อสอบ 5 ข้อรวม 15 ข้อ

สำหรับการวิจัยครั้งนี้ผู้วิจัยวิเคราะห์ข้อมูลแต่ละเงื่อนไขจำนวน 25 ครั้ง เพื่อตรวจสอบความถูกต้อง แม่นยำ ของการจำแนกกลุ่มความสามารถแฝง ผลการวิเคราะห์ข้อมูลจำลองที่นำเสนอต่อไปนี้เป็นตัวอย่างผลจากการวิเคราะห์ข้อมูล 1 ครั้ง จากการวิเคราะห์ทั้งหมด 25 ครั้ง ของแต่ละเงื่อนไข

ในที่นี้ผู้วิจัยได้นำเสนอตามขั้นตอนการวิเคราะห์ข้อมูล โดยแสดงตามหัวข้อ ดังนี้

2.1 ผลการวิเคราะห์ข้อมูลจำลองผลการตอบข้อสอบ แบบ 6 ข้อ

2.2 ผลการวิเคราะห์ข้อมูลจำลองผลการตอบข้อสอบ แบบ 15 ข้อ

(2.1) ผลการวิเคราะห์ข้อมูลจำลองผลการตอบข้อสอบ แบบ 6 ข้อ

2.1.1 การจำลองข้อมูล การประมาณค่าพารามิเตอร์ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ

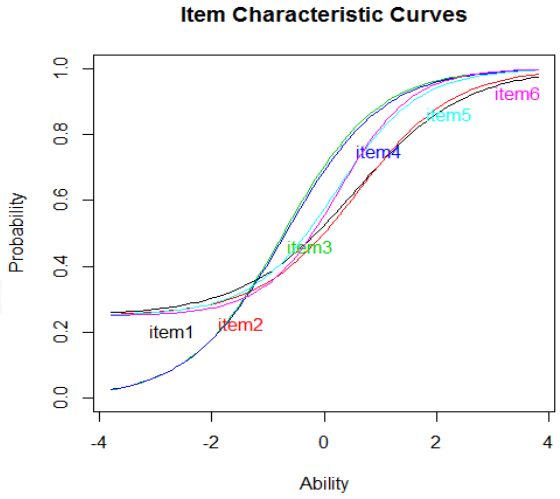
(1) รูปแบบการตอบของแต่ละโรงเรียนได้ถูกจำลองขึ้นตามความสามารถของนักเรียนในแต่ละโรงเรียนที่มีความเชี่ยวชาญความรู้ในด้านที่ต่างกัน รวมทั้งขนาดของโรงเรียนที่ต่างกัน โดยผู้วิจัยได้จำลองข้อมูลบนโปรแกรม R

(2) ประมาณค่าพารามิเตอร์ ในขั้นตอนนี้ผู้วิจัยได้ทำการประมาณค่าพารามิเตอร์ต่างๆ โดยใช้แพคเกจ ltm ในโปรแกรม R ในเริ่มต้นโปรแกรมจะทำการสุ่มค่าเริ่มต้นสำหรับการประมาณค่าพารามิเตอร์มาทั้งหมด 10 ค่า แล้วนำค่าพารามิเตอร์ที่ถูกประมาณขึ้นแล้วทำให้ log likelihood มีค่าสูงสุดไปใช้ต่อไป ค่าพารามิเตอร์ที่ประมาณแสดงดังตารางที่ 4.1 และภาพที่ 4.1

ตารางที่ 4.1 ผลการประมาณค่าพารามิเตอร์ของข้อมูลจำลองจำนวน 6 ข้อ

ข้อที่	ค่าสัมประสิทธิ์(Coefficients)		
	การเดา (Guessing)	ค่าความยาก (Difficulty)	ค่าอำนาจจำแนก (Discriminant)
1	0.250	0.536	1.020
2	0.250	0.591	1.163
3	0.000	-0.727	1.191
4	0.000	-0.686	1.164
5	0.250	0.197	1.358
6	0.250	0.244	1.537
Log.Lik: -126318			

จากตารางที่ 4.1 ค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.00 ถึง 0.25 ค่าความยาก มีค่าตั้งแต่ -0.727 ถึง 0.591 และ ค่าอำนาจจำแนก มีค่าตั้งแต่ 1.020 ถึง 1.537 และมีค่า maximum log likelihood เท่ากับ -126,318



ภาพที่ 4.1 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลจำลอง 6 ข้อ

(3) ผลการประมาณค่าความสามารถ (θ) ของแต่ละรูปแบบการตอบ แสดงดังตารางที่ 4.2

ตารางที่ 4.2 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบสำหรับข้อมูลจำลอง จำนวน 6 ข้อ

รูปแบบที่	ข้อ						จำนวน (คน)	ค่าความสามารถ(θ)	ความคลาดเคลื่อน
	1	2	3	4	5	6			
1	0	0	0	0	0	0	820	-1.310	0.671
2	0	0	0	0	0	1	386	-1.084	0.727
3	0	0	0	0	1	0	486	-1.066	0.709
4	0	0	0	0	1	1	404	-0.725	0.769
5	0	0	0	1	0	0	429	-0.832	0.614
6	0	0	0	1	0	1	272	-0.528	0.657
7	0	0	0	1	1	0	306	-0.541	0.640
8	0	0	0	1	1	1	312	-0.132	0.668
9	0	0	1	0	0	0	474	-0.822	0.613
10	0	0	1	0	0	1	286	-0.516	0.656
11	0	0	1	0	1	0	332	-0.530	0.638
12	0	0	1	0	1	1	316	-0.121	0.666
13	0	0	1	1	0	0	809	-0.410	0.579
14	0	0	1	1	0	1	912	-0.051	0.613
15	0	0	1	1	1	0	918	-0.086	0.601
16	0	0	1	1	1	1	1428	0.362	0.631
17	0	1	0	0	0	0	390	-1.133	0.695
18	0	1	0	0	0	1	202	-0.829	0.758
19	0	1	0	0	1	0	232	-0.824	0.733
20	0	1	0	0	1	1	395	-0.369	0.769
21	0	1	0	1	0	0	247	-0.622	0.634
22	0	1	0	1	0	1	184	-0.239	0.671
23	0	1	0	1	1	0	195	-0.270	0.655
24	0	1	0	1	1	1	310	0.222	0.674
25	0	1	1	0	0	0	286	-0.612	0.633
26	0	1	1	0	0	1	155	-0.227	0.670
27	0	1	1	0	1	0	185	-0.258	0.653
28	0	1	1	0	1	1	316	0.234	0.673
29	0	1	1	1	0	0	476	-0.173	0.599
30	0	1	1	1	0	1	564	0.261	0.632
31	0	1	1	1	1	0	612	0.210	0.622
32	0	1	1	1	1	1	1130	0.745	0.663
33	1	0	0	0	0	0	472	-1.122	0.685
34	1	0	0	0	0	1	240	-0.827	0.743
35	1	0	0	0	1	0	255	-0.822	0.719

ตารางที่ 4.2 (ต่อ) ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบสำหรับข้อมูลจำลอง จำนวน 6 ข้อ

รูปแบบที่	ข้อ						จำนวน (คน)	ค่าความสามารถ(θ)	ความคลาดเคลื่อน
	1	2	3	4	5	6			
36	1	0	0	0	1	1	380	-0.388	0.754
37	1	0	0	1	0	0	283	-0.627	0.625
38	1	0	0	1	0	1	173	-0.258	0.661
39	1	0	0	1	1	0	214	-0.287	0.645
40	1	0	0	1	1	1	322	0.186	0.666
41	1	0	1	0	0	0	311	-0.617	0.623
42	1	0	1	0	0	1	198	-0.246	0.660
43	1	0	1	0	1	0	242	-0.276	0.644
44	1	0	1	0	1	1	363	0.198	0.665
45	1	0	1	1	0	0	504	-0.190	0.591
46	1	0	1	1	0	1	666	0.228	0.624
47	1	0	1	1	1	0	658	0.180	0.614
48	1	0	1	1	1	1	1119	0.697	0.654
49	1	1	0	0	0	0	396	-0.907	0.708
50	1	1	0	0	0	1	263	-0.511	0.759
51	1	1	0	0	1	0	251	-0.529	0.733
52	1	1	0	0	1	1	547	0.011	0.738
53	1	1	0	1	0	0	402	-0.381	0.643
54	1	1	0	1	0	1	316	0.073	0.671
55	1	1	0	1	1	0	286	0.025	0.659
56	1	1	0	1	1	1	528	0.585	0.684
57	1	1	1	0	0	0	435	-0.370	0.642
58	1	1	1	0	0	1	279	0.085	0.670
59	1	1	1	0	1	0	298	0.037	0.658
60	1	1	1	0	1	1	652	0.598	0.684
61	1	1	1	1	0	0	683	0.085	0.613
62	1	1	1	1	0	1	784	0.588	0.653
63	1	1	1	1	1	0	804	0.524	0.644
64	1	1	1	1	1	1	4277	1.154	0.709
รวม							32,370		

จากตารางพบว่าผลการตอบข้อสอบของข้อมูลจำลอง 32,370 คน มีทั้งหมด 64 รูปแบบ ซึ่งมีค่าความสามารถ (θ) ตั้งแต่ -1.310 ถึง 1.154

(4) การคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ สำหรับผลการคำนวณความน่าจะเป็นแบบมีเงื่อนไขแสดงดังตารางที่ 4.3

ตารางที่ 4.3 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก สำหรับข้อมูลจำลองจำนวน 6 ข้อ

รูปแบบที่	ข้อ					
	1	2	3	4	5	6
1	0.349	0.324	0.333	0.326	0.336	0.313
2	0.371	0.344	0.396	0.386	0.362	0.336
3	0.373	0.345	0.401	0.391	0.364	0.338
4	0.413	0.384	0.501	0.489	0.417	0.388
5	0.399	0.370	0.469	0.458	0.399	0.371
6	0.440	0.411	0.559	0.546	0.454	0.426
7	0.438	0.409	0.555	0.542	0.451	0.423
8	0.502	0.476	0.670	0.656	0.542	0.520
9	0.400	0.372	0.472	0.461	0.400	0.372
10	0.441	0.412	0.563	0.549	0.456	0.428
11	0.439	0.410	0.559	0.546	0.454	0.425
12	0.504	0.478	0.673	0.659	0.545	0.523
13	0.457	0.428	0.593	0.580	0.479	0.451
14	0.516	0.491	0.691	0.677	0.562	0.541
15	0.510	0.485	0.682	0.668	0.554	0.532
16	0.592	0.575	0.785	0.772	0.667	0.659
17	0.366	0.339	0.382	0.373	0.356	0.331
18	0.399	0.371	0.470	0.459	0.399	0.371
19	0.400	0.371	0.471	0.460	0.400	0.372
20	0.463	0.435	0.605	0.591	0.488	0.460
21	0.426	0.397	0.531	0.519	0.435	0.407
22	0.484	0.457	0.642	0.627	0.517	0.492
23	0.479	0.452	0.633	0.619	0.510	0.484
24	0.566	0.546	0.756	0.742	0.631	0.619
25	0.428	0.399	0.534	0.522	0.437	0.409
26	0.486	0.459	0.645	0.631	0.520	0.495
27	0.481	0.454	0.636	0.622	0.513	0.487
28	0.568	0.548	0.759	0.745	0.634	0.622
29	0.495	0.469	0.659	0.645	0.533	0.509
30	0.573	0.554	0.764	0.751	0.641	0.630
31	0.563	0.543	0.753	0.740	0.628	0.615

ตารางที่ 4.3 (ต่อ) ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก สำหรับข้อมูลจำลองจำนวน 6 ข้อ

รูปแบบที่	ข้อ					
	1	2	3	4	5	6
32	0.665	0.658	0.852	0.841	0.758	0.763
33	0.367	0.340	0.385	0.376	0.357	0.332
34	0.400	0.371	0.471	0.459	0.400	0.371
35	0.400	0.371	0.472	0.461	0.400	0.372
36	0.460	0.432	0.600	0.586	0.483	0.456
37	0.426	0.396	0.530	0.517	0.435	0.406
38	0.481	0.454	0.636	0.622	0.513	0.487
39	0.476	0.449	0.628	0.614	0.506	0.480
40	0.559	0.538	0.748	0.734	0.622	0.608
41	0.427	0.398	0.533	0.520	0.437	0.408
42	0.483	0.456	0.639	0.625	0.515	0.490
43	0.478	0.451	0.631	0.617	0.509	0.483
44	0.561	0.541	0.751	0.737	0.625	0.612
45	0.492	0.465	0.655	0.641	0.529	0.504
46	0.567	0.547	0.757	0.744	0.633	0.620
47	0.558	0.537	0.747	0.733	0.621	0.607
48	0.656	0.648	0.845	0.833	0.748	0.750
49	0.390	0.362	0.447	0.436	0.387	0.359
50	0.442	0.413	0.564	0.551	0.457	0.429
51	0.439	0.410	0.559	0.546	0.454	0.425
52	0.527	0.503	0.707	0.693	0.578	0.559
53	0.462	0.433	0.602	0.588	0.485	0.458
54	0.538	0.515	0.722	0.708	0.594	0.576
55	0.530	0.506	0.710	0.696	0.581	0.563
56	0.634	0.624	0.827	0.815	0.722	0.721
57	0.463	0.435	0.605	0.591	0.487	0.460
58	0.540	0.518	0.725	0.711	0.597	0.580
59	0.532	0.508	0.713	0.699	0.584	0.566
60	0.637	0.626	0.829	0.817	0.725	0.724
61	0.540	0.518	0.725	0.711	0.596	0.579
62	0.635	0.624	0.827	0.815	0.722	0.722
63	0.623	0.610	0.816	0.804	0.707	0.704
64	0.740	0.744	0.904	0.895	0.839	0.852

จากตาราง 4.3 แสดงความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน พบว่า นักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 1 ซึ่งมีความสามารถในการตอบข้อสอบต่ำที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง 6 ถูก เท่ากับ 0.349, 0.324, 0.333, 0.326, 0.336 และ 0.313 ตามลำดับ และนักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 64 ซึ่งมีความสามารถในการตอบข้อสอบสูงที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง 6 ถูก เท่ากับ 0.740, 0.744, 0.904 0.895 0.839 และ 0.852 ตามลำดับ

2.1.2 การจำแนกกลุ่มความสามารถแฝงระดับบุคคล g (Individual Class) โดยนำความน่าจะเป็นในขั้นตอนที่ (4) มาใช้ในการจำแนกกลุ่มความสามารถแฝง โดยใช้แพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล แสดงดังตารางที่ 4.4 4.5 และ 4.6

ตารางที่ 4.4 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 6 ข้อ

VEV,8	EEV,9	EEV,10
4884.123	4879.533	4866.949

จากตารางที่ 4.4 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 8 กลุ่ม เป็นค่าที่เหมาะสมในการแบ่ง เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VEV มีค่า BIC เท่ากับ 4884.123 ดังนั้นจึงเลือกการจำแนกกลุ่มแฝงระดับบุคคลออกเป็น 8 กลุ่มแฝง

ตารางที่ 4.5 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มแฝงระดับบุคคล สำหรับข้อมูลจำลองจำนวน 6 ข้อ

กลุ่มที่	จำนวนรูปแบบ	สัดส่วน
1	5	0.078
2	8	0.125
3	10	0.156
4	13	0.203
5	5	0.078
6	7	0.109
7	9	0.141
8	7	0.109
รวม	64	1
Log. Likelihood : 2832.997 , BIC : 4884.123		

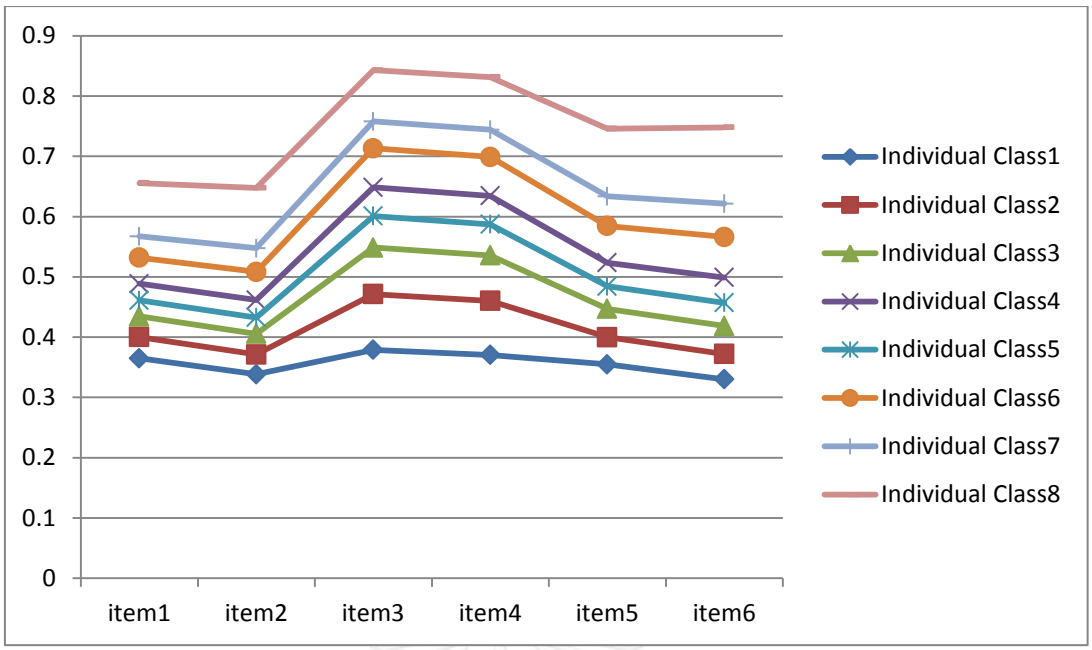
จากตารางที่ 4.5 พบว่า ค่า log likelihood เท่ากับ 2832.997 ค่า BIC เท่ากับ 2832.997 และกลุ่มแฝงที่ 1 ถึง 8 มีจำนวนรูปแบบการตอบเท่ากับ 5, 8, 10, 13, 5, 7, 9 และ 7 รูปแบบ ตามลำดับ และมีจำนวนรูปแบบในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 8 เท่ากับ 0.078, 0.125, 0.156, 0.203, 0.078, 0.109, 0.141 และ 0.109 ตามลำดับ

ตารางที่ 4.6 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 6 ข้อ

	กลุ่มที่ 1	กลุ่มที่ 2	กลุ่มที่ 3	กลุ่มที่ 4	กลุ่มที่ 5	กลุ่มที่ 6	กลุ่มที่ 7	กลุ่มที่ 8
ข้อที่ 1	0.365	0.400	0.434	0.489	0.461	0.532	0.567	0.656
ข้อที่ 2	0.338	0.371	0.405	0.462	0.433	0.508	0.548	0.648
ข้อที่ 3	0.379	0.471	0.549	0.648	0.601	0.713	0.758	0.843
ข้อที่ 4	0.370	0.460	0.536	0.634	0.587	0.699	0.744	0.831
ข้อที่ 5	0.355	0.400	0.447	0.523	0.484	0.585	0.634	0.746
ข้อที่ 6	0.330	0.372	0.418	0.499	0.457	0.566	0.621	0.748

จากตารางที่ 4.6 พบว่าค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มของกลุ่มที่ 1 มีค่าตั้งแต่ 0.330 ถึง 0.379 กลุ่มที่ 2 มีค่าตั้งแต่ 0.371 ถึง 0.471 กลุ่มที่ 3 มีค่าตั้งแต่ 0.405 ถึง 0.549 กลุ่มที่ 4 มีค่าตั้งแต่ 0.462 ถึง 0.648 กลุ่มที่ 5 มีค่าตั้งแต่ 0.433 ถึง 0.601 กลุ่มที่ 6 มีค่าตั้งแต่ 0.532 ถึง 0.713 กลุ่มที่ 7 มีค่าตั้งแต่ 0.548 ถึง 0.758 กลุ่มที่ 8 มีค่าตั้งแต่ 0.648 ถึง 0.843 กล่าวได้ว่า กลุ่มที่ 8 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกมากที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถสูงที่สุด ในทางกลับกัน กลุ่มที่ 1 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบถูกน้อยที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถต่ำที่สุด

เมื่อนำค่าความน่าจะเป็นจากตารางที่ 4.6 มาเขียนกราฟ จะได้ดังภาพที่ 4.2



ภาพที่ 4.2 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝง สำหรับข้อมูลจำลองจำนวน 6 ข้อ

2.1.3 การคำนวณหาความสามารถของโรงเรียนโดยใช้สัดส่วนของจำนวนกลุ่มแฝงระดับบุคคล สำหรับผลการคำนวณ แสดงดังตารางที่ 4.7 และ 4.8

ตารางที่ 4.7 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ

โรงเรียนที่	กลุ่มแฝง							
	1	2	3	4	5	6	7	8
1	4	9	9	5	16	39	2	36
2	4	29	43	18	67	130	7	102
3	13	45	80	27	116	262	10	247
4	23	121	185	59	338	638	37	599
5	0	4	2	11	17	14	38	34
6	4	29	10	30	68	61	80	118
7	10	65	16	70	131	82	175	251
8	37	159	55	196	275	272	430	576
9	3	4	14	48	24	46	110	111
10	1	15	43	215	67	126	377	356
11	9	32	75	407	166	237	769	705
12	14	89	189	1086	351	612	1937	1722
13	8	14	15	26	11	15	31	30
14	18	33	66	83	23	54	106	117

ตารางที่ 4.7 (ต่อ) จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ

โรงเรียนที่	กลุ่มแฝง							
	1	2	3	4	5	6	7	8
15	42	61	104	176	59	114	238	206
16	87	160	273	427	157	264	474	658
17	0	0	0	0	0	0	0	120
18	0	0	0	0	1	4	1	394
19	0	0	0	0	0	3	5	792
20	0	0	0	1	0	4	12	1983
21	84	63	59	42	21	17	8	6
22	277	206	198	151	60	37	55	16
23	585	434	404	269	118	84	83	23
24	1331	1060	997	725	335	206	254	92

จากตารางที่ 4.7 แสดงจำนวนคนในแต่ละกลุ่มแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน พบว่า โรงเรียนที่ 1 ถึง 4 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 1 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 6, 8 และ 5 ตามลำดับ และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 และ7 จำนวนน้อยที่สุด โรงเรียนที่ 5 ถึง 8 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 2 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8, 7 และ 5 และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 จำนวนน้อยที่สุด โรงเรียนที่ 9 ถึง 12 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 3 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8, 7 และ 4 และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 จำนวนน้อยที่สุด โรงเรียนที่ 13 ถึง 16 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 4 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 7, 8 และ 4 ตามลำดับ และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 จำนวนน้อยที่สุด โรงเรียนที่ 17 ถึง 20 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 5 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 อย่างชัดเจน โรงเรียนที่ 21 ถึง 24 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 6 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2 และ 3 ตามลำดับ และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 จำนวนน้อยที่สุด

ซึ่งจำนวนคนในกลุ่มแฝงระดับบุคคลของแต่ละโรงเรียนจะนำไปใช้ในการคำนวณความสามารถของโรงเรียน โดยคำนวณจากสัดส่วนจำนวนคนในแต่ละกลุ่มแฝง ดังตารางที่ 4.8

ตารางที่ 4.8 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝงระดับบุคคล สำหรับข้อมูลจำลองจำนวน 6 ข้อ

โรงเรียนที่	ความสามารถที่จำลองขึ้น	ขนาดโรงเรียน	กลุ่มแฝงที่ 1	กลุ่มแฝงที่ 2	กลุ่มแฝงที่ 3	กลุ่มแฝงที่ 4	กลุ่มแฝงที่ 5	กลุ่มแฝงที่ 6	กลุ่มแฝงที่ 7	กลุ่มแฝงที่ 8
1	กลุ่มแฝงจำลองกลุ่มที่ 1	1	0.033	0.075	0.075	0.042	0.133	0.325	0.017	0.300
2		2	0.010	0.073	0.108	0.045	0.168	0.325	0.018	0.255
3		3	0.016	0.056	0.100	0.034	0.145	0.328	0.013	0.309
4		4	0.012	0.061	0.093	0.030	0.169	0.319	0.019	0.300
5	กลุ่มแฝงจำลองกลุ่มที่ 2	1	0.000	0.033	0.017	0.092	0.142	0.117	0.317	0.283
6		2	0.010	0.073	0.025	0.075	0.170	0.153	0.200	0.295
7		3	0.013	0.081	0.020	0.088	0.164	0.103	0.219	0.314
8		4	0.019	0.080	0.028	0.098	0.138	0.136	0.215	0.288
9	กลุ่มแฝงจำลองกลุ่มที่ 3	1	0.008	0.011	0.039	0.133	0.067	0.128	0.306	0.308
10		2	0.001	0.013	0.036	0.179	0.056	0.105	0.314	0.297
11		3	0.004	0.013	0.031	0.170	0.069	0.099	0.320	0.294
12		4	0.002	0.015	0.032	0.181	0.059	0.102	0.323	0.287
13	กลุ่มแฝงจำลองกลุ่มที่ 4	1	0.053	0.093	0.100	0.173	0.073	0.100	0.207	0.200
14		2	0.036	0.066	0.132	0.166	0.046	0.108	0.212	0.234
15		3	0.042	0.061	0.104	0.176	0.059	0.114	0.238	0.206
16		4	0.035	0.064	0.109	0.171	0.063	0.106	0.190	0.263
17	กลุ่มแฝงจำลองกลุ่มที่ 5	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
18		2	0.000	0.000	0.000	0.000	0.003	0.010	0.003	0.985
19		3	0.000	0.000	0.000	0.000	0.000	0.004	0.006	0.990
20		4	0.000	0.000	0.000	0.001	0.000	0.002	0.006	0.992
21	กลุ่มแฝงจำลองกลุ่มที่ 6	1	0.280	0.210	0.197	0.140	0.070	0.057	0.027	0.020
22		2	0.277	0.206	0.198	0.151	0.060	0.037	0.055	0.016
23		3	0.293	0.217	0.202	0.135	0.059	0.042	0.042	0.012
24		4	0.266	0.212	0.199	0.145	0.067	0.041	0.051	0.018

จากตารางที่ 4.7 และ 4.8 พบว่าความสามารถของโรงเรียนซึ่งคำนวณจากสัดส่วนของจำนวนคนในกลุ่มแฝงระดับบุคคลของทุกโรงเรียน มีความสอดคล้องกับข้อมูลที่จำลองขึ้น กลุ่มแฝงจำลองกลุ่มที่ 5 ซึ่งจำลองเป็นกลุ่มเก่งทุกด้าน เมื่อวิเคราะห์จำแนกกลุ่มความสามารถแฝงแล้ว พบว่ามีนักเรียนอยู่ในกลุ่มแฝงที่ 8 จำนวนมากที่สุด ซึ่งเป็นกลุ่มที่มีความสามารถสูงที่สุด โดยโรงเรียนขนาดที่ 1 นักเรียนทุกคนถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 คิดเป็นสัดส่วนความสามารถเท่ากับ 1 โรงเรียนขนาดที่ 2 มีนักเรียนจำนวน 394 คน จากทั้งหมด 400 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.985 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 โรงเรียนขนาดที่ 3 มีนักเรียนจำนวน 792 คน จากทั้งหมด 800 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.990 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 โรงเรียนขนาดที่ 4 มี

นักเรียนจำนวน 1,983 คน จากทั้งหมด 2,000 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.992 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8

สำหรับกลุ่มแฝงจำลองกลุ่มที่ 6 ซึ่งจำลองเป็นกลุ่มอ่อนทุกด้าน เมื่อวิเคราะห์จำแนกกลุ่มความสามารถแฝงแล้ว พบว่านักเรียนอยู่ในกลุ่มที่ 8 จำนวนน้อยที่สุด และอยู่ในกลุ่มแฝงที่ 1 จำนวนมากที่สุด โดยโรงเรียนขนาดที่ 1 จากนักเรียนทั้งหมด 300 คน มีนักเรียนจำนวน 6 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.020 และจำนวน 84 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.280 โรงเรียนขนาดที่ 2 จากนักเรียนทั้งหมด 1,000 คน มีนักเรียนจำนวน 16 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.016 และจำนวน 277 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.277 โรงเรียนขนาดที่ 3 จากนักเรียนทั้งหมด 2,000 คน มีนักเรียนจำนวน 23 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.012 และจำนวน 585 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.293 โรงเรียนขนาดที่ 4 จากนักเรียนทั้งหมด 5,000 คน มีนักเรียนจำนวน 82 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 8 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.018 และจำนวน 1,331 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.266

2.1.4 การจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน k (School Class) สำหรับการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน นำความสามารถของโรงเรียนในขั้นตอนที่ 2.1.3 มาวิเคราะห์จำแนกกลุ่มความสามารถแฝงระดับโรงเรียน โดยแพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล ผลการจำแนกกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.9 4.10 และ 4.11

Chulalongkorn University

ตารางที่ 4.9 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 6 ข้อ

VEI,6	VII,6	VEI,5
1017.884	998.733	945.071

จากตารางที่ 4.9 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 6 กลุ่ม เป็นค่าที่เหมาะสมในการแบ่ง เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VEI มีค่า BIC เท่ากับ 1017.884

ตารางที่ 4.10 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับโรงเรียน สำหรับข้อมูลจำลอง จำนวน 6 ข้อ

กลุ่มที่	จำนวนรูปแบบ	สัดส่วน
1	4	0.167
2	4	0.167
3	4	0.167
4	4	0.167
5	4	0.167
6	4	0.167
รวม	24	0.167
Log. Likelihood : 613.8179 , BIC : 1017.884		

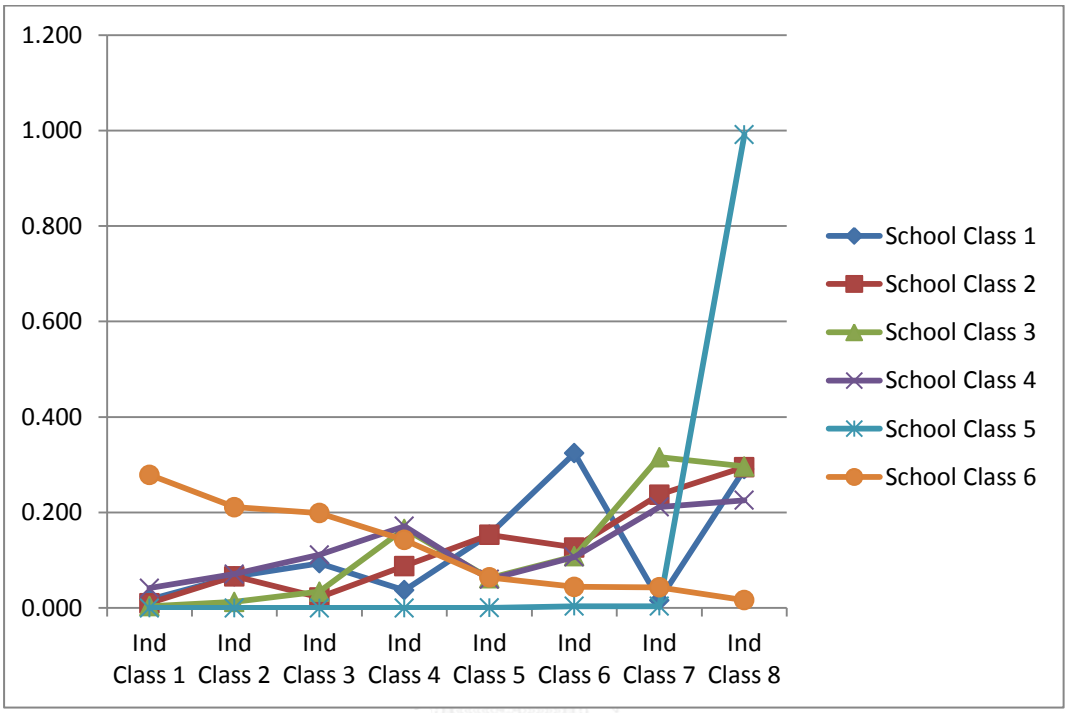
จากตารางที่ 4.10 พบว่าค่า log likelihood เท่ากับ 613.8179 ค่า BIC เท่ากับ 1017.884 และทุกกลุ่มแฝงมีรูปแบบการตอบจำนวน 4 รูปแบบ และมีจำนวนรูปแบบในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 6 เท่ากับ 0.167

ตารางที่ 4.11 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 6 ข้อ

กลุ่มแฝงระดับบุคคล	กลุ่มแฝงระดับโรงเรียน					
	1	2	3	4	5	6
1	0.018	0.010	0.004	0.042	0.000	0.279
2	0.066	0.067	0.013	0.071	0.000	0.211
3	0.094	0.022	0.034	0.111	0.000	0.199
4	0.037	0.088	0.166	0.172	0.000	0.143
5	0.154	0.153	0.063	0.060	0.001	0.064
6	0.324	0.127	0.108	0.107	0.004	0.044
7	0.016	0.238	0.316	0.212	0.004	0.043
8	0.291	0.295	0.296	0.226	0.992	0.016

จากตารางที่ 4.11 พบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 1 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อของแต่ละกลุ่มแฝงระดับบุคคล ตั้งแต่ 0.018 ถึง 0.324 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 2 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.010 ถึง 0.295 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 3 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.004 ถึง 0.316 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 4 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.042 ถึง 0.226 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 5 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.992 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 6 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.016 ถึง 0.279

จากค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มข้างต้น สามารถแสดงดังภาพที่ 4.3



ภาพที่ 4.3 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ

สำหรับผลการจำแนกกลุ่มแฝงระดับโรงเรียน แสดงดังตารางที่ 4.12

ตารางที่ 4.12 ผลการจำแนกกลุ่มแฝงระดับโรงเรียนจำแนกตามโรงเรียนสำหรับข้อมูลจำลองจำนวน 6 ข้อ

กลุ่มแฝงที่	กลุ่มแฝงระดับโรงเรียนที่ผู้วิจัยจำลองขึ้น		กลุ่มแฝงระดับโรงเรียนที่ได้จากการจำแนกโดยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ	
	จำนวน	ลำดับที่	จำนวน	ลำดับที่
1	4	1, 2, 3, 4	4	1, 2, 3, 4
2	4	5, 6, 7, 8	4	5, 6, 7, 8
3	4	9, 10, 11, 12	4	9, 10, 11, 12
4	4	13, 14, 15, 16	4	13, 14, 15, 16
5	4	17, 18, 19, 20	4	17, 18, 19, 20
6	4	21, 22, 23, 24	4	21, 22, 23, 24

จากตาราง 4.12 พบว่าผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนโรงเรียนที่ 1, 2, 3, และ 4 ถูกจำแนกอยู่กลุ่มแฝงที่ 1 โรงเรียนที่ 5, 6, 7 และ 8 ถูกจำแนกอยู่กลุ่มแฝงที่ 2 โรงเรียนที่ 9, 10 ,11 และ 12 ถูกจำแนกอยู่กลุ่มแฝงที่ 3 โรงเรียนที่ 13, 14, 15, 16 ถูกจำแนกอยู่กลุ่มแฝงที่ 4 โรงเรียนที่ 17, 18, 19, 20 ถูกจำแนกอยู่กลุ่มแฝงที่ 5 โรงเรียนที่ 21, 22, 23, 24 ถูกจำแนกอยู่กลุ่มแฝงที่ 6

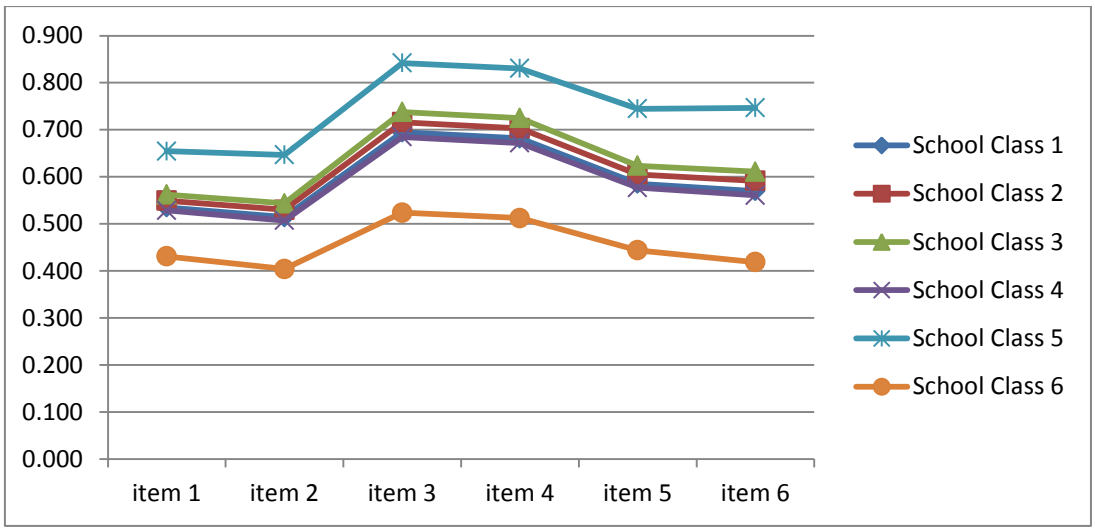
2.1.5 การคำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน ผลการจำแนกกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.13 และภาพที่ 4.4

ตารางที่ 4.13 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 6 ข้อ

กลุ่มแฝงระดับโรงเรียน	ข้อ					
	1	2	3	4	5	6
1	0.535	0.514	0.695	0.682	0.585	0.570
2	0.549	0.530	0.716	0.703	0.605	0.592
3	0.562	0.544	0.738	0.725	0.623	0.611
4	0.528	0.507	0.685	0.672	0.577	0.561
5	0.655	0.647	0.842	0.830	0.745	0.747
6	0.431	0.404	0.524	0.512	0.444	0.419

จากตารางพบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มแฝงที่ 1 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.514 ถึง 0.695 กลุ่มแฝงที่ 2 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.530 ถึง 0.716 กลุ่มแฝงที่ 3 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.544 ถึง 0.738 กลุ่มแฝงที่ 4 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.507 ถึง 0.685 กลุ่มแฝงที่ 5 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.647 ถึง 0.842 กลุ่มแฝงที่ 6 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.404 ถึง 0.524

เมื่อนำค่าความน่าจะเป็นจากตารางที่ 4.13 มาเขียนกราฟ จะได้ดังภาพที่ 4.5



ภาพที่ 4.4 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน จากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 6 ข้อ

ผลการตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝง

ผู้วิจัยตรวจสอบความถูกต้องในการจำแนกกลุ่มเทียบกับข้อมูลที่จำลองในเบื้องต้น โดยทำการวิเคราะห์ซ้ำ 25 ครั้ง ซึ่งพบว่า ผลการจำแนกกลุ่มแฝงระดับโรงเรียน โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ตรงตามข้อมูลที่ผู้วิจัยกำหนดในเงื่อนไขการจำลองข้อมูลในเริ่มต้น ทั้งสิ้น 24 ครั้ง คิดเป็นร้อยละ 96 ในที่นี้ผู้วิจัยนำเสนอรายละเอียดที่ได้จากการวิเคราะห์ซ้ำ 25 ครั้ง โดยแสดงเป็นค่าเฉลี่ย และผลการจำแนกกลุ่มความสามารถแฝงแสดงดังตารางต่อไปนี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.14 ผลการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากการวิเคราะห์ซ้ำ 25 ครั้ง สำหรับข้อมูลจำลอง 6 ข้อ

รายละเอียด	ผลการวิเคราะห์ข้อมูล 25 ครั้ง	
1.การประมาณค่าพารามิเตอร์	ข้อ 1	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.03 ค่าความยากเฉลี่ยเท่ากับ 0.55 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 2	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.11 ค่าความยากเฉลี่ยเท่ากับ 0.59 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 3	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.24 ค่าความยากเฉลี่ยเท่ากับ -0.71 ค่าการเดาเฉลี่ยเท่ากับ 0.00
	ข้อ 4	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.24 ค่าความยากเฉลี่ยเท่ากับ -0.66 ค่าการเดาเฉลี่ยเท่ากับ 0.00
	ข้อ 5	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.32 ค่าความยากเฉลี่ยเท่ากับ 0.21 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 6	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.42 ค่าความยากเฉลี่ยเท่ากับ 0.25 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ค่า log likelihood	มีค่าตั้งแต่ -126,596.10 ถึง –126,081.70 มีค่าเฉลี่ยเท่ากับ -126,268.94
2. จำนวนรูปแบบ	64 รูปแบบ	
3. ผลการจำแนกกลุ่มความสามารถแฝงระดับบุคคล	ตั้งแต่ 8 กลุ่มแฝง ถึง 10 กลุ่มแฝง ค่า Maximum BIC เฉลี่ยเท่ากับ 4932.666 ค่า log likelihood เฉลี่ยเท่ากับ 2866.999	
4. ผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน	จำแนกได้ 5 กลุ่มแฝง จำนวน 1 ครั้ง และ 6 กลุ่มแฝง จำนวน 24 ครั้ง คิดเป็นร้อยละ 96 ค่า Maximum BIC เฉลี่ยเท่ากับ 1035.063 ค่า log likelihood เฉลี่ยเท่ากับ 622.471	

(2.2) ผลการวิเคราะห์ข้อมูลจำลองผลการตอบข้อสอบ แบบ 15 ข้อ

2.2.1 การจำลองข้อมูล การประมาณค่าพารามิเตอร์ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ

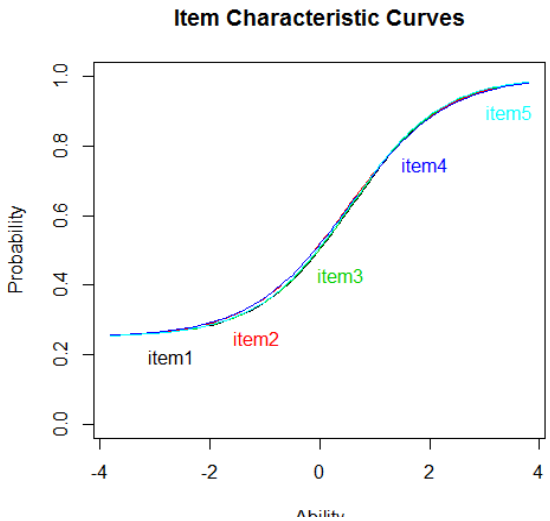
(1) รูปแบบการตอบของแต่ละโรงเรียนได้ถูกจำลองขึ้นตามความสามารถของนักเรียนในแต่ละโรงเรียนที่มีความเชี่ยวชาญความรู้ในด้านที่ต่างกัน รวมทั้งขนาดของโรงเรียนที่ต่างกัน โดยผู้วิจัยได้จำลองข้อมูลบนโปรแกรม R

(2) ประมาณค่าพารามิเตอร์ ในขั้นตอนนี้ผู้วิจัยได้ทำการประมาณค่าพารามิเตอร์ต่างๆ โดยใช้แพคเกจ ltm ในโปรแกรม R ในเริ่มต้นโปรแกรมจะทำการสุ่มค่าเริ่มต้นสำหรับการประมาณค่าพารามิเตอร์มาทั้งหมด 10 ค่า แล้วนำค่าพารามิเตอร์ที่ถูกประมาณขึ้นแล้วทำให้ log likelihood มีค่าสูงสุดไปใช้ต่อไป ค่าพารามิเตอร์ที่ประมาณแสดงดังตารางที่ 4.15 และภาพที่ 4.5 4.6 และ 4.7

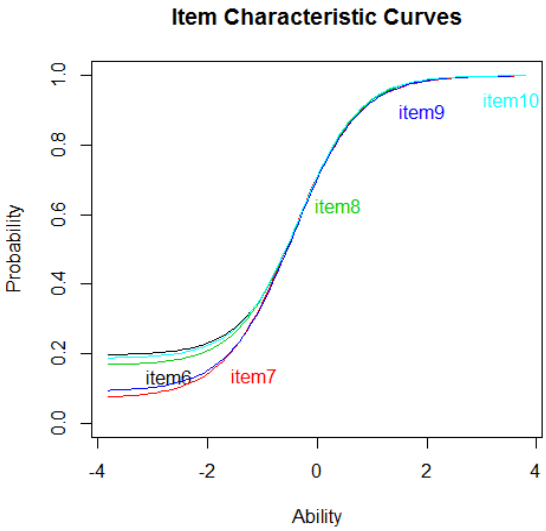
ตารางที่ 4.15 ผลการประมาณค่าพารามิเตอร์ สำหรับข้อมูลจำลองจำนวน 15 ข้อ

ข้อที่	ค่าสัมประสิทธิ์(Coefficients)		
	การเดา (Guessing)	ค่าความยาก (Difficulty)	ค่าอำนาจจำแนก (Discriminant)
1	0.250	0.575	1.189
2	0.250	0.515	1.146
3	0.250	0.563	1.193
4	0.250	0.530	1.127
5	0.250	0.544	1.194
6	0.196	-0.287	1.832
7	0.071	-0.473	1.640
8	0.166	-0.340	1.773
9	0.091	-0.421	1.683
10	0.186	-0.317	1.822
11	0.250	0.227	1.433
12	0.250	0.196	1.310
13	0.250	0.262	1.414
14	0.250	0.220	1.313
15	0.250	0.236	1.397
Log.Lik: -304231			

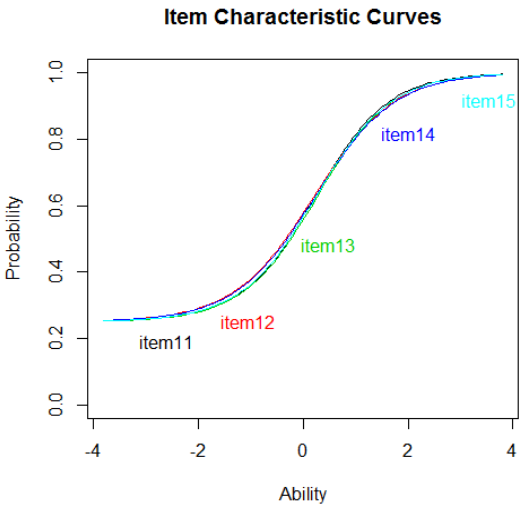
จากตารางที่ 4.15 ค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.071 ถึง 0.25 ความความยาก มีค่าตั้งแต่ -0.473 ถึง 0.575 และ ค่าอำนาจจำแนก มีค่าตั้งแต่ 1.127 ถึง 1.832 และมีค่า maximum log.likelihood เท่ากับ -304231



ภาพที่ 4.5 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง15 ข้อ (ข้อที่ 1 ถึง 5)



ภาพที่ 4.6 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง 15 ข้อ (ข้อที่ 6 ถึง 10)



ภาพที่ 4.7 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์ของข้อมูลจำลอง 15 ข้อ (ข้อที่ 11 ถึง 15)

(3) ผลการประมาณค่าความสามารถ (θ) ของแต่ละรูปแบบการตอบ แสดงดังตารางที่ 4.15

ตารางที่ 4.16 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบ สำหรับข้อมูลจำลอง จำนวน 15 ข้อ

รูปแบบ	ข้อ															จำนวน (คน)	ค่าความสามารถ(θ)	ความคลาดเคลื่อน
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	-1.678	0.550
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	10	-1.581	0.555
3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	14	-1.573	0.552
4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	4	-1.464	0.555
5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	9	-1.585	0.556
6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	4	-1.475	0.559
7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	11	-1.468	0.556
8	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	-1.344	0.556
9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	10	-1.571	0.552
10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	6	-1.461	0.554
11	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	8	-1.454	0.551
12	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	2	-1.331	0.550
13	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	5	-1.465	0.555
14	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	3	-1.341	0.555
15	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	-1.335	0.551
16	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	3	-1.197	0.547
17	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	4	-1.582	0.556
18	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	3	-1.472	0.559
19	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	10	-1.465	0.556
20	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	4	-1.340	0.556
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
13208	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	64	1.253	0.527
13209	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2630	1.674	0.599

จากตารางพบว่าผลการตอบข้อสอบของข้อมูลจำลองจำนวน 32,370 คน มีทั้งหมด 13,209 รูปแบบ ซึ่งมีค่าความสามารถ (θ) ตั้งแต่ -1.678 ถึง 1.674

(4) การคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ สำหรับผลการคำนวณความน่าจะเป็นแบบมีเงื่อนไขแสดงดังตารางที่ 4.17

ตารางที่ 4.17 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก สำหรับข้อมูลจำลองจำนวน 15 ข้อ

รูปแบบ	ข้อ														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.298	0.306	0.298	0.308	0.299	0.254	0.184	0.238	0.189	0.249	0.296	0.309	0.295	0.307	0.298
2	0.304	0.312	0.304	0.314	0.305	0.265	0.201	0.250	0.204	0.260	0.302	0.317	0.302	0.315	0.305
3	0.304	0.313	0.304	0.314	0.306	0.266	0.202	0.251	0.205	0.261	0.303	0.317	0.302	0.315	0.305
4	0.311	0.320	0.311	0.322	0.313	0.279	0.224	0.266	0.225	0.276	0.311	0.327	0.310	0.324	0.314
5	0.303	0.312	0.304	0.313	0.305	0.264	0.200	0.249	0.204	0.260	0.302	0.316	0.301	0.314	0.305
6	0.310	0.320	0.311	0.321	0.312	0.278	0.222	0.265	0.223	0.274	0.310	0.326	0.309	0.323	0.313
7	0.311	0.320	0.311	0.321	0.312	0.279	0.223	0.266	0.224	0.275	0.311	0.326	0.310	0.324	0.314
8	0.319	0.330	0.320	0.331	0.321	0.297	0.251	0.287	0.250	0.295	0.321	0.338	0.320	0.335	0.324
9	0.304	0.313	0.305	0.314	0.306	0.266	0.203	0.251	0.206	0.261	0.303	0.317	0.302	0.315	0.306
10	0.311	0.321	0.312	0.322	0.313	0.280	0.225	0.267	0.226	0.276	0.311	0.327	0.310	0.324	0.314
11	0.312	0.321	0.312	0.322	0.313	0.281	0.226	0.268	0.227	0.277	0.312	0.327	0.311	0.325	0.315
12	0.320	0.331	0.321	0.332	0.322	0.299	0.254	0.289	0.253	0.297	0.323	0.339	0.321	0.337	0.326
13	0.311	0.320	0.311	0.322	0.313	0.279	0.224	0.266	0.225	0.276	0.311	0.326	0.310	0.324	0.314
14	0.320	0.330	0.320	0.331	0.322	0.298	0.251	0.287	0.251	0.295	0.322	0.338	0.320	0.336	0.325
15	0.320	0.330	0.321	0.332	0.322	0.299	0.253	0.288	0.252	0.296	0.322	0.339	0.321	0.336	0.325
16	0.331	0.343	0.332	0.344	0.334	0.324	0.288	0.316	0.285	0.323	0.336	0.354	0.335	0.351	0.339
17	0.304	0.312	0.304	0.314	0.305	0.265	0.201	0.249	0.204	0.260	0.302	0.317	0.302	0.314	0.305
18	0.310	0.320	0.311	0.321	0.312	0.278	0.222	0.265	0.224	0.275	0.310	0.326	0.310	0.323	0.313
19	0.311	0.320	0.311	0.322	0.313	0.279	0.224	0.266	0.225	0.276	0.311	0.326	0.310	0.324	0.314
20	0.320	0.330	0.320	0.331	0.322	0.298	0.252	0.287	0.251	0.295	0.322	0.338	0.321	0.336	0.325
...	...														
13208	0.769	0.775	0.771	0.770	0.775	0.955	0.948	0.953	0.949	0.956	0.860	0.850	0.852	0.846	0.854
13209	0.840	0.843	0.843	0.838	0.846	0.978	0.973	0.977	0.974	0.979	0.916	0.906	0.910	0.903	0.911

จากตารางแสดงความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน พบว่า นักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 1 ซึ่งมีความสามารถในการตอบข้อสอบต่ำที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง 15 ถูกเท่ากับ 0.298, 0.306, 0.298, 0.308, 0.299, 0.254, 0.184, 0.238, 0.189, 0.249, 0.296, 0.309, 0.295, 0.307 และ 0.298 ตามลำดับ และนักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 13,209 ซึ่งมีความสามารถในการตอบข้อสอบสูงที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง

15 ถูก เท่ากับ 0.840, 0.843, 0.843, 0.838, 0.846, 0.978, 0.973, 0.977, 0.974, 0.979, 0.916, 0.906, 0.910, 0.903 และ 0.911 ตามลำดับ

2.2.2 การจำแนกกลุ่มความสามารถแฝงระดับบุคคล g (Individual Class) โดยนำความน่าจะเป็นในขั้นตอนที่ (4) มาใช้ในการจำแนกกลุ่มความสามารถแฝง โดยใช้แพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล แสดงดังตารางที่ 4.18 4.19 และ 4.20

ตารางที่ 4.18 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 15 ข้อ

VVV,16	VVV,15	VVV,14
3257869	3256142	3254323

จากตารางที่ 4.18 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 16 กลุ่ม เป็นค่าที่เหมาะสมในการแบ่ง เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VVV มีค่า BIC เท่ากับ 3257869 ดังนั้นจึงเลือกการจำแนกกลุ่มแฝงระดับบุคคลออกเป็น 16 กลุ่มแฝง

ตารางที่ 4.19 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับบุคคล g สำหรับข้อมูลจำลองจำนวน 15 ข้อ

กลุ่มแฝงที่	จำนวนรูปแบบ	จำนวนรูปแบบสัมพัทธ์
1	1032	0.078
2	1137	0.086
3	1228	0.093
4	566	0.043
5	828	0.063
6	589	0.045
7	1036	0.078
8	552	0.042
9	759	0.057
10	811	0.061
11	1138	0.086
12	680	0.052
13	817	0.062
14	589	0.044
15	994	0.076
16	453	0.034
Log. Likelihood : 1639253 , BIC : 3257869		

จากตารางที่ 4.19 พบว่าค่า log likelihood เท่ากับ 1639253 ค่า BIC เท่ากับ 3257869 และกลุ่มแฝงที่ 1 ถึง 16 มีรูปแบบการตอบจำนวน 1,032, 1,137, 1,228, 566 , 828, 589, 1,036, 552, 759, 811, 1138, 680, 817, 589, 994 และ 453 รูปแบบ ตามลำดับ และมีจำนวนรูปแบบในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 16 เท่ากับ 0.078, 0.086, 0.093, 0.043,0.063, 0.045, 0.078, 0.042, 0.057, 0.061, 0.086, 0.052, 0.062, 0.044, 0.076 และ 0.034 ตามลำดับ

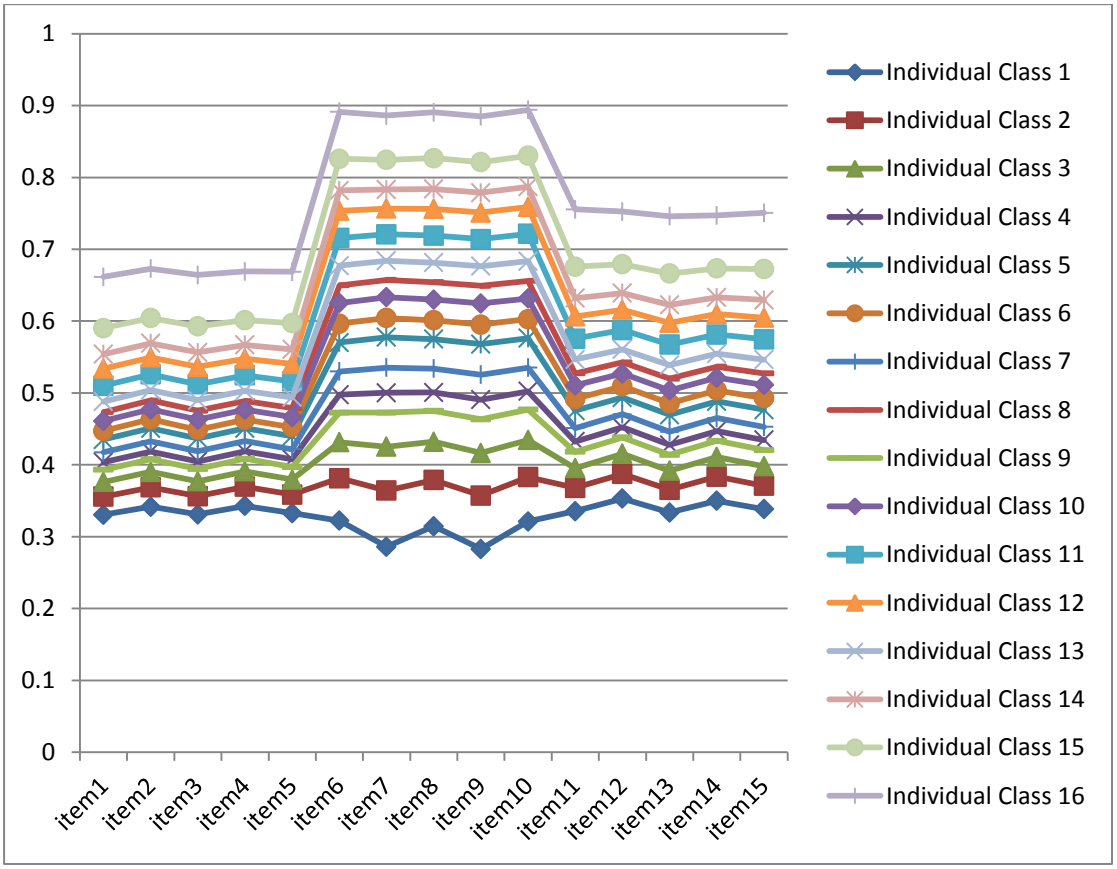
ตารางที่ 4.20 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 15 ข้อ

ข้อ	กลุ่มแฝงที่															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.331	0.356	0.376	0.404	0.435	0.447	0.417	0.473	0.393	0.461	0.510	0.534	0.488	0.554	0.590	0.662
2	0.342	0.369	0.390	0.419	0.451	0.463	0.433	0.490	0.408	0.477	0.526	0.550	0.504	0.569	0.604	0.673
3	0.331	0.356	0.377	0.405	0.437	0.449	0.419	0.475	0.394	0.463	0.512	0.536	0.490	0.556	0.593	0.664
4	0.343	0.370	0.391	0.419	0.451	0.463	0.433	0.489	0.408	0.477	0.525	0.548	0.503	0.567	0.602	0.669
5	0.333	0.359	0.380	0.408	0.440	0.452	0.422	0.479	0.397	0.466	0.516	0.540	0.494	0.561	0.597	0.669
6	0.322	0.381	0.432	0.498	0.570	0.597	0.530	0.650	0.473	0.625	0.716	0.753	0.677	0.782	0.826	0.891
7	0.286	0.364	0.425	0.500	0.578	0.604	0.535	0.657	0.473	0.633	0.721	0.757	0.684	0.783	0.824	0.886
8	0.315	0.379	0.432	0.501	0.575	0.601	0.534	0.654	0.475	0.630	0.719	0.756	0.681	0.784	0.827	0.891
9	0.283	0.357	0.417	0.491	0.568	0.595	0.526	0.649	0.463	0.625	0.714	0.751	0.676	0.779	0.821	0.885
10	0.321	0.383	0.435	0.502	0.576	0.603	0.535	0.656	0.477	0.632	0.722	0.759	0.683	0.787	0.830	0.894
11	0.335	0.368	0.395	0.432	0.475	0.492	0.451	0.527	0.418	0.511	0.576	0.607	0.547	0.632	0.676	0.756
12	0.353	0.387	0.415	0.452	0.494	0.509	0.470	0.543	0.438	0.527	0.588	0.616	0.561	0.639	0.679	0.752
13	0.334	0.365	0.392	0.428	0.470	0.486	0.446	0.520	0.414	0.504	0.567	0.598	0.539	0.623	0.666	0.746
14	0.350	0.384	0.411	0.447	0.488	0.504	0.465	0.537	0.433	0.521	0.581	0.610	0.555	0.633	0.673	0.747
15	0.338	0.371	0.398	0.435	0.477	0.493	0.453	0.527	0.421	0.511	0.575	0.605	0.546	0.629	0.672	0.751

จากตารางที่ 4.20 พบว่าค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มของกลุ่มที่ 1 มีค่าตั้งแต่ 0.283 ถึง 0.353 กลุ่มที่ 2 มีค่าตั้งแต่ 0.356 ถึง 0.387 กลุ่มที่ 3 มีค่าตั้งแต่ 0.376 ถึง 0.435 กลุ่มที่ 4 มีค่าตั้งแต่ 0.404 ถึง 0.502 กลุ่มที่ 5 มีค่าตั้งแต่ 0.435 ถึง 0.578 กลุ่มที่ 6 มีค่าตั้งแต่ 0.447 ถึง 0.604 กลุ่มที่ 7 มีค่าตั้งแต่ 0.417 ถึง 0.535 กลุ่มที่ 8 มีค่าตั้งแต่ 0.473 ถึง 0.657 กลุ่มที่ 9 มีค่าตั้งแต่ 0.393 ถึง 0.477 กลุ่มที่ 10 มีค่าตั้งแต่ 0.461 ถึง 0.633 กลุ่มที่ 11 มีค่าตั้งแต่ 0.510 ถึง 0.722 กลุ่มที่ 12 มีค่าตั้งแต่ 0.534 ถึง 0.759 กลุ่มที่ 13 มีค่าตั้งแต่ 0.488 ถึง 0.684 กลุ่มที่ 14 มีค่าตั้งแต่ 0.554 ถึง 0.787 กลุ่มที่ 15 มีค่าตั้งแต่ 0.590 ถึง 0.830 กลุ่มที่ 16 มีค่าตั้งแต่ 0.662 ถึง 0.894 กล่าวได้ว่า กลุ่มที่ 16

เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกมากที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถสูงที่สุด ในทางกลับกัน กลุ่มที่ 1 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบถูกน้อยที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถต่ำที่สุด

เมื่อนำค่าความน่าจะเป็นจากตารางที่ 4.20 มาเขียนกราฟ จะได้ดังภาพที่ 4.8



ภาพที่ 4.8 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 15 ข้อ

2.2.3 การคำนวณหาความสามารถของโรงเรียนโดยใช้สัดส่วนของจำนวนกลุ่มแฝงระดับบุคคล สำหรับผลการคำนวณ แสดงดังตารางที่ 4.21 และ 4.22

ตารางที่ 4.21 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน สำหรับข้อมูลจำลองจำนวน 15 ข้อ

โรงเรียน	กลุ่มแฝงที่															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	2	4	4	6	8	8	6	2	10	13	0	6	10	27	14
2	4	9	23	19	14	30	23	27	3	18	48	27	24	19	66	46
3	4	13	27	49	34	55	51	58	8	24	96	59	51	60	126	85
4	17	43	71	118	86	133	103	140	31	74	226	145	141	110	339	223
5	3	2	4	3	4	17	10	8	4	0	11	22	0	4	16	12
6	4	9	22	11	5	52	29	51	12	5	58	44	3	3	62	30
7	9	20	26	27	20	77	66	102	36	10	100	115	5	5	118	64
8	18	48	88	63	48	227	183	263	61	10	268	247	11	17	297	151
9	0	4	1	1	11	3	2	8	1	12	35	24	20	31	111	96
10	0	1	7	1	16	27	14	25	9	44	137	104	70	115	347	283
11	1	9	8	13	40	29	41	64	10	88	292	196	156	239	669	545
12	5	12	27	19	105	81	68	146	23	215	706	503	383	607	1565	1535
13	1	3	4	6	10	8	9	5	6	10	16	5	3	13	16	35
14	7	11	15	10	33	23	38	23	19	33	44	32	39	41	68	64
15	22	26	55	18	50	37	54	30	30	80	105	59	78	74	128	154
16	43	78	96	69	164	84	146	80	82	180	281	170	205	190	308	324
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	119
18	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	399
19	0	0	0	0	0	0	0	0	0	0	1	2	0	0	2	795
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	7	1992
21	103	52	49	11	16	5	22	2	20	5	5	3	6	1	0	0
22	323	177	149	58	49	17	73	10	76	29	11	6	18	2	2	0
23	642	362	310	98	93	39	146	26	159	60	26	6	24	3	5	1
24	1580	952	726	201	239	125	398	48	386	169	53	21	74	14	12	2

จากตารางที่ 4.21 แสดงจำนวนคนในแต่ละกลุ่มแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน พบว่า โรงเรียนที่ 1 ถึง 4 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 1 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 15, 11 และ 16 และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 และ9 จำนวนน้อยที่สุด โรงเรียนที่ 5 ถึง 8 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 2 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 15, 12, 11, 8 และ 6 และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 10, 13 และ 14 จำนวนน้อยที่สุด โรงเรียนที่ 9 ถึง 12 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 3 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 15, 16 และ 11 ตามลำดับ และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1

จำนวนน้อยที่สุด โรงเรียนที่ 13 ถึง 16 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 4 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 15, 16 และ 11 และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 จำนวนน้อยที่สุด โรงเรียนที่ 17 ถึง 20 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 5 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 อย่างชัดเจน โรงเรียนที่ 21 ถึง 24 ซึ่งเป็นโรงเรียนที่ถูกจำลองเป็นกลุ่มแฝงที่ 6 นักเรียนส่วนใหญ่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2 และ 3 ตามลำดับ และมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 จำนวนน้อยที่สุด

ซึ่งจำนวนคนในกลุ่มความสามารถแฝงระดับบุคคลของแต่ละโรงเรียนจะนำไปใช้ในการคำนวณความสามารถของโรงเรียน โดยคำนวณจากสัดส่วนจำนวนคนในแต่ละกลุ่มแฝงดังตารางที่ 4.22

ตารางที่ 4.22 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝง สำหรับข้อมูลจำลอง จำนวน 15 ข้อ

ความสามารถที่จำลองขึ้น	โรงเรียนที่	ขนาดโรงเรียน	กลุ่มแฝงระดับบุคคล															
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
กลุ่มแฝงจำลองกลุ่มที่ 1	1	1	0.000	0.017	0.033	0.033	0.050	0.067	0.067	0.050	0.017	0.083	0.108	0.000	0.050	0.083	0.225	0.117
	2	2	0.010	0.023	0.058	0.048	0.035	0.075	0.058	0.068	0.008	0.045	0.120	0.068	0.060	0.048	0.165	0.115
	3	3	0.005	0.016	0.034	0.061	0.043	0.069	0.064	0.073	0.010	0.030	0.120	0.074	0.064	0.075	0.158	0.106
	4	4	0.009	0.022	0.036	0.059	0.043	0.067	0.052	0.070	0.016	0.037	0.113	0.073	0.071	0.055	0.170	0.112
กลุ่มแฝงจำลองกลุ่มที่ 2	5	1	0.025	0.017	0.033	0.025	0.033	0.142	0.083	0.067	0.033	0.000	0.092	0.183	0.000	0.033	0.133	0.100
	6	2	0.010	0.023	0.055	0.028	0.013	0.130	0.073	0.128	0.030	0.013	0.145	0.110	0.008	0.008	0.155	0.075
	7	3	0.011	0.025	0.033	0.034	0.025	0.096	0.083	0.128	0.045	0.013	0.125	0.144	0.006	0.006	0.148	0.080
	8	4	0.009	0.024	0.044	0.032	0.024	0.114	0.092	0.132	0.031	0.005	0.134	0.124	0.006	0.009	0.149	0.076
กลุ่มแฝงจำลองกลุ่มที่ 3	9	1	0.000	0.011	0.003	0.003	0.031	0.008	0.006	0.022	0.003	0.033	0.097	0.067	0.056	0.086	0.308	0.267
	10	2	0.000	0.001	0.006	0.001	0.013	0.023	0.012	0.021	0.008	0.037	0.114	0.087	0.058	0.096	0.289	0.236
	11	3	0.000	0.004	0.003	0.005	0.017	0.012	0.017	0.027	0.004	0.037	0.122	0.082	0.065	0.100	0.279	0.227
	12	4	0.001	0.002	0.005	0.003	0.018	0.014	0.011	0.024	0.004	0.036	0.118	0.084	0.064	0.101	0.261	0.256
กลุ่มแฝงจำลองกลุ่มที่ 4	13	1	0.007	0.020	0.027	0.040	0.067	0.053	0.060	0.033	0.040	0.067	0.107	0.033	0.020	0.087	0.107	0.233
	14	2	0.014	0.022	0.030	0.020	0.066	0.046	0.076	0.046	0.038	0.066	0.088	0.064	0.078	0.082	0.136	0.128
	15	3	0.022	0.026	0.055	0.018	0.050	0.037	0.054	0.030	0.030	0.080	0.105	0.059	0.078	0.074	0.128	0.154
	16	4	0.017	0.031	0.038	0.028	0.066	0.034	0.058	0.032	0.033	0.072	0.112	0.068	0.082	0.076	0.123	0.130
กลุ่มแฝงจำลองกลุ่มที่ 5	17	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.992
	18	2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.998
	19	3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.003	0.994
	20	4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.996
กลุ่มแฝงจำลองกลุ่มที่ 6	21	1	0.343	0.173	0.163	0.037	0.053	0.017	0.073	0.007	0.067	0.017	0.017	0.010	0.020	0.003	0.000	0.000
	22	2	0.323	0.177	0.149	0.058	0.049	0.017	0.073	0.010	0.076	0.029	0.011	0.006	0.018	0.002	0.002	0.000
	23	3	0.321	0.181	0.155	0.049	0.047	0.020	0.073	0.013	0.080	0.030	0.013	0.003	0.012	0.002	0.003	0.001
	24	4	0.316	0.190	0.145	0.040	0.048	0.025	0.080	0.010	0.077	0.034	0.011	0.004	0.015	0.003	0.002	0.000

จากตารางที่ 4.22 พบว่าความสามารถของโรงเรียนซึ่งคำนวณจากสัดส่วนของจำนวนคนในกลุ่มแฝงระดับบุคคลของทุกโรงเรียน มีความสอดคล้องกับข้อมูลที่จำลองขึ้น ซึ่งผู้วิจัยจำลองให้กลุ่มที่ 1, 2, และ 3 เป็นกลุ่มที่มีความสามารถแบบคละ กลุ่มที่ 4 เป็นกลุ่มที่มีความสามารถปานกลาง กลุ่มที่ 5 เป็นกลุ่มที่มีความสามารถสูง และกลุ่มที่ 6 เป็นกลุ่มที่มีความสามารถอ่อน ซึ่งสังเกตได้ชัดจากกลุ่มแฝงที่ 5 และ 6 ดังนี้

กลุ่มแฝงจำลองกลุ่มที่ 5 ซึ่งผู้วิจัยจำลองเป็นกลุ่มเก่งทุกด้าน เมื่อวิเคราะห์จำแนกกลุ่มความสามารถแฝงแล้ว พบว่ามีนักเรียนอยู่ในกลุ่มแฝงระดับบุคคลกลุ่มที่ 16 จำนวนมากที่สุด ซึ่งเป็นกลุ่มที่มีความสามารถสูงที่สุด โดยโรงเรียนขนาดที่ 1 นักเรียนทุกคนถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 มีนักเรียนจำนวน 119 จากทั้งหมด 120 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.992 โรงเรียนขนาดที่ 2 มีนักเรียนจำนวน 399 คน จากทั้งหมด 400 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.998 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 โรงเรียนขนาดที่ 3 มีนักเรียนจำนวน 795 คน จากทั้งหมด 800 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.994 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 โรงเรียนขนาดที่ 4 มีนักเรียนจำนวน 1,992 คน จากทั้งหมด 2,000 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.996 ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16

สำหรับกลุ่มแฝงจำลองกลุ่มที่ 6 ซึ่งผู้วิจัยจำลองเป็นกลุ่มอ่อนทุกด้าน เมื่อวิเคราะห์จำแนกกลุ่มความสามารถแฝงแล้ว พบว่านักเรียนอยู่ในกลุ่มแฝงระดับบุคคลกลุ่มที่ 16 จำนวนน้อยที่สุด และอยู่ในกลุ่มแฝงที่ 1 จำนวนมากที่สุด โดยโรงเรียนขนาดที่ 1 จากนักเรียนทั้งหมด 300 คน ไม่มีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.000 และนักเรียนจำนวน 103 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.343 โรงเรียนขนาดที่ 2 จากนักเรียนทั้งหมด 1,000 คน ไม่มีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.000 และนักเรียนจำนวน 323 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.323 โรงเรียนขนาดที่ 3 จากนักเรียนทั้งหมด 2,000 คน มีนักเรียนจำนวน 1 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.001 และจำนวน 642 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.321 โรงเรียนขนาดที่ 4 จากนักเรียนทั้งหมด 5,000 คน มีนักเรียนจำนวน 2 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 16 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.000 และจำนวน 1,580 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.316

2.2.4 การจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน k (School Class) สำหรับการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน นำความสามารถของโรงเรียนในขั้นตอนที่ 2.2.3 มาวิเคราะห์จำแนกกลุ่มความสามารถแฝงระดับโรงเรียน โดยใช้แพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล ผลการแบ่งกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.23 4.24 และ 4.25

ตารางที่ 4.23 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน 3 อันดับสูงสุด สำหรับข้อมูลจำลอง 15 ข้อ

VEI,6	VEI,5	VEI,4
2473.793	2464.716	2350.752

จากตารางที่ 4.23 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 6 กลุ่ม เป็นค่าที่เหมาะสมในการแบ่ง เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VEI มีค่า BIC เท่ากับ 2473.793

ตารางที่ 4.24 จำนวนรูปแบบ และสัดส่วนในแต่ละกลุ่มความสามารถแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 15 ข้อ

กลุ่มที่	จำนวนรูปแบบ	สัดส่วน
1	4	0.167
2	4	0.167
3	4	0.167
4	4	0.167
5	4	0.167
6	4	0.167
รวม	24	1
Log. Likelihood : 1430.758, BIC : 2473.793		

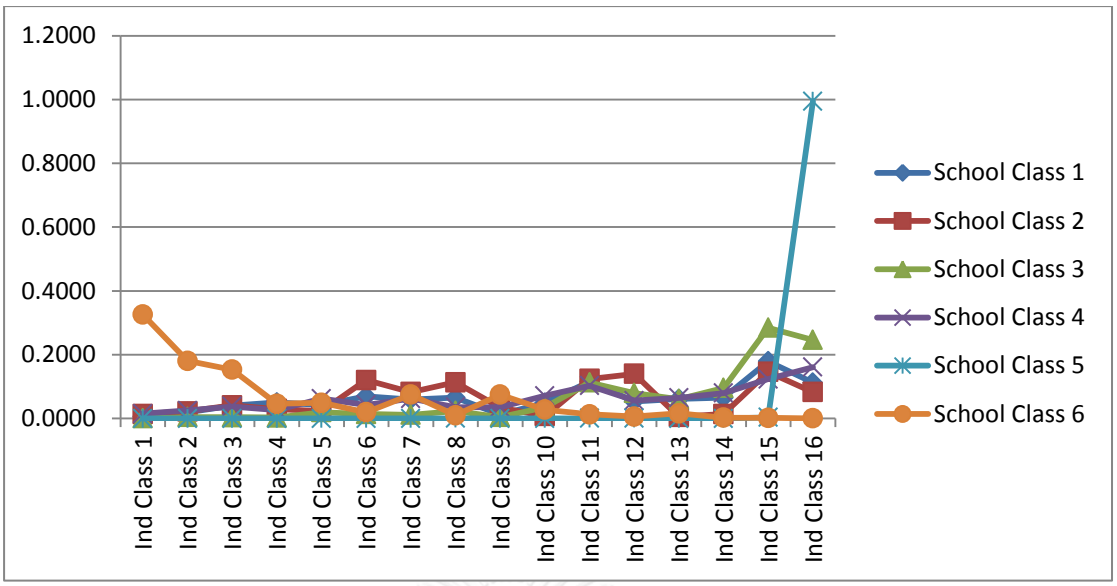
จากตารางที่ 4.24 พบว่าค่า log likelihood เท่ากับ 1430.758 ค่า BIC เท่ากับ 2473.793 และทุกกลุ่มแฝงมีรูปแบบการตอบจำนวน 4 รูปแบบ และมีจำนวนรูปแบบในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 6 เท่ากับ 0.167

ตารางที่ 4.25 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 15 ข้อ

กลุ่มแฝงระดับบุคคล	กลุ่มแฝงระดับโรงเรียน					
	1	2	3	4	5	6
1	0.006	0.014	0.000	0.015	0.000	0.326
2	0.019	0.022	0.004	0.025	0.000	0.180
3	0.040	0.041	0.004	0.038	0.000	0.153
4	0.050	0.029	0.003	0.026	0.000	0.046
5	0.043	0.024	0.020	0.062	0.000	0.049
6	0.069	0.120	0.014	0.042	0.000	0.020
7	0.060	0.082	0.011	0.062	0.000	0.075
8	0.065	0.113	0.024	0.035	0.001	0.010
9	0.012	0.035	0.005	0.035	0.000	0.075
10	0.049	0.008	0.036	0.071	0.000	0.027
11	0.115	0.124	0.113	0.103	0.000	0.013
12	0.053	0.140	0.080	0.056	0.001	0.006
13	0.061	0.005	0.061	0.064	0.000	0.016
14	0.065	0.014	0.096	0.080	0.000	0.002
15	0.179	0.146	0.284	0.123	0.004	0.002
16	0.112	0.083	0.246	0.161	0.995	0.000

จากตารางที่ 4.25 พบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 1 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อของแต่ละกลุ่มแฝงระดับบุคคล ตั้งแต่ 0.006 ถึง 0.179 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 2 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.005 ถึง 0.146 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 3 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.284 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 4 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.015 ถึง 0.161 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 5 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.995 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 6 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.326

จากค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มข้างต้น สามารถแสดงดังภาพที่ 4.9



ภาพที่ 4.9 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลจำลองจำนวน 15 ข้อ

สำหรับผลการจำแนกกลุ่มแฝงระดับโรงเรียนแสดงดังตารางที่ 4.26

ตารางที่ 4.26 ผลการจำแนกกลุ่มแฝงระดับโรงเรียนจำแนกตามโรงเรียน สำหรับข้อมูลจำลองจำนวน 15 ข้อ

กลุ่มแฝงที่	กลุ่มแฝงระดับโรงเรียนที่ผู้วิจัยจำลองขึ้น		กลุ่มแฝงระดับโรงเรียนที่ได้จากการจำแนกโดยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ	
	จำนวน	ลำดับที่	จำนวน	ลำดับที่
1	4	1, 2, 3, 4	4	1, 2, 3, 4
2	4	5, 6, 7, 8	4	5, 6, 7, 8
3	4	9, 10, 11, 12	4	9, 10, 11, 12
4	4	13, 14, 15, 16	4	13, 14, 15, 16
5	4	17, 18, 19, 20	4	17, 18, 19, 20
6	4	21, 22, 23, 24	4	21, 22, 23, 24

จากตาราง 4.26 พบว่าผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนโรงเรียนที่ 1, 2, 3, และ 4 ถูกจำแนกอยู่กลุ่มแฝงที่ 1 โรงเรียนที่ 5, 6, 7 และ 8 ถูกจำแนกอยู่กลุ่มแฝงที่ 2 โรงเรียนที่ 9, 10 ,11 และ 12 ถูกจำแนกอยู่กลุ่มแฝงที่ 3 โรงเรียนที่ 13, 14, 15, 16 ถูกจำแนกอยู่กลุ่มแฝงที่ 4 โรงเรียนที่ 17, 18, 19, 20 ถูกจำแนกอยู่กลุ่มแฝงที่ 5 โรงเรียนที่ 21, 22, 23, 24 ถูกจำแนกอยู่กลุ่มแฝงที่ 6

2.2.5 การคำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน ผลการจำแนกกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.27 และภาพที่ 4.10

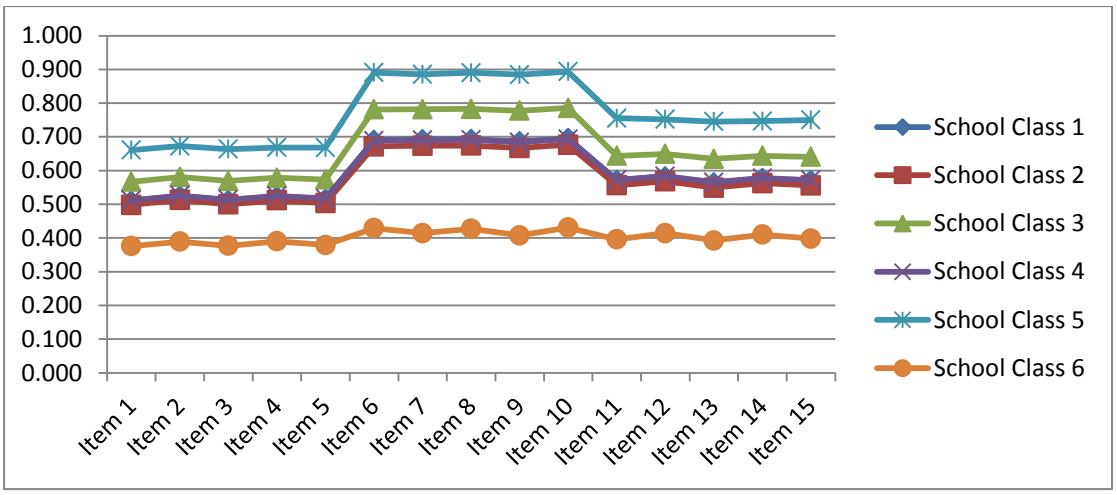
ตารางที่ 4.27 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลจำลองจำนวน 15 ข้อ

กลุ่มแฝงระดับโรงเรียน	ข้อ														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.510	0.525	0.512	0.524	0.516	0.691	0.693	0.693	0.686	0.696	0.572	0.583	0.564	0.577	0.571
2	0.498	0.513	0.500	0.512	0.504	0.672	0.674	0.674	0.667	0.677	0.556	0.568	0.549	0.563	0.556
3	0.567	0.581	0.569	0.579	0.573	0.781	0.782	0.783	0.777	0.786	0.644	0.650	0.635	0.644	0.641
4	0.512	0.526	0.514	0.525	0.517	0.689	0.690	0.691	0.684	0.694	0.573	0.583	0.565	0.578	0.572
5	0.661	0.672	0.664	0.669	0.668	0.891	0.886	0.890	0.884	0.893	0.755	0.752	0.745	0.747	0.750
6	0.376	0.389	0.377	0.390	0.379	0.429	0.415	0.427	0.408	0.431	0.396	0.414	0.392	0.410	0.398

จากตารางพบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มแฝงที่ 1 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.510 ถึง 0.696 กลุ่มแฝงที่ 2 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.498 ถึง 0.677 กลุ่มแฝงที่ 3 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.567 ถึง 0.786 กลุ่มแฝงที่ 4 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.512 ถึง 0.694 กลุ่มแฝงที่ 5 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.661 ถึง 0.893 กลุ่มแฝงที่ 6 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.376 ถึง 0.431

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

เมื่อนำค่าความน่าจะเป็นจากตารางที่ 4.27 มาเขียนกราฟ จะได้ดังภาพที่ 4.10



ภาพที่ 4.10 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียนจากการจำแนกกลุ่มแฝงของข้อมูลจำลองจำนวน 15 ข้อ

ผลการตรวจสอบความถูกต้องในการจำแนกกลุ่มความสามารถแฝง

ผู้วิจัยตรวจสอบความถูกต้องในการจำแนกกลุ่มเทียบกับข้อมูลที่จำลองในเบื้องต้น โดยทำการวิเคราะห์ซ้ำ 25 ครั้ง ซึ่งพบว่า ผลการจำแนกกลุ่มแฝงระดับโรงเรียน โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ตรงตามข้อมูลที่ผู้วิจัยกำหนดในเงื่อนไขการจำลองข้อมูลในเริ่มต้นทุกครั้งคิดเป็นร้อยละ 100 ในที่นี้ผู้วิจัยนำเสนอรายละเอียด ที่ได้จากการวิเคราะห์ซ้ำ 25 ครั้ง โดยแสดงเป็นค่าเฉลี่ย และผลการจำแนกกลุ่มความสามารถแฝงแสดงดังตารางต่อไปนี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.28 ผลการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากการวิเคราะห์ซ้ำ 25 ครั้ง สำหรับข้อมูลจำลอง 15 ข้อ

รายละเอียด	ผลการวิเคราะห์ข้อมูล 25 ครั้ง	
1.การประมาณค่าพารามิเตอร์	ข้อ 1	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.22 ค่าความยากเฉลี่ยเท่ากับ 0.55 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 2	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.16 ค่าความยากเฉลี่ยเท่ากับ 0.51 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 3	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.20 ค่าความยากเฉลี่ยเท่ากับ 0.56 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 4	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.15 ค่าความยากเฉลี่ยเท่ากับ 0.52 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 5	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.22 ค่าความยากเฉลี่ยเท่ากับ 0.55 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 6	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.74 ค่าความยากเฉลี่ยเท่ากับ -0.34 ค่าการเดาเฉลี่ยเท่ากับ 0.17
	ข้อ 7	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.59 ค่าความยากเฉลี่ยเท่ากับ -0.48 ค่าการเดาเฉลี่ยเท่ากับ 0.07
	ข้อ 8	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.67 ค่าความยากเฉลี่ยเท่ากับ -0.39 ค่าการเดาเฉลี่ยเท่ากับ 0.15
	ข้อ 9	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.57 ค่าความยากเฉลี่ยเท่ากับ -0.51 ค่าการเดาเฉลี่ยเท่ากับ 0.06
	ข้อ 10	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.69 ค่าความยากเฉลี่ยเท่ากับ -0.38 ค่าการเดาเฉลี่ยเท่ากับ 0.16

ตารางที่ 4.28 (ต่อ) ผลการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากการวิเคราะห์ซ้ำ 25 ครั้ง สำหรับข้อมูลจำลอง 15 ข้อ

รายละเอียด	ผลการวิเคราะห์ข้อมูล 25 ครั้ง	
	ข้อ 11	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.44 ค่าความยากเฉลี่ยเท่ากับ 0.23 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 12	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.36 ค่าความยากเฉลี่ยเท่ากับ 0.20 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 13	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.46 ค่าความยากเฉลี่ยเท่ากับ 0.24 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 14	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.37 ค่าความยากเฉลี่ยเท่ากับ 0.19 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ข้อ 15	ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 1.43 ค่าความยากเฉลี่ยเท่ากับ 0.24 ค่าการเดาเฉลี่ยเท่ากับ 0.25
	ค่า log likelihood	มีค่าตั้งแต่ -304,415.7 ถึง -303,881.3 มีค่าเฉลี่ยเท่ากับ –304,142.24
2. จำนวนรูปแบบ	เฉลี่ย 13,153 รูปแบบ	
3. ผลการจำแนกกลุ่มความสามารถแฝงระดับบุคคล	ตั้งแต่ 14 กลุ่มแฝง ถึง 20 กลุ่มแฝง ค่า Maximum BIC เฉลี่ย 3,242,889.29 ค่า log likelihood เฉลี่ย 1,632,509.43	
4. ผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน	จำแนกได้ 6 กลุ่มแฝง จำนวน 25 ครั้ง จากการวิเคราะห์ 25 ครั้ง ซึ่งคิดเป็นร้อยละ 100 ค่า Maximum BIC เฉลี่ย 2,829.09 ค่า log likelihood เฉลี่ย 1,624.30	

ตอนที่ 3 ผลการวิเคราะห์ข้อมูลเชิงประจักษ์ เพื่อจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น

สำหรับผลการวิเคราะห์ข้อมูลเชิงประจักษ์ เพื่อจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคล ด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่ผู้วิจัยพัฒนาขึ้น ผู้วิจัยนำเสนอตามขั้นตอนการวิเคราะห์ข้อมูล ดังต่อไปนี้

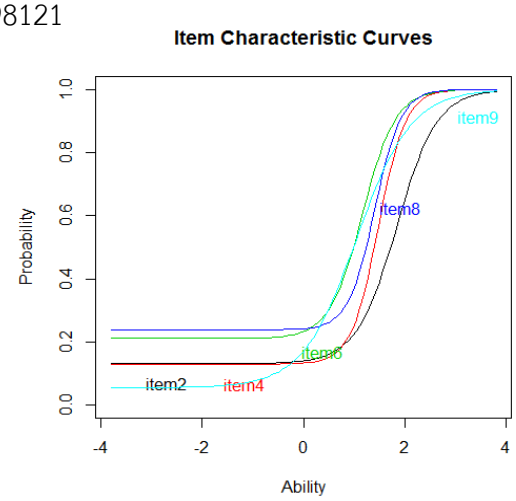
3.1 การประมาณค่าพารามิเตอร์ และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ

(1) การประมาณค่าพารามิเตอร์ ในขั้นตอนนี้ผู้วิจัยนำรูปแบบการตอบข้อสอบจำนวน 15 ข้อที่ผู้วิจัยเลือกให้ครอบคลุมสาระการเรียนรู้ทั้งหมด ซึ่งเป็นข้อมูลที่ได้จากข้อมูลเชิงประจักษ์ และผู้วิจัยทำการสุ่มผลการตอบโดยใช้โปรแกรม Microsoft access มาประมาณค่าพารามิเตอร์ของ IRT โดยใช้ แพคเกจ ltm ในโปแกรม R ในเริ่มต้นโปรแกรมจะทำการสุ่มค่าเริ่มต้นสำหรับการประมาณค่าพารามิเตอร์มาทั้งหมด 10 ค่า แล้วนำค่าพารามิเตอร์ที่ถูกประมาณขึ้นแล้วทำให้ log likelihood มีค่าสูงสุดไปใช้ต่อไป ค่าพารามิเตอร์ที่ประมาณแสดงดังตารางที่ 4.29 และภาพที่ 4.11 4.12 4.13

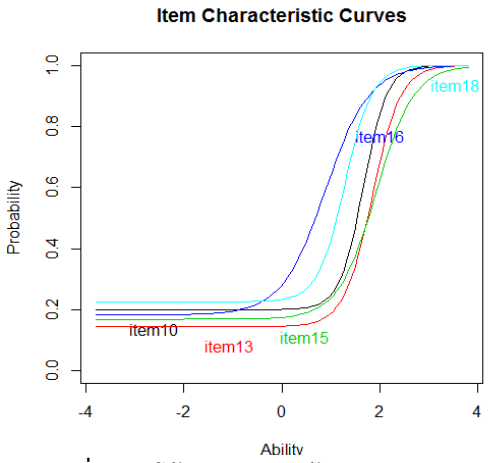
ตารางที่ 4.29 ผลการประมาณค่าพารามิเตอร์ สำหรับข้อมูลเชิงประจักษ์จำนวน 15 ข้อ

ข้อที่	ค่าสัมประสิทธิ์(Coefficients)		
	การเดา (Guessing)	ค่าความยาก (Difficulty)	ค่าอำนาจจำแนก (Discriminant)
2	0.132	1.840	2.513
4	0.130	1.476	3.752
6	0.211	1.163	3.062
8	0.239	1.403	3.850
9	0.055	1.047	1.873
10	0.200	1.660	4.162
13	0.145	1.845	3.462
15	0.169	1.922	2.598
16	0.183	0.889	2.232
18	0.226	1.288	3.599
20	0.250	0.953	3.316
22	0.229	1.232	3.289
27	0.203	1.563	2.135
29	0.000	1.766	0.353
31	0.169	1.720	2.715
Log. Lik: -598121			

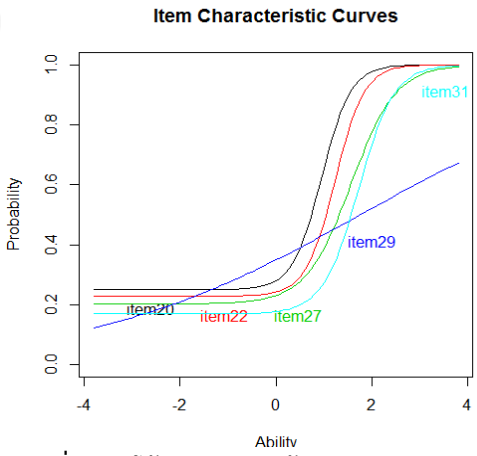
จากตารางที่ 4.29 พบว่าค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.00 ถึง 0.25 ซึ่งตามทฤษฎีค่าการเดามีค่าตั้งแต่ 0 ถึง 1 แต่ในทางปฏิบัติค่าการเดามีค่าตั้งแต่ 0 ถึง 0.3 และถ้ามีค่ามากกว่า 0.3 แสดงถึงข้อสอบที่ไม่ดี เพราะผู้สอบที่มีความสามารถอ่อนมีโอกาสมากที่จะตอบข้อสอบข้อนั้นถูก แต่ถ้าการเดาเป็น 0 แสดงถึงข้อสอบที่ดีมาก เพราะผู้สอบต้องทำข้อสอบด้วยความสามารถเท่านั้นจึงจะมีโอกาสตอบข้อสอบได้ถูกต้อง ความความยาก มีค่าตั้งแต่ 0.889 ถึง 1.922 ซึ่งค่าความยากที่ใกล้ 2 หมายถึงข้อสอบที่มีความยากมาก และค่าความยากที่ใกล้ -2 หมายถึงข้อสอบที่มีความง่ายมาก และค่าอำนาจจำแนก มีค่าตั้งแต่ 0.353 ถึง 4.162 ซึ่งค่าอำนาจจำแนกที่ใกล้ 0 หมายถึงมีอำนาจจำแนกที่ต่ำ และค่าอำนาจจำแนกที่มีค่ามาก หมายถึงข้อสอบสามารถจำแนกได้ดี และมีค่า maximum log likelihood เท่ากับ -598121



ภาพที่ 4.11 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 2, 4, 6, 8 และ 9)



ภาพที่ 4.12 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 10, 13, 15, 16 และ 18)



ภาพที่ 4.13 โค้งคุณลักษณะข้อสอบจากการประมาณค่าพารามิเตอร์สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ (ข้อ 20, 22, 27, 29 และ 31)

(2) ผลการประมาณค่าความสามารถ (θ) ของแต่ละรูปแบบการตอบ แสดงดังตารางที่ 4.29

ตารางที่ 4.30 ผลการประมาณค่าความสามารถของแต่ละรูปแบบการตอบ สำหรับข้อมูลเชิงประจักษ์

รูปแบบ	ข้อที่															จำนวน (คน)	ค่าความสามารถ(θ)	ความคลาดเคลื่อน
	2	4	6	8	9	10	13	15	16	18	20	22	27	29	31			
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1416	-0.453	0.742
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	311	-0.428	0.757
3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	722	-0.277	0.669
4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	139	-0.244	0.684
5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	408	-0.370	0.771
6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	123	-0.335	0.794
7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	257	-0.182	0.687
8	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	52	-0.136	0.707
9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	485	-0.419	0.770
10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	120	-0.389	0.792
11	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	180	-0.228	0.702
12	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	43	-0.186	0.727
13	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	155	-0.321	0.816
14	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	16	-0.276	0.852
15	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	94	-0.108	0.736
16	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	18	-0.041	0.773
17	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	635	-0.370	0.818
18	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	165	-0.331	0.854
19	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	303	-0.151	0.758
20	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	51	-0.087	0.802
...																		
12552	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	45	2.136	0.296
12553	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	305	2.431	0.390

จากตารางพบว่าผลการตอบข้อสอบของข้อมูล 72,295 คน มีทั้งหมด 12,553 รูปแบบ ซึ่งมีค่าความสามารถ (θ) ตั้งแต่ -0.453 ถึง 2.431

(3) การคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ที่แต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ สำหรับผลการคำนวณความน่าจะเป็นแบบมีเงื่อนไขแสดงดังตารางที่ 4.30

ตารางที่ 4.31 ผลการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

รูปแบบ	ข้อ														
	2	4	6	8	9	10	13	15	16	18	20	22	27	29	31
1	0.134	0.130	0.216	0.239	0.109	0.200	0.145	0.171	0.222	0.227	0.257	0.232	0.213	0.314	0.171
2	0.135	0.131	0.217	0.240	0.111	0.200	0.145	0.171	0.224	0.228	0.258	0.232	0.214	0.316	0.172
3	0.136	0.131	0.220	0.240	0.128	0.201	0.145	0.172	0.239	0.229	0.263	0.234	0.218	0.327	0.173
4	0.136	0.131	0.221	0.240	0.132	0.201	0.145	0.172	0.243	0.229	0.264	0.235	0.219	0.330	0.173
5	0.135	0.131	0.218	0.240	0.117	0.201	0.145	0.171	0.230	0.228	0.259	0.233	0.215	0.320	0.172
6	0.135	0.131	0.219	0.240	0.121	0.201	0.145	0.171	0.233	0.228	0.260	0.233	0.216	0.323	0.172
7	0.137	0.132	0.224	0.241	0.141	0.201	0.146	0.172	0.252	0.230	0.267	0.236	0.222	0.335	0.174
8	0.138	0.132	0.225	0.241	0.148	0.201	0.146	0.173	0.258	0.230	0.270	0.237	0.223	0.338	0.175
9	0.135	0.131	0.217	0.240	0.112	0.200	0.145	0.171	0.225	0.228	0.258	0.232	0.214	0.316	0.172
10	0.135	0.131	0.218	0.240	0.115	0.201	0.145	0.171	0.228	0.228	0.259	0.232	0.215	0.319	0.172
11	0.136	0.131	0.222	0.240	0.135	0.201	0.146	0.172	0.246	0.229	0.265	0.235	0.220	0.331	0.173
12	0.137	0.132	0.223	0.241	0.141	0.201	0.146	0.172	0.251	0.230	0.267	0.236	0.221	0.334	0.174
13	0.135	0.131	0.219	0.240	0.123	0.201	0.145	0.171	0.235	0.228	0.261	0.233	0.217	0.324	0.172
14	0.136	0.131	0.220	0.240	0.128	0.201	0.145	0.172	0.240	0.229	0.263	0.234	0.218	0.327	0.173
15	0.138	0.132	0.227	0.241	0.153	0.201	0.146	0.173	0.263	0.231	0.272	0.238	0.225	0.340	0.175
16	0.139	0.133	0.230	0.242	0.164	0.201	0.146	0.174	0.274	0.232	0.277	0.240	0.228	0.346	0.176
17	0.135	0.131	0.218	0.240	0.117	0.201	0.145	0.171	0.230	0.228	0.259	0.233	0.215	0.320	0.172
18	0.135	0.131	0.219	0.240	0.122	0.201	0.145	0.171	0.234	0.228	0.261	0.233	0.217	0.323	0.172
19	0.137	0.132	0.225	0.241	0.146	0.201	0.146	0.173	0.256	0.230	0.269	0.237	0.223	0.337	0.174
20	0.138	0.132	0.228	0.241	0.156	0.201	0.146	0.173	0.266	0.231	0.273	0.239	0.226	0.342	0.175
...															
12552	0.720	0.933	0.962	0.957	0.891	0.903	0.771	0.697	0.952	0.965	0.985	0.962	0.819	0.533	0.797
12553	0.840	0.977	0.984	0.986	0.934	0.969	0.901	0.825	0.975	0.988	0.995	0.985	0.892	0.559	0.895

จากตารางที่ 4.31 แสดงความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน พบว่า นักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 1 ซึ่งมีความสามารถในการตอบข้อสอบต่ำที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง 15 ถูกเท่ากับ 0.134, 0.130, 0.216, 0.239, 0.109, 0.200, 0.145, 0.171, 0.222, 0.227, 0.257, 0.232,

0.213, 0.314 และ 0.171 ตามลำดับ และนักเรียนที่มีความสามารถตรงกับรูปแบบการตอบที่ 12,553 ซึ่งมีความสามารถในการตอบข้อสอบสูงที่สุด มีความน่าจะเป็นในการตอบข้อสอบข้อที่ 1 ถึง 15 ถูก เท่ากับ 0.840, 0.977, 0.984, 0.986, 0.934, 0.969, 0.901, 0.825, 0.975, 0.988, 0.995, 0.985, 0.892, 0.559 และ 0.895 ตามลำดับ

2.1.2 การจำแนกกลุ่มความสามารถแฝงระดับบุคคล g (Individual Class) โดยนำความน่าจะเป็นในขั้นตอนที่ (3) มาใช้ในการจำแนกกลุ่มความสามารถแฝง โดยใช้แพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล แสดงดังตารางที่ 4.32 4.33 และ 4.34

ตารางที่ 4.32 ค่า BIC ของการจำแนกกลุ่มระดับบุคคล 3อันดับสูงสุด สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

VVV,20	VVV,19	VVV,18
3059008	3057408	3054077

จากตารางที่ 4.32 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 20 กลุ่ม เป็นค่าที่เหมาะสมในการแบ่ง เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VVV มีค่า BIC เท่ากับ 3059008 ดังนั้นจึงเลือกการจำแนกกลุ่มแฝงระดับบุคคลออกเป็น 20 กลุ่มแฝง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.33 จำนวนรูปแบบ สัดส่วนในแต่ละกลุ่มแฝงระดับบุคคล สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มที่	จำนวนรูปแบบ	สัดส่วน	ร้อยละ
1	1377	0.109	10.9
2	796	0.064	6.4
3	549	0.044	4.4
4	592	0.047	4.7
5	591	0.047	4.7
6	369	0.030	3.0
7	563	0.045	4.5
8	598	0.047	4.7
9	663	0.053	5.3
10	657	0.052	5.2
11	649	0.052	5.2
12	490	0.039	3.9
13	568	0.045	4.5
14	594	0.047	4.7
15	719	0.057	5.7
16	653	0.052	5.2
17	551	0.044	4.4
18	589	0.047	4.7
19	586	0.047	4.7
20	399	0.032	3.2
รวม	12,553	1	100
Log. Likelihood : 1542335, BIC : 3059008			

จากตารางพบว่า ค่า log likelihood เท่ากับ 1542335 ค่า BIC เท่ากับ 3059008 และกลุ่มแฝงที่ 1 ถึง 20 มีจำนวนรูปแบบการตอบเท่ากับ 1377, 796, 549, 592, 591, 369, 563, 598, 663, 657, 649, 490, 568, 594, 719, 653, 551, 589, 586 และ 399 รูปแบบ ตามลำดับ และมีจำนวนรูปแบบในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 20 คิดเป็นร้อยละ 10.9, 6.4, 4.4, 4.7, 4.7, 3.0, 4.5, 4.7, 5.3, 5.2, 5.2, 3.9, 4.5, 4.7, 5.7, 5.2, 4.4, 4.7, 4.7 และ 3.2 ตามลำดับ

ตารางที่ 4.34 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม

ข้อที่	กลุ่มแฝงที่																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	0.136	0.139	0.144	0.151	0.160	0.168	0.197	0.176	0.186	0.208	0.222	0.235	0.249	0.265	0.288	0.317	0.349	0.389	0.450	0.562
4	0.131	0.132	0.136	0.142	0.151	0.160	0.203	0.171	0.186	0.223	0.248	0.273	0.300	0.335	0.381	0.442	0.506	0.581	0.680	0.813
8	0.240	0.241	0.245	0.252	0.261	0.271	0.317	0.283	0.299	0.338	0.365	0.390	0.419	0.454	0.501	0.559	0.619	0.686	0.769	0.872
9	0.127	0.159	0.206	0.255	0.302	0.338	0.437	0.370	0.405	0.467	0.499	0.525	0.550	0.578	0.610	0.646	0.679	0.715	0.759	0.822
เฉลี่ย	**0.159**	**0.168**	**0.183**	**0.200**	**0.219**	**0.234**	**0.289**	**0.250**	**0.269**	**0.309**	**0.334**	**0.356**	**0.380**	**0.408**	**0.445**	**0.491**	**0.538**	**0.593**	**0.665**	**0.767**
16	0.239	0.269	0.317	0.370	0.422	0.461	0.569	0.496	0.535	0.601	0.634	0.660	0.685	0.712	0.742	0.774	0.803	0.832	0.866	0.910
18	0.229	0.232	0.239	0.251	0.268	0.284	0.355	0.303	0.328	0.384	0.419	0.451	0.486	0.526	0.576	0.635	0.691	0.751	0.819	0.900
เฉลี่ย	**0.234**	**0.251**	**0.278**	**0.311**	**0.345**	**0.373**	**0.462**	**0.400**	**0.432**	**0.493**	**0.527**	**0.556**	**0.586**	**0.619**	**0.659**	**0.705**	**0.747**	**0.792**	**0.843**	**0.905**
6	0.220	0.229	0.246	0.271	0.301	0.328	0.425	0.356	0.391	0.460	0.499	0.533	0.567	0.605	0.649	0.699	0.744	0.791	0.844	0.907
10	0.201	0.201	0.202	0.204	0.207	0.210	0.226	0.213	0.219	0.234	0.245	0.257	0.271	0.290	0.317	0.357	0.405	0.469	0.568	0.732
13	0.145	0.146	0.147	0.150	0.153	0.156	0.171	0.160	0.165	0.177	0.186	0.195	0.205	0.218	0.237	0.265	0.297	0.342	0.415	0.563
15	0.172	0.174	0.177	0.183	0.189	0.194	0.216	0.200	0.208	0.225	0.235	0.245	0.257	0.270	0.288	0.313	0.340	0.375	0.430	0.537
20	0.262	0.274	0.301	0.338	0.384	0.424	0.556	0.464	0.511	0.599	0.643	0.680	0.714	0.750	0.789	0.829	0.862	0.893	0.926	0.960
22	0.234	0.239	0.251	0.268	0.291	0.313	0.396	0.336	0.365	0.428	0.464	0.497	0.531	0.570	0.616	0.670	0.720	0.772	0.831	0.903
เฉลี่ย	**0.206**	**0.211**	**0.221**	**0.236**	**0.254**	**0.271**	**0.332**	**0.288**	**0.310**	**0.354**	**0.379**	**0.401**	**0.424**	**0.451**	**0.483**	**0.522**	**0.561**	**0.607**	**0.669**	**0.767**
27	0.218	0.226	0.241	0.258	0.276	0.292	0.343	0.307	0.325	0.361	0.382	0.400	0.419	0.440	0.467	0.500	0.534	0.573	0.626	0.712
29	0.326	0.343	0.362	0.377	0.389	0.398	0.419	0.405	0.412	0.425	0.431	0.436	0.441	0.447	0.453	0.461	0.468	0.476	0.487	0.506
31	0.173	0.176	0.181	0.189	0.198	0.206	0.239	0.215	0.227	0.252	0.268	0.283	0.299	0.319	0.345	0.379	0.416	0.462	0.529	0.646
เฉลี่ย	**0.239**	**0.248**	**0.261**	**0.275**	**0.288**	**0.299**	**0.334**	**0.309**	**0.321**	**0.346**	**0.360**	**0.373**	**0.386**	**0.402**	**0.422**	**0.447**	**0.473**	**0.504**	**0.547**	**0.621**
จำนวน	1377	796	549	592	591	369	563	598	663	657	649	490	568	594	719	653	551	589	586	399
ร้อยละ	10.969	6.341	4.373	4.716	4.708	2.940	4.485	4.764	5.282	5.234	5.170	3.903	4.525	4.732	5.728	5.202	4.389	4.692	4.668	3.179

จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4.34 พบว่าค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มของกลุ่มที่ 1 มีค่าตั้งแต่ 0.127 ถึง 0.326 กลุ่มที่ 2 มีค่าตั้งแต่ 0.132 ถึง 0.343 กลุ่มที่ 3 มีค่าตั้งแต่ 0.136 ถึง 0.362 กลุ่มที่ 4 มีค่าตั้งแต่ 0.142 ถึง 0.377 กลุ่มที่ 5 มีค่าตั้งแต่ 0.151 ถึง 0.422 กลุ่มที่ 6 มีค่าตั้งแต่ 0.156 ถึง 0.461 กลุ่มที่ 7 มีค่าตั้งแต่ 0.171 ถึง 0.569 กลุ่มที่ 8 มีค่าตั้งแต่ 0.160 ถึง 0.496 กลุ่มที่ 9 มีค่าตั้งแต่ 0.165 ถึง 0.535 กลุ่มที่ 10 มีค่าตั้งแต่ 0.177 ถึง 0.601 กลุ่มที่ 11 มีค่าตั้งแต่ 0.186 ถึง 0.643 กลุ่มที่ 12 มีค่าตั้งแต่ 0.195 ถึง 0.680 กลุ่มที่ 13 มีค่าตั้งแต่ 0.205 ถึง 0.714 กลุ่มที่ 14 มีค่าตั้งแต่ 0.218 ถึง 0.750 กลุ่มที่ 15 มีค่าตั้งแต่ 0.237 ถึง 0.789 กลุ่มที่ 16 มีค่าตั้งแต่ 0.265 ถึง 0.829 กลุ่มที่ 17 มีค่าตั้งแต่ 0.297 ถึง 0.862 กลุ่มที่ 18 มีค่าตั้งแต่ 0.342 ถึง 0.893 กลุ่มที่ 19 มีค่าตั้งแต่ 0.415 ถึง 0.926 กลุ่มที่ 20 มีค่าตั้งแต่ 0.506 ถึง 0.960 กล่าวได้ว่า กลุ่มที่ 20 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกมากที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถสูงที่สุด ในทางกลับกัน กลุ่มที่ 1 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบถูกน้อยที่สุด ดังนั้นเป็นกลุ่มที่มีความสามารถต่ำที่สุด

เมื่อพิจารณาการจำแนกกลุ่มความสามารถแฝงระดับบุคคลกับข้อสอบตามเนื้อหาที่ใช้ในการสอบ เป็นรายสาระการเรียนรู้ ซึ่งแต่ละสาระการเรียนรู้ มีรายละเอียดดังนี้ เรื่องจำนวนและการดำเนินการ ได้แก่ ข้อสอบข้อที่ 2, 4, 8, และ 9 สาระการเรียนรู้เรื่องการวัด ได้แก่ข้อ 16 และ 18 สาระการเรียนรู้เรื่องพีชคณิต ได้แก่ข้อ 6, 10, 13, 15, 20 และ 22 สาระการเรียนรู้เรื่อง การวิเคราะห์ข้อมูลและความน่าจะเป็น ได้แก่ข้อ 27, 29 และ 31 และในที่นี้กำหนดให้ กลุ่มเก่ง(A) คือ กลุ่มที่มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกอยู่ในระดับสูง มีค่าตั้งแต่ 0.61 ขึ้นไป กลุ่มปานกลาง (B) คือ กลุ่มที่มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกอยู่ในระดับปานกลาง โดยมีค่าตั้งแต่ 0.40-0.60 และ กลุ่มอ่อน (C) คือกลุ่มที่มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกอยู่ในระดับต่ำ โดยมีค่าน้อยกว่า 0.40 ซึ่งพบว่า กลุ่มแฝงระดับบุคคลกลุ่มที่ 20 หรือกลุ่มเก่งทั้ง 4 สาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของทั้ง 4 สาระการเรียนรู้มีค่าอยู่ในระดับสูง เป็นจำนวน 399 คน คิดเป็นร้อยละ 3.179

กลุ่มแฝงระดับบุคคลกลุ่มที่ 19 หรือกลุ่มเก่ง 3 สาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของสาระการเรียนรู้การวิเคราะห์ข้อมูลและความน่าจะเป็น มีค่าอยู่ในระดับปานกลาง แต่สาระอื่นๆ อยู่ในระดับสูง เป็นจำนวน 586 คน คิดเป็นร้อยละ 4.668

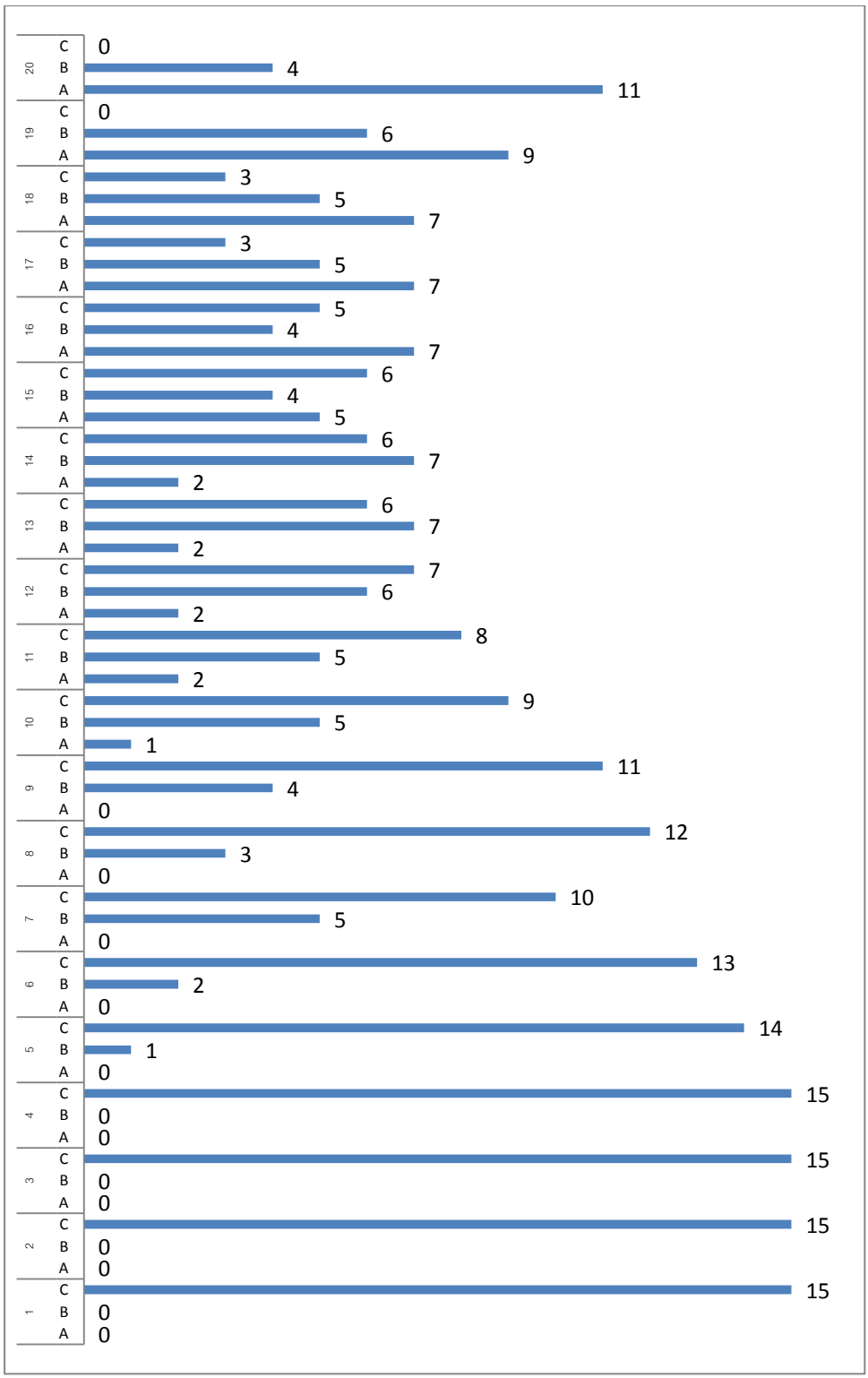
กลุ่มแฝงระดับบุคคลกลุ่มที่ 14, 15, 16, 17 และ 18 หรือกลุ่มเก่ง 1 สาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของสาระการเรียนรู้เรื่องการวัด อยู่ในระดับสูง แต่สาระอื่นๆ อยู่ในระดับปานกลาง เป็นจำนวน 3,106 คน คิดเป็นร้อยละ 24.743

กลุ่มแฝงระดับบุคคลกลุ่มที่ 12 และ 13 หรือ กลุ่มปานกลาง 2 สาระ อ่อน 2 สาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของสาระการเรียนรู้เรื่องการวัด และพีชคณิตอยู่ในระดับปานกลาง แต่สาระอื่นๆ อยู่ในระดับต่ำ เป็นจำนวน 1,058 คน คิดเป็นร้อยละ 8.428

กลุ่มแฝงระดับบุคคลกลุ่มที่ 7, 8, 9, 10 และ 11 หรือ กลุ่มปานกลาง 1 สาระ อ่อน 3 สาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของสาระการเรียนรู้เรื่องการวัดอยู่ในระดับปานกลาง แต่สาระอื่น ๆ อยู่ในระดับต่ำ เป็นจำนวน 3,130 คน คิดเป็นร้อยละ 24.934

กลุ่มแฝงระดับบุคคลกลุ่มที่ 1, 2, 3, 4, 5 และ 6 หรือกลุ่มอ่อนทุกสาระ มีค่าเฉลี่ยของความน่าจะเป็นในการตอบข้อสอบถูกของสาระการเรียนรู้ทุกสาระอยู่ในระดับต่ำ เป็นจำนวน 4,274 คน คิดเป็นร้อยละ 34.048

จากตารางที่ 4.32 และ 4.33 แสดงรายละเอียดการจำแนกกลุ่มความสามารถแฝงระดับบุคคลกับข้อสอบตามเนื้อหาที่ใช้ในการสอบเป็นรายข้อ ได้ดังภาพที่ 4.14 และกำหนดให้ความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง ปานกลาง ต่ำ แทนด้วย A, B และ C ตามลำดับ



ภาพที่ 4.14 กราฟแท่งแสดงจำนวนข้อสอบจำแนกตามระดับความน่าจะเป็นในการตอบข้อสอบถูกของกลุ่มแฝงระดับบุคคล 20 กลุ่มแฝง

หมายเหตุ จากภาพระดับความน่าจะเป็นในการตอบข้อสอบถูกระดับสูง ปานกลาง ต่ำ แทนด้วย A, B และ C ตามลำดับ

จากภาพพบว่ากลุ่มแฝงกลุ่มที่ 20 เป็นกลุ่มเก่งเกือบทุกข้อ มีลักษณะ 11A 4B 0C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 11 ข้อ ระดับปานกลาง 4 ข้อ และไม่มีระดับต่ำเลย โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลางได้แก่ข้อสอบเรื่อง การให้เหตุผล ความสัมพันธ์และฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 399 คน คิดเป็นร้อยละเท่ากับ 3.179

กลุ่มแฝงระดับบุคคลกลุ่มที่ 19 เป็นกลุ่มเก่งเกือบทุกข้อ มีลักษณะ 9A 6B 0C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 9 ข้อ ระดับปานกลาง 6 ข้อ และไม่มีระดับต่ำเลย โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลางได้แก่ข้อสอบเรื่อง การให้เหตุผล สมการกำลังสองความสัมพันธ์และฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 586 คน คิดเป็นร้อยละเท่ากับ 4.668

กลุ่มแฝงระดับบุคคลกลุ่มที่ 18 เป็นกลุ่มอ่อนบางข้อ มีลักษณะ 7A 5B 3C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 7 ข้อ ระดับปานกลาง 5 ข้อ และระดับต่ำ 3 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล ความสัมพันธ์ระหว่างฟังก์ชันในรูปของสมการ และตรีโกณมิติกับการหาความยาวด้าน ซึ่งคิดเป็นจำนวน 589 คน คิดเป็นร้อยละเท่ากับ 4.692

กลุ่มแฝงระดับบุคคลกลุ่มที่ 17 เป็นกลุ่มอ่อนบางข้อ มีลักษณะ 7A 5B 3C เช่นเดียวกับกลุ่มที่ 18 นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 7 ข้อ ระดับปานกลาง 5 ข้อ และระดับต่ำ 3 ข้อ แต่เมื่อพิจารณารายละเอียดข้อสอบพบว่าเป็นคนละข้อกัน โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล และความสัมพันธ์ระหว่างฟังก์ชันในรูปของกราฟ และสมการ ซึ่งคิดเป็นจำนวน 551 คน คิดเป็นร้อยละเท่ากับ 4.389

กลุ่มแฝงระดับบุคคลกลุ่มที่ 16 เป็นกลุ่มอ่อนบางข้อ มีลักษณะ 7A 4B 5C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 7 ข้อ ระดับปานกลาง 4 ข้อ และระดับต่ำ 5 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับน้อยได้แก่ข้อสอบเรื่อง การให้เหตุผล ความสัมพันธ์ระหว่างฟังก์ชัน และสถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 653 คน คิดเป็นร้อยละเท่ากับ 5.202

กลุ่มแฝงระดับบุคคลกลุ่มที่ 15 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 5A 4B 6C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 5 ข้อ ระดับปานกลาง 4 ข้อ และระดับต่ำ 6 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสองความสัมพันธ์ระหว่างฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 719 คน คิดเป็นร้อยละเท่ากับ 5.728

กลุ่มแฝงระดับบุคคลกลุ่มที่ 14 เป็นกลุ่มปานกลางเป็นส่วนใหญ่ มีลักษณะ 2A 7B 6C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 2 ข้อ ระดับปานกลาง 7 ข้อ และระดับต่ำ 6 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสอง ความสัมพันธ์ระหว่างฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 594 คน คิดเป็นร้อยละเท่ากับ 4.732

กลุ่มแฝงระดับบุคคลกลุ่มที่ 13 เป็นกลุ่มปานกลางเป็นส่วนใหญ่ มีลักษณะ 2A 7B 6C เช่นเดียวกับกลุ่มที่ 14 นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 2 ข้อ ระดับปานกลาง 7 ข้อ และระดับต่ำ 6 ข้อ เมื่อพิจารณารายละเอียดแล้วพบว่ามีค่าความน่าจะเป็นในการตอบถูกแต่ละข้อแตกต่างกัน แต่มีค่าใกล้เคียงกัน โดยกลุ่มที่ 14 มีความน่าจะเป็นในการตอบข้อสอบถูกสูงกว่ากลุ่มที่ 13 ข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสอง ความสัมพันธ์ระหว่างฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 568 คน คิดเป็นร้อยละเท่ากับ 4.525

กลุ่มแฝงระดับบุคคลกลุ่มที่ 12 เป็นกลุ่มปานกลางเป็นส่วนใหญ่ มีลักษณะ 2A 6B 7C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 2 ข้อ ระดับปานกลาง 6 ข้อ และระดับต่ำ 7 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสอง ความสัมพันธ์ระหว่างฟังก์ชัน สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 490 คน คิดเป็นร้อยละเท่ากับ 3.903

กลุ่มแฝงระดับบุคคลกลุ่มที่ 11 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 2A 5B 8C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 2 ข้อ ระดับปานกลาง 5 ข้อ และระดับต่ำ 8 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสอง ความสัมพันธ์ระหว่างฟังก์ชัน ความน่าจะเป็น สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 649 คน คิดเป็นร้อยละเท่ากับ 5.170

กลุ่มแฝงระดับบุคคลกลุ่มที่ 10 เป็นกลุ่มเก่ง 1 ข้อ มีลักษณะ 1A 5B 9C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 1 ข้อ ระดับปานกลาง 5 ข้อ และระดับต่ำ 9 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับต่ำได้แก่ข้อสอบเรื่อง การให้เหตุผล จำนวนจริง สมการกำลังสอง ความสัมพันธ์ระหว่างฟังก์ชัน การประยุกต์ตรีโกณมิติกับการหามุม ความน่าจะเป็น สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 657 คน คิดเป็นร้อยละเท่ากับ 5.234

กลุ่มแฝงระดับบุคคลกลุ่มที่ 9 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 0A 4B 11C นั่นคือ ไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงเลย มีระดับปานกลาง 4 ข้อ และข้อที่เหลือเป็นระดับต่ำโดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลาง ได้แก่ ข้อสอบ

เรื่องความสัมพันธ์และฟังก์ชัน การประยุกต์ตรีโกณมิติกับการหาความยาวด้าน ลำดับ สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 663 คน คิดเป็นร้อยละเท่ากับ 5.282

กลุ่มแฝงระดับบุคคลกลุ่มที่ 8 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 0A 3B 12C นั่นคือ ไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงเลย มีระดับปานกลาง 3 ข้อ และข้อที่เหลือเป็นระดับต่ำ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลาง ได้แก่ ข้อสอบเรื่อง การประยุกต์ตรีโกณมิติกับการหาความยาวด้าน ลำดับ สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 598 คน คิดเป็นร้อยละเท่ากับ 4.764

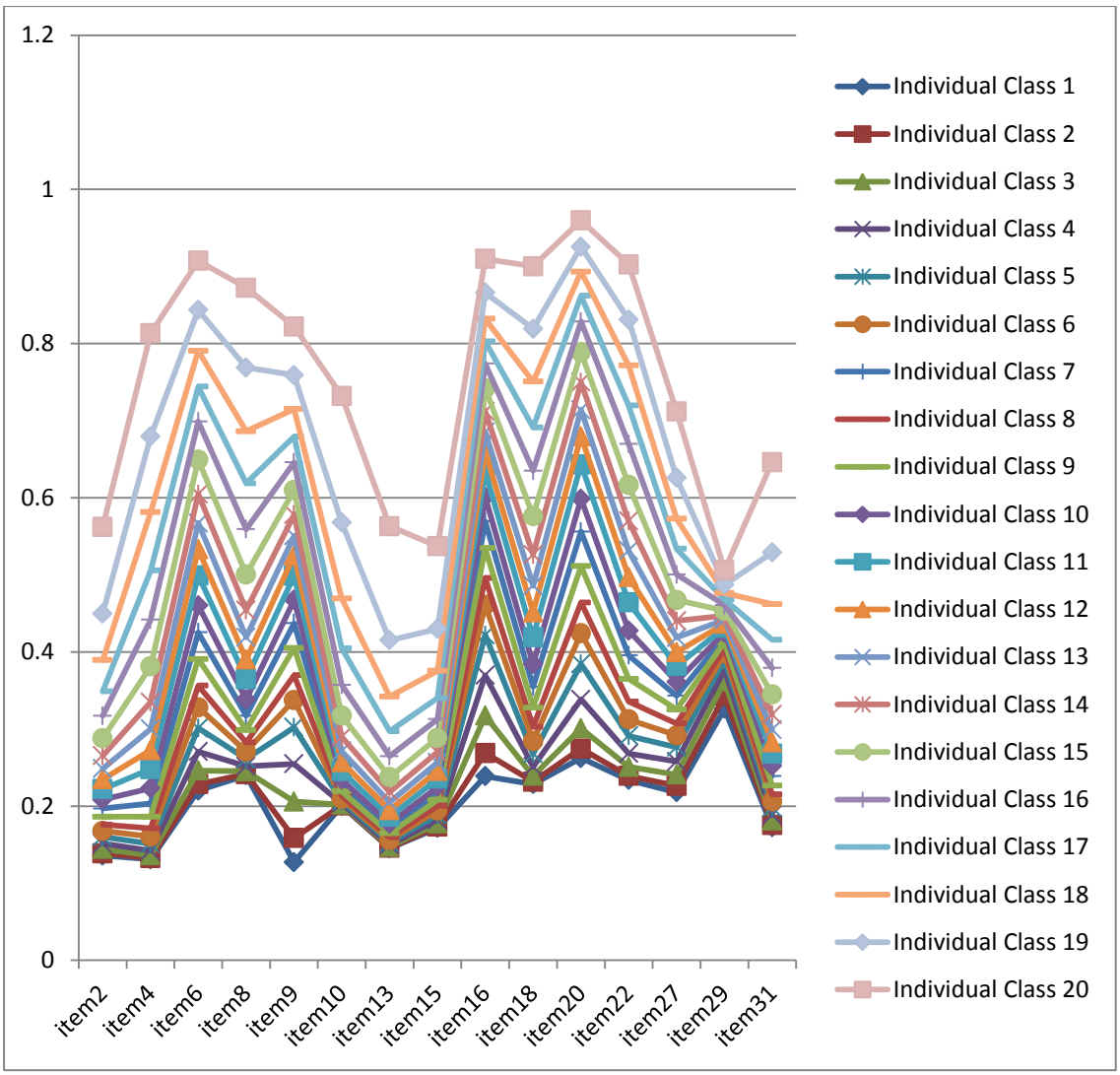
กลุ่มแฝงระดับบุคคลกลุ่มที่ 7 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 0A 5B 10C นั่นคือไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงเลย มีระดับปานกลาง 5 ข้อ และระดับต่ำ 10 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลางได้แก่ข้อสอบเรื่องจำนวนจริง ความสัมพันธ์ระหว่างฟังก์ชัน ตรีโกณมิติกับการหาความยาวด้าน ลำดับ สถิติและการวิเคราะห์ข้อมูลเบื้องต้น ซึ่งคิดเป็นจำนวน 563 คน คิดเป็นร้อยละเท่ากับ 4.485

กลุ่มแฝงระดับบุคคลกลุ่มที่ 6 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 0A 2B 13C นั่นคือไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงเลย มีระดับปานกลาง 2 ข้อ และระดับต่ำ 13 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลางได้แก่ข้อสอบเรื่องตรีโกณมิติกับการหาความยาวด้าน และลำดับ ซึ่งคิดเป็นจำนวน 369 คน คิดเป็นร้อยละเท่ากับ 2.940

กลุ่มแฝงระดับบุคคลกลุ่มที่ 5 เป็นกลุ่มอ่อนเป็นส่วนใหญ่ มีลักษณะ 0A 1B 14C นั่นคือ โดยไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงเลย มีระดับปานกลาง 1 ข้อ และระดับต่ำ 14 ข้อ โดยข้อที่มีความน่าจะเป็นในการตอบข้อสอบถูกระดับปานกลางได้แก่ข้อสอบเรื่อง ตรีโกณมิติกับการหาความยาวด้าน ซึ่งคิดเป็นจำนวน 591 คน คิดเป็นร้อยละเท่ากับ 4.708

กลุ่มแฝงระดับบุคคลกลุ่มที่ 4, 3 ,2 และ 1 เป็นกลุ่มอ่อนทุกด้าน มีลักษณะ 0A 0B 15C นั่นคือไม่มีนักเรียนคนใดมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง และปานกลางเลย เมื่อพิจารณาความน่าจะเป็นในการตอบข้อสอบถูกพบว่าแต่ละกลุ่มมีค่าใกล้เคียงกัน แต่กลุ่ม 4 มีค่าสูงกว่า กลุ่ม 3 กลุ่ม 2 และ กลุ่ม 1 ตามลำดับ และปานกลางเลย ซึ่งคิดเป็นจำนวน 592 คน 549 คน 796 คน และ 1,377 คน ตามลำดับ คิดเป็นร้อยละเท่ากับ 4.716, 4.373, 6.341 และ 10.969 ตามลำดับ ในที่นี้ผู้วิจัยกำหนดรหัสต่อท้ายด้วยเลข 1 ถึง 4 เพื่อแสดงถึงค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบถูกตามลำดับจากมากไปน้อย

จากค่าความน่าจะเป็นตารางที่ 4.34 แสดงดังภาพที่ 4.15



Chulalongkorn University

ภาพที่ 4.15 ความน่าจะเป็นที่จะตอบข้อสอบในแต่ละข้อถูกของแต่ละกลุ่มแฝงระดับบุคคลจากการจำแนกกลุ่มแฝงของข้อมูลเชิงประจักษ์จำนวน 15 ข้อ

2.1.3 การคำนวณหาความสามารถของโรงเรียนโดยใช้สัดส่วนของจำนวนกลุ่มแฝงระดับบุคคล สำหรับผลการคำนวณ แสดงดังตารางที่ 4.35 และ 4.36

ตารางที่ 4.35 จำนวนคนของแต่ละกลุ่มความสามารถแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียน สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

โรงเรียนที่	กลุ่มแฝงที่																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	80	26	26	17	9	5	4	4	5	5	4	0	2	2	0	0	0	0	0	0
2	58	14	9	2	7	1	0	4	6	1	1	0	1	1	0	0	0	0	0	0
3	62	18	18	11	6	1	4	4	2	3	3	0	3	1	1	0	0	0	0	0
4	61	28	12	12	13	2	3	5	8	2	1	0	1	0	0	0	0	0	0	0
5	141	39	40	25	19	12	7	6	8	5	3	0	1	2	1	0	0	0	0	0
...																				
45	93	25	15	12	15	8	17	19	20	13	15	13	15	14	27	45	34	68	98	781
...																				
399	50	9	11	4	0	1	0	3	5	0	2	1	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	2	4	6	215

จากตารางที่ 4.35 แสดงจำนวนรูปแบบการตอบในแต่ละกลุ่มแฝงระดับบุคคลซึ่งจำแนกตามโรงเรียนจำนวน 400 โรงเรียน จะพบว่า โรงเรียนที่ 1 ไม่มีนักเรียนคนใดเลยถูกจำแนกอยู่ในกลุ่มแฝงที่ 12, 15, 16, 17, 18, 19 และ 20 แต่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2 และ 3 มากที่สุดตามลำดับ โรงเรียนที่ 2 ไม่มีนักเรียนคนใดเลขถูกจำแนกอยู่ในกลุ่มแฝงที่ 7, 12, 15, 16, 17, 18, 19 และ 20 แต่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2 และ 3 มากที่สุดตามลำดับ โรงเรียนที่ 3 ไม่มีนักเรียนคนใดเลยถูกจำแนกอยู่กลุ่มแฝงที่ 12, 16, 17, 18, 19 และ 20 แต่ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2 และ 3 มากที่สุดตามลำดับ โรงเรียนที่ 45 ถูกจำแนกอยู่กลุ่มแฝงที่ 6 และ 4 จำนวนน้อยที่สุด และกลุ่มแฝงที่ 20, 19 และ 1 มากที่สุดตามลำดับ โรงเรียนที่ 400 ไม่มีนักเรียนคนใดเลขถูกจำแนกอยู่ในกลุ่มแฝงที่ 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15 และ 16 แต่ถูกจำแนกอยู่กลุ่มแฝงที่ 20, 19 และ 18 มากที่สุดตามลำดับ ซึ่งจำนวนคนในกลุ่มแฝงระดับบุคคลของแต่ละโรงเรียนจะนำไปใช้ในการคำนวณความสามารถของโรงเรียน โดยคำนวณจากสัดส่วนจำนวนคนในแต่ละกลุ่มแฝง สำหรับจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนต่อไป ดังตารางที่ 4.36

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ 4.36 สัดส่วนของจำนวนนักเรียนในแต่ละกลุ่มความสามารถแฝง
สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

โรงเรียนที่	กลุ่มแฝงที่																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.423	0.138	0.138	0.090	0.048	0.026	0.021	0.021	0.026	0.026	0.021	0.000	0.011	0.011	0.000	0.000	0.000	0.000	0.000	0.000
2	0.552	0.133	0.086	0.019	0.067	0.010	0.000	0.038	0.057	0.010	0.010	0.000	0.010	0.010	0.000	0.000	0.000	0.000	0.000	0.000
3	0.453	0.131	0.131	0.080	0.044	0.007	0.029	0.029	0.015	0.022	0.022	0.000	0.022	0.007	0.007	0.000	0.000	0.000	0.000	0.000
4	0.412	0.189	0.081	0.081	0.088	0.014	0.020	0.034	0.054	0.014	0.007	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.456	0.126	0.129	0.081	0.061	0.039	0.023	0.019	0.026	0.016	0.010	0.000	0.003	0.006	0.003	0.000	0.000	0.000	0.000	0.000
6	0.165	0.044	0.033	0.099	0.022	0.022	0.022	0.055	0.055	0.066	0.044	0.066	0.033	0.033	0.055	0.033	0.044	0.055	0.000	0.055
7	0.542	0.136	0.096	0.056	0.062	0.011	0.017	0.017	0.028	0.011	0.006	0.011	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8	0.677	0.097	0.097	0.000	0.032	0.000	0.000	0.032	0.032	0.000	0.000	0.000	0.000	0.032	0.000	0.000	0.000	0.000	0.000	0.000
9	0.408	0.140	0.124	0.064	0.073	0.041	0.010	0.041	0.025	0.006	0.025	0.016	0.010	0.003	0.000	0.000	0.006	0.003	0.003	0.000
10	0.606	0.085	0.117	0.011	0.043	0.053	0.011	0.064	0.000	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000
...																				
45	0.069	0.019	0.011	0.009	0.011	0.006	0.013	0.014	0.015	0.010	0.011	0.010	0.011	0.010	0.020	0.033	0.025	0.050	0.073	0.580
...																				
399	0.581	0.105	0.128	0.047	0.000	0.012	0.000	0.035	0.058	0.000	0.023	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.004	0.000	0.000	0.000	0.009	0.017	0.026	0.939

จากตารางที่ 4.36 พบว่าความสามารถของโรงเรียนซึ่งคำนวณจากสัดส่วนของจำนวนคนในกลุ่มแฝงระดับบุคคลของทุกโรงเรียน กลุ่มแฝงกลุ่มที่ 20 เป็นกลุ่มที่มีความสามารถสูงที่สุด และกลุ่มแฝงกลุ่มที่ 1 เป็นกลุ่มที่มีความสามารถต่ำที่สุด ซึ่งโรงเรียนที่ 400 มีสัดส่วนความน่าจะเป็นของนักเรียนในกลุ่ม 20 มากที่สุด โดยมีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 20 จำนวน 215 คน จากทั้งหมด 229 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.939 และไม่มีนักเรียนถูกจำแนกอยู่ในกลุ่มที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.000 และรองลงมาคือโรงเรียนที่ 45 มีนักเรียนถูกจำแนกอยู่ในกลุ่มแฝงที่ 20 จำนวน 781 คน จากทั้งหมด 1,347 คน คิดเป็นสัดส่วนความสามารถเท่ากับ 0.580 และมีนักเรียนจำนวน 93 คน ถูกจำแนกอยู่ในกลุ่มแฝงที่ 1 คิดเป็นสัดส่วนความสามารถเท่ากับ 0.069

2.1.4 การจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน k (School Class) สำหรับการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน นำความสามารถของโรงเรียนในขั้นตอนที่ 2.1.3 มาวิเคราะห์จำแนกกลุ่มความสามารถแฝงระดับโรงเรียน โดยใช้แพคเกจ Mclust ในโปรแกรม R ในการวิเคราะห์ข้อมูล ผลการจำแนกกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.37 4.38 และ 4.39

ตารางที่ 4.37 ค่า BIC ของการจำแนกกลุ่มระดับโรงเรียน สูงสุด 3 อันดับแรก สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

VEI,9	VEI,8	VEI,6
45543.49	45380.78	45133.03

จากตารางที่ 4.36 พบว่า ค่า BIC ของการจำแนกกลุ่มแฝงจำนวน 9 กลุ่ม เป็นค่าที่เหมาะสมในการจำแนกกลุ่ม เนื่องจากเป็นค่าที่มากที่สุด (maximum) ซึ่งสำหรับการแบ่งแบบ VEI มีค่า BIC เท่ากับ 45543.49

ตารางที่ 4.38 จำนวนโรงเรียนและสัดส่วนของแต่ละกลุ่มความสามารถแฝงระดับโรงเรียน สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มที่	จำนวนโรงเรียน	สัดส่วน	ร้อยละ
1	81	0.198	19.8
2	90	0.224	22.4
3	28	0.070	7.0
4	26	0.063	6.3
5	47	0.122	12.2
6	91	0.230	23.0
7	23	0.058	5.8
8	12	0.030	3.0
9	2	0.005	0.5
รวม	400	1	100
Log. Likelihood : 23418.83, BIC : 45543.49			

จากตารางที่ 4.38 พบว่าค่า log likelihood เท่ากับ 23418.83 ค่า BIC เท่ากับ 45543.49 และกลุ่มแฝงที่ 1 ถึง 9 มีจำนวนโรงเรียน 81, 90, 28, 26, 47, 91, 23, 12 และ 2 โรงเรียน และมีจำนวนโรงเรียนในแต่ละกลุ่มแฝงเมื่อเทียบกับจำนวนทั้งหมดของกลุ่มแฝงที่ 1 ถึง 9 คิดเป็นร้อยละ 19.8, 22.4, 7.0, 6.3, 1.22, 23.0, 5.8, 3.0 และ 0.5 ตามลำดับ

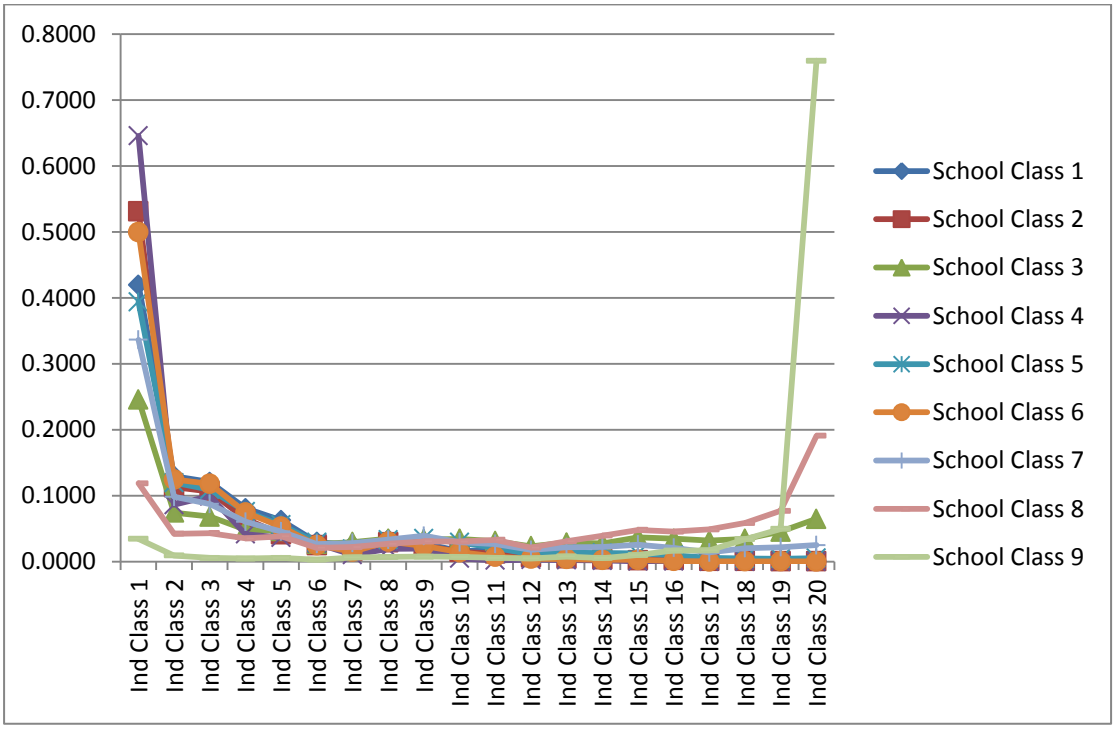
ตารางที่ 4.39 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกในแต่ละข้อของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับบุคคล	ลักษณะ	ร้อยละของจำนวนคนในกลุ่มแฝงระดับบุคคล	กลุ่มแฝงระดับโรงเรียน								
			1	2	3	4	5	6	7	8	9
1	0A 0B 15C L	10.969	0.419	0.531	0.246	0.646	0.394	0.499	0.336	0.119	0.035
2	0A 0B 15C M2	6.341	0.129	0.113	0.074	0.086	0.120	0.124	0.098	0.042	0.009
3	0A 0B 15C M1	4.373	0.121	0.106	0.068	0.099	0.109	0.118	0.087	0.043	0.006
4	0A 0B 15C H	4.716	0.081	0.065	0.048	0.042	0.076	0.074	0.060	0.035	0.004
5	0A 1B 14C	4.708	0.063	0.041	0.046	0.036	0.056	0.053	0.046	0.038	0.006
6	0A 2B 13C	2.940	0.030	0.025	0.027	0.028	0.029	0.025	0.027	0.021	0.003
7	0A 5B 10C	4.485	0.025	0.016	0.030	0.010	0.019	0.015	0.028	0.022	0.006
8	0A 3B 12C	4.764	0.034	0.029	0.035	0.019	0.033	0.030	0.033	0.027	0.007
9	0A 4B 11C	5.282	0.027	0.024	0.033	0.020	0.035	0.024	0.039	0.031	0.007
10	1A 5B 9C	5.234	0.016	0.015	0.034	0.005	0.029	0.014	0.029	0.030	0.007
11	2A 5B 8C	5.170	0.017	0.012	0.031	0.002	0.021	0.007	0.027	0.033	0.006
12	2A 6B 7C	3.903	0.008	0.007	0.023	0.003	0.011	0.004	0.018	0.021	0.005
13	2A 7B 6C L	4.525	0.011	0.005	0.029	0.002	0.014	0.004	0.022	0.031	0.008
14	2A 7B 6C H	4.732	0.008	0.004	0.028	0.002	0.013	0.003	0.022	0.039	0.005
15	5A 4B 6C	5.728	0.004	0.002	0.037	0.000	0.012	0.003	0.026	0.048	0.010
16	6A 4B 5C	5.202	0.003	0.002	0.035	0.000	0.010	0.001	0.021	0.045	0.017
17	7A 5B 3C L	4.389	0.001	0.001	0.032	0.000	0.006	0.000	0.014	0.049	0.017
18	7A 5B 3C H	4.692	0.002	0.001	0.034	0.000	0.005	0.000	0.020	0.058	0.034
19	9A 6B 0C	4.668	0.002	0.000	0.045	0.000	0.004	0.000	0.022	0.077	0.049
20	11A 4B 0C	3.179	0.001	0.000	0.065	0.000	0.005	0.000	0.025	0.191	0.759

จากตารางที่ 4.39 พบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 1 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อของแต่ละกลุ่มแฝงระดับบุคคล ตั้งแต่ 0.001 ถึง 0.419 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 2 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.531 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 3 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.023 ถึง 0.246 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 4 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.646 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 5 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่

0.004 ถึง 0.394 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 6 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.000 ถึง 0.499 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 7 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.014 ถึง 0.336 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 8 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.021 ถึง 0.191 กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 9 มีค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบแต่ละข้อตั้งแต่ 0.003 ถึง 0.759

จากค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่มข้างต้น สามารถแสดงดังภาพที่ 4.16



ภาพที่ 4.16 ค่าเฉลี่ยความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบแต่ละข้อถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน ในแต่ละกลุ่ม สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

สำหรับผลการจำแนกกลุ่มแฝงระดับโรงเรียนแสดงดังตารางที่ 4.40

ตารางที่ 4.40 รหัสโรงเรียนที่ได้จากการจำแนกกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับโรงเรียน								
1	2	3	4	5	6	7	8	9
1	2	6	8	14	5	47	40	45
3	11	42	10	18	7	50	59	400
4	13	49	19	29	15	58	63	
9	17	55	20	35	16	61	64	
12	24	57	38	36	21	78	330	
25	27	60	130	39	22	97	333	
28	33	62	175	43	23	108	334	
31	34	67	177	46	26	117	337	
44	37	81	185	53	30	136	338	
51	41	85	197	65	32	137	341	
52	48	99	199	66	54	147	369	
56	69	107	218	74	71	155	389	
68	77	112	220	84	82	158		
70	80	116	237	109	86	159		
72	87	123	257	110	89	171		
73	94	188	272	111	90	217		
75	103	215	284	115	91	228		
76	105	223	309	127	93	343		
79	114	296	316	173	96	348		
83	119	331	321	192	98	350		
88	121	332	323	203	101	351		
92	125	335	326	209	104	363		
95	129	336	377	216	131	370		
100	133	339	385	224	135			
102	139	340	388	232	140			
106	141	342	390	239	142			
113	144	354		260	143			

ตารางที่ 4.40 (ต่อ) รหัสโรงเรียนที่ได้จากการจำแนกกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับโรงเรียน								
1	2	3	4	5	6	7	8	9
118	148	371		261	146			
120	151			271	150			
122	152			273	154			
124	153			279	163			
126	156			285	164			
128	157			287	165			
132	160			290	167			
134	161			312	168			
138	162			314	170			
145	166			317	172			
149	174			320	176			
169	181			327	179			
178	184			344	182			
180	190			347	186			
183	191			359	187			
189	193			360	195			
198	194			362	200			
204	196			372	201			
206	211			374	202			
219	213			380	205			
222	221				207			
226	227				208			
234	230				210			
240	231				212			
245	236				214			
250	238				225			
252	243				229			
270	244				233			
274	247				235			
275	248				241			

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.40 (ต่อ) รหัสโรงเรียนที่ได้จากการจำแนกกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับโรงเรียน								
1	2	3	4	5	6	7	8	9
276	249				242			
280	254				246			
281	263				251			
282	264				253			
288	267				255			
289	278				256			
298	292				258			
299	293				259			
302	294				262			
303	300				265			
305	313				266			
306	315				268			
307	319				269			
318	322				277			
346	324				283			
349	325				286			
352	328				291			
355	353				295			
364	356				297			
381	358				301			
384	366				304			
392	373				308			
394	375				310			
398	378				311			
	379				329			
	382				345			
	383				357			
	386				361			
	387				365			
	395				367			

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.40 (ต่อ) รหัสโรงเรียนที่ได้จากการจำแนกกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับโรงเรียน								
1	2	3	4	5	6	7	8	9
	396 397 399				368 376 391 393			
81 โรงเรียน	90 โรงเรียน	28 โรงเรียน	26 โรงเรียน	47 โรงเรียน	91 โรงเรียน	23 โรงเรียน	12 โรงเรียน	2 โรงเรียน
20.25%	22.50%	7%	6.50%	11.75%	22.75%	5.75%	3%	0.50%

จากตาราง 4.40 พบว่าผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนโรงเรียนที่ถูกจำแนกอยู่ในกลุ่มแฝงระดับโรงเรียนกลุ่มที่ 1 มีจำนวน 81 โรงเรียน คิดเป็นร้อยละ 20.25 กลุ่มที่ 2 จำนวน 90 โรงเรียน คิดเป็นร้อยละ 22.5 กลุ่มที่ 3 จำนวน 28 โรงเรียนคิดเป็นร้อยละ 7 กลุ่มที่ 4 จำนวน 26 โรงเรียน คิดเป็นร้อยละ 6.5 กลุ่มที่ 5 จำนวน 47 โรงเรียน คิดเป็นร้อยละ 11.75 กลุ่มที่ 6 จำนวน 91 โรงเรียน คิดเป็นร้อยละ 22.75 กลุ่มที่ 7 จำนวน 23 โรงเรียน คิดเป็นร้อยละ 5.75 กลุ่มที่ 8 จำนวน 12 โรงเรียน คิดเป็นร้อยละ 3 และกลุ่มที่ 9 จำนวน 2 โรงเรียน คิดเป็นร้อยละ 0.5

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

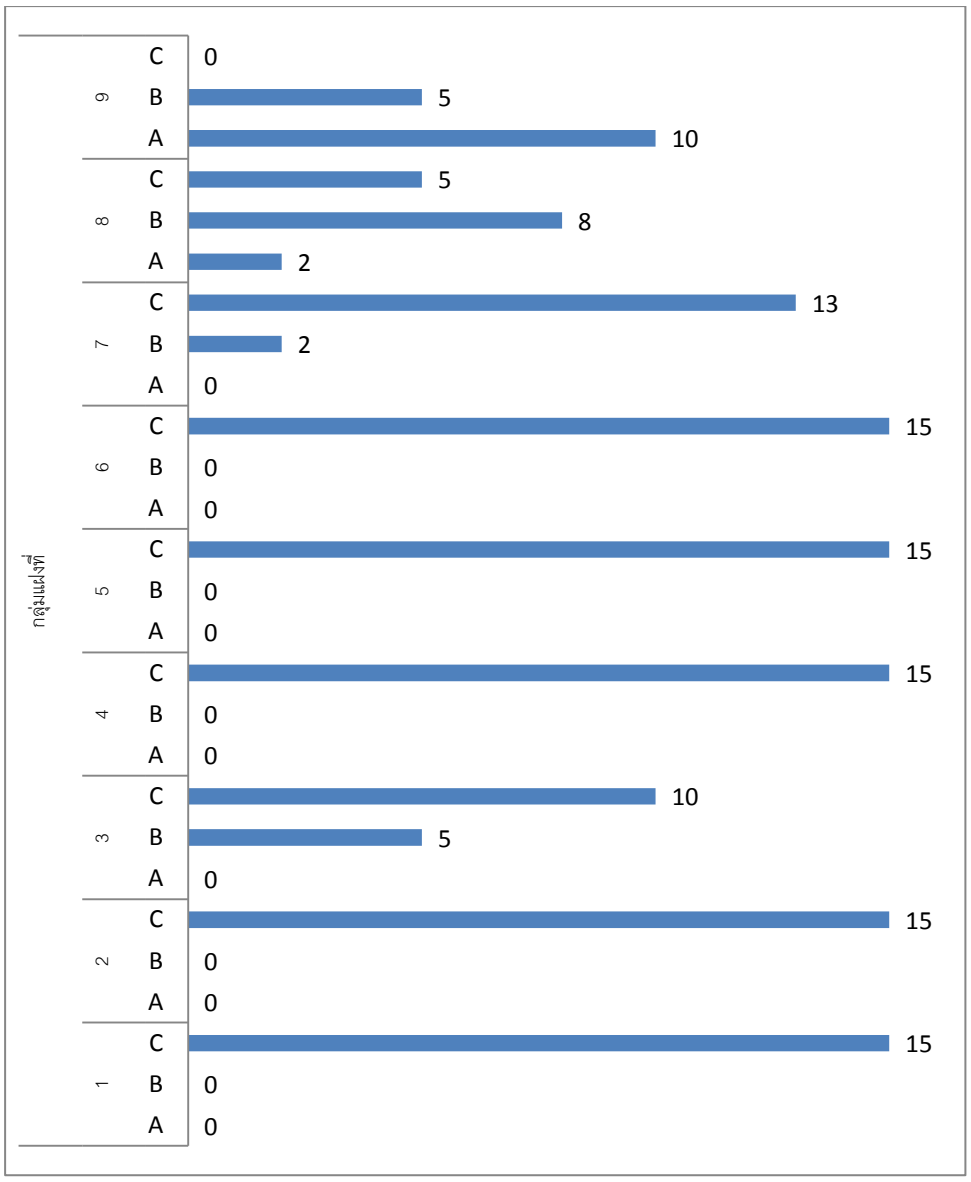
2.1.5 การคำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน ผลการจำแนกกลุ่มระดับโรงเรียน แสดงดังตารางที่ 4.41 และภาพที่ 4.17

ตารางที่ 4.41 ผลการคำนวณความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียน สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

กลุ่มแฝงระดับโรงเรียน	ลักษณะ	ข้อ															เฉลี่ย
		2	4	6	8	9	10	13	15	16	18	20	22	27	29	31	
1	0A 0B 15C 2	0.153	0.152	0.275	0.262	0.221	0.209	0.153	0.185	0.337	0.263	0.339	0.277	0.254	0.358	0.192	0.242
2	0A 0B 15C 3	0.149	0.145	0.260	0.255	0.198	0.206	0.151	0.181	0.313	0.253	0.319	0.265	0.245	0.351	0.186	0.232
3	0A 5B 10C	0.228	0.276	0.435	0.383	0.386	0.288	0.214	0.245	0.504	0.406	0.518	0.426	0.355	0.399	0.276	0.356
4	0A 0B 15C 5	0.143	0.138	0.244	0.248	0.174	0.203	0.148	0.177	0.287	0.242	0.297	0.252	0.234	0.343	0.180	0.221
5	0A 0B 15C 1	0.163	0.167	0.298	0.277	0.246	0.217	0.160	0.192	0.362	0.282	0.366	0.298	0.268	0.364	0.202	0.257
6	0A 0B 15C 4	0.148	0.144	0.258	0.253	0.198	0.205	0.150	0.180	0.312	0.250	0.316	0.262	0.243	0.351	0.185	0.230
7	0A 2B 13C	0.190	0.212	0.357	0.321	0.306	0.246	0.182	0.214	0.423	0.334	0.432	0.352	0.305	0.379	0.233	0.299
8	2A 8B 5C	0.304	0.403	0.571	0.504	0.520	0.380	0.285	0.310	0.635	0.537	0.660	0.558	0.448	0.433	0.361	0.461
9	10A 5B 0C	0.496	0.711	0.827	0.782	0.750	0.641	0.490	0.479	0.845	0.813	0.891	0.820	0.646	0.488	0.574	0.684

จากตารางพบว่า กลุ่มแฝงระดับโรงเรียนกลุ่มแฝงที่ 1 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.152 ถึง 0.358 กลุ่มแฝงที่ 2 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.145 ถึง 0.351 กลุ่มแฝงที่ 3 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.214 ถึง 0.518 กลุ่มแฝงที่ 4 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.138 ถึง 0.343 กลุ่มแฝงที่ 5 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.160 ถึง 0.366 กลุ่มแฝงที่ 6 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.144 ถึง 0.351 กลุ่มแฝงที่ 7 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.182 ถึง 0.432 กลุ่มแฝงที่ 8 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.285 ถึง 0.660 และกลุ่มแฝงที่ 9 มีความน่าจะเป็นในการตอบข้อสอบแต่ละข้อถูกตั้งแต่ 0.479 ถึง 0.891

จากตารางที่ 4.41 แสดงรายละเอียดการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนกับข้อสอบตามเนื้อหาที่ใช้ในการสอบรายข้อ ได้ดังภาพที่ 4.17 และกำหนดให้โอกาสตอบข้อสอบถูกในระดับมาก ปานกลาง น้อย แทนด้วย A, B และ C ตามลำดับ



ภาพที่ 4.17 กราฟแท่งแสดงจำนวนข้อสอบจำแนกตามระดับความน่าจะเป็นในการตอบข้อสอบถูกของกลุ่มแฝงระดับโรงเรียน 9 กลุ่มแฝง

หมายเหตุ จากภาพระดับความน่าจะเป็นในการตอบข้อสอบถูกระดับสูง ปานกลาง ต่ำ แทนด้วย A, B และ C ตามลำดับ

จากภาพพบว่า กลุ่มแฝงที่ 9 เป็นกลุ่มที่มีลักษณะ 10A 5B 0C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 10 ข้อ ปานกลาง 5 ข้อ

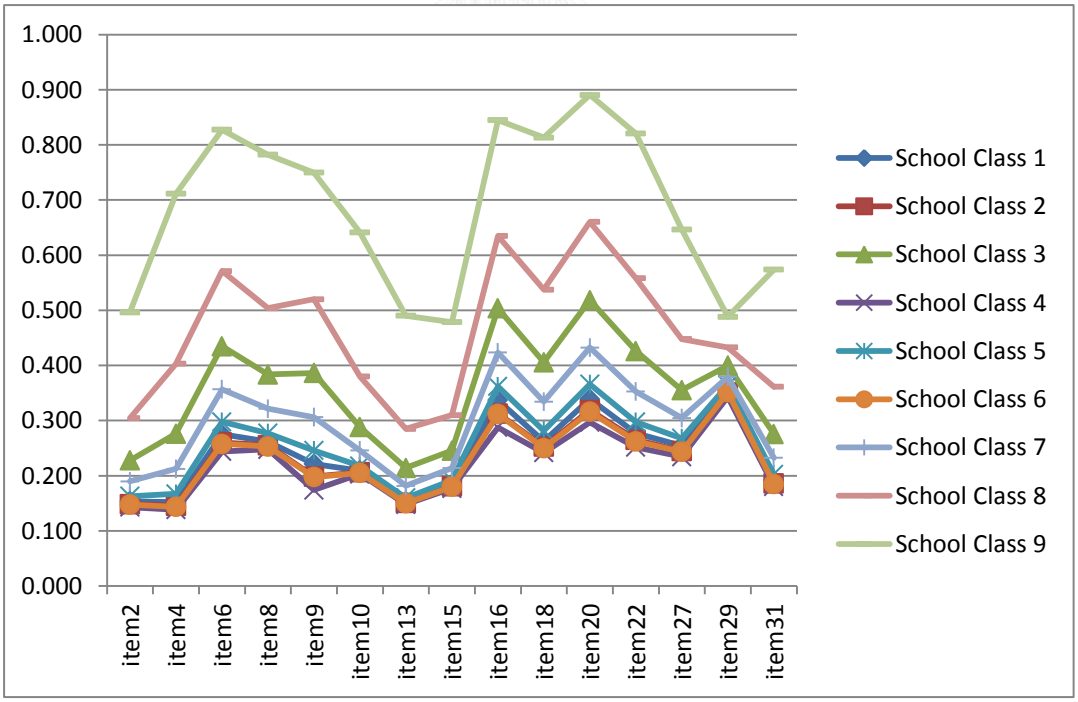
กลุ่มแฝงที่ 8 เป็นกลุ่มที่มีลักษณะ 2A 8B 5C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูง 2 ข้อ ปานกลาง 8 ข้อ และต่ำ 5 ข้อ

กลุ่มแฝงที่ 7 เป็นกลุ่มที่มีลักษณะ 0A 2B 13C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับปานกลาง 2 ข้อ และต่ำ 13 ข้อ

กลุ่มแฝงที่ 3 เป็นกลุ่มที่มีลักษณะ 0A 5B 10C นั่นคือมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับปานกลาง 5 ข้อ และต่ำ 10 ข้อ

กลุ่มแฝงระดับโรงเรียนกลุ่มที่ 1, 2, 4, 5 และ 6 เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบถูกต่ำในทุกข้อ อาจกล่าวได้ว่าทั้ง 5 กลุ่ม มีลักษณะ 0A 0B 15C คือไม่มีความน่าจะเป็นในการตอบข้อสอบข้อใดในระดับสูงและปานกลางเลย เมื่อพิจารณาค่าความน่าจะเป็นเฉลี่ยพบว่ากลุ่มแฝงระดับโรงเรียนกลุ่มที่ 5 มีค่าสูงสุด รองลงมาเป็นกลุ่มที่ 1, 2, 6 และ 4 ตามลำดับ ในที่นี้ผู้วิจัยกำหนดรหัสต่อท้ายด้วยตัวเลข 1 หลัก คือเลข 1 ถึง 5 เพื่อแสดงถึงค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบถูกเรียงตามลำดับจากมากไปน้อย

เมื่อนำค่าความน่าจะเป็นจากตารางที่ 4.41 มาเขียนกราฟ จะได้ดังภาพที่ 4.18



ภาพที่ 4.18 ความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละกลุ่มแฝงระดับโรงเรียนจากการจำแนกกลุ่มแฝง สำหรับข้อมูลเชิงประจักษ์ 15 ข้อ

จากตารางที่ 4.41 และภาพที่ 4.18 พบว่า เมื่อจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน กลุ่มแฝงที่ 9 เป็นกลุ่มที่มีความสามารถสูงสุดในทุกสาระ และกลุ่มแฝงที่ 8 เป็นกลุ่มที่มีความสามารถสูงเป็นอันดับ 2 ในทุกสาระ รองลงมาเป็นกลุ่มแฝงที่ 3 และ 7 ตามลำดับ สำหรับกลุ่มแฝงที่ 5, 1, 2, 6 และ 4 เป็นกลุ่มที่มีความสามารถใกล้เคียงกันตามลำดับจากสูงไปต่ำ

โดยสรุปผลการจำแนกกลุ่มความสามารถแฝงจากข้อมูลเชิงประจักษ์นี้เมื่อพิจารณาลักษณะของแต่ละกลุ่มแฝง เทียบกับข้อมูลสังเกตได้ของแต่ละกลุ่มแฝงพบว่า ในกลุ่มแฝงระดับโรงเรียน กลุ่มแฝงที่ 9 (10A 5B 0C) ซึ่งเป็นกลุ่มที่มีความสามารถสูงที่สุด โดยมีความน่าจะเป็นในการตอบข้อสอบถูกในระดับสูงจำนวน 10 ข้อ ระดับปานกลาง 5 ข้อ และไม่มีความน่าจะเป็นในการตอบข้อสอบข้อใดอยู่ในระดับต่ำ โดยข้อสอบที่มีความน่าจะเป็นในการตอบข้อสอบถูกในระดับปานกลาง ได้แก่ข้อที่ 2 วัดเนื้อหาเรื่องการให้เหตุผล, ข้อ 13 วัดเนื้อหาเรื่องกราฟของความสัมพันธ์และฟังก์ชัน, ข้อ 15 วัดเนื้อหาเรื่อง ความสัมพันธ์และฟังก์ชัน ประยุกต์กับสมการ, ข้อ 29 และ ข้อ 31 วัดเนื้อหาเรื่อง สถิติและการวิเคราะห์ข้อมูลเบื้องต้น และข้อสอบที่มีความน่าจะเป็นในการตอบถูกอยู่ระดับสูงวัดเนื้อหาเรื่องตรีโกณมิติ อสมการ จำนวนจริง ความน่าจะเป็น และลำดับ มีโรงเรียนที่ถูกจัดอยู่ในกลุ่มแฝงที่ 9 เพียง 2 โรงเรียน (โรงเรียนที่ 45 และโรงเรียนที่ 400) คิดเป็นร้อยละ 0.5 โดยทั้งสองโรงเรียนเป็นโรงเรียนขนาดใหญ่ที่มีชื่อเสียงโดยเฉพาะด้านคณิตศาสตร์ และวิทยาศาสตร์ มีระบบการสอบคัดเลือกนักเรียนเข้าศึกษาต่อในระดับชั้นมัธยมศึกษาปีที่ 4 ทำให้มีนักเรียนที่มีความสามารถสูงจำนวนมาก โดยจากผลการวิเคราะห์พบว่ามีนักเรียนที่ถูกจัดอยู่ในกลุ่มแฝงระดับบุคคลกลุ่มแฝงที่ 20 จำนวนมาก เมื่อสังเกตกลุ่มแฝงระดับบุคคล กลุ่มแฝงที่ 20 พบว่าเป็นกลุ่มแฝงที่มีความสามารถสูงในทุกด้าน โดยมีความสามารถสูงในด้าน พีชคณิต และการวัดมากที่สุด แต่ในด้านการวิเคราะห์ข้อมูลและความน่าจะเป็น มีความสามารถไม่แตกต่างกันกับกลุ่มแฝงอื่นมากนัก ซึ่งสอดคล้องกับสภาพจริงของโรงเรียน

กลุ่มแฝงระดับโรงเรียน กลุ่มแฝงที่ 8 (2A 8B 5C) เป็นกลุ่มที่มีความสามารถรองลงมา มีโรงเรียนที่ถูกจัดอยู่ในกลุ่มนี้จำนวน 12 โรงเรียน คิดเป็นร้อยละ 3 ซึ่งเป็นโรงเรียนในสังกัดคณะกรรมการการอุดมศึกษา สังกัดเอกชน และมีทั้งโรงเรียนขนาดใหญ่ และโรงเรียนขนาดกลาง นอกจากนี้ในกลุ่มแฝงที่ 8 มีโรงเรียนที่ใช้หลักสูตรทางเลือกและใช้หลักสูตรแกนกลางในการเรียน ซึ่งสะท้อนให้เห็นว่า นักเรียนที่เรียนในโรงเรียนทางเลือก มีความสามารถในการแข่งขันทางวิชาการได้เท่าเทียมกับโรงเรียนปกติ และเมื่อพิจารณาโรงเรียนในสังกัดคณะกรรมการการอุดมศึกษา พบว่าถูกจำแนกอยู่ในกลุ่มแฝงระดับโรงเรียนกลุ่มที่ 9, 8 และ 3 ซึ่งเป็นกลุ่มแฝงที่มีความสามารถสูงเป็น 3 อันดับแรก

นอกจากนี้พบว่ากลุ่มแฝงระดับโรงเรียน กลุ่มแฝงที่ 4 เป็นกลุ่มที่มีความสามารถต่ำสุดซึ่งมีโรงเรียนถูกจัดอยู่ในกลุ่มนี้เพียง 26 โรงเรียน ซึ่งเป็นโรงเรียนในสังกัดกรมส่งเสริมการปกครองส่วน

ท้องถิ่น จำนวน 5 โรงเรียน จากทั้งหมด 302 โรงเรียน คิดเป็นร้อยละ 1.66 โดยเป็นโรงเรียนขนาดเล็ก 4 โรงเรียนจากทั้งหมด 70 โรงเรียน คิดเป็นร้อยละ 5.71 และขนาดกลาง 1 โรงเรียน จากทั้งหมด 164 โรงเรียน คิดเป็นร้อยละ 0.61 โรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 17 โรงเรียน จากทั้งหมด 2,497 โรงเรียน คิดเป็นร้อยละ 0.68 โดยเป็นโรงเรียนขนาดเล็ก 8 โรงเรียนจากทั้งหมด 516 โรงเรียน คิดเป็นร้อยละ 1.55 และขนาดกลาง 9 โรงเรียน จากทั้งหมด 1,161 โรงเรียน คิดเป็นร้อยละ 0.78 โรงเรียนสังกัดเอกชน จำนวน 4 โรงเรียน จากทั้งหมด 571 โรงเรียน คิดเป็นร้อยละ 0.70 โดยเป็นโรงเรียนขนาดเล็ก 2 โรงเรียน จากทั้งหมด 121 โรงเรียน คิดเป็นร้อยละ 1.65 และขนาดกลาง 2 โรงเรียน จากทั้งหมด 216 โรงเรียน คิดเป็นร้อยละ 0.93 และเมื่อเทียบสัดส่วนโรงเรียนที่ถูกจำแนกอยู่กลุ่มแฝงที่ 4 ในแต่ละสังกัดตามขนาดโรงเรียนพบว่าโรงเรียนในสังกัด กรมส่งเสริมการปกครองส่วนท้องถิ่นที่มีขนาดเล็ก เป็นโรงเรียนที่มีสัดส่วนมากที่สุด จึงควรได้รับการพัฒนาการศึกษาด้านคณิตศาสตร์เป็นลำดับแรก สำหรับโรงเรียนในกลุ่มแฝงระดับโรงเรียนกลุ่มที่ 4 ซึ่งมีนักเรียนที่ถูกจัดอยู่ในกลุ่มแฝงระดับบุคคลกลุ่มที่ 1, 2 และ 3 จำนวนมาก และ เมื่อสังเกตกลุ่มแฝงระดับบุคคลกลุ่มที่ 1, 2 และ 3 พบว่าเป็นกลุ่มแฝงที่มีความสามารถต่ำในทุกด้าน แต่ในด้านสถิติและการวิเคราะห์ข้อมูล มีความสามารถไม่แตกต่างจากกลุ่มอื่นมากนัก

นอกจากนี้จากผลการจำแนกกลุ่มความสามารถแฝงนี้สังเกตได้ว่าไม่มีโรงเรียนขนาดใหญ่และขนาดใหญ่พิเศษ และโรงเรียนในสังกัดคณะกรรมการการอุดมศึกษาถูกจำแนกอยู่ในกลุ่มแฝงระดับโรงเรียน กลุ่มแฝงที่ 4 แสดงให้เห็นว่าโรงเรียนขนาดใหญ่ และขนาดใหญ่พิเศษมีนักเรียนที่มีความสามารถคณิตศาสตร์สูงกว่าโรงเรียนขนาดเล็ก และปานกลาง

จากที่กล่าวข้างต้นจะเห็นว่าควรเร่งพัฒนาการศึกษาทางคณิตศาสตร์ให้กับโรงเรียนที่อยู่ในกลุ่มแฝงที่ 4 จำนวน 26 โรงเรียน ทั้งนี้ควรพัฒนาในด้านจำนวนและการดำเนินการ พีชคณิต การวัด โดยศึกษาแนวทางการพัฒนาความสามารถทางคณิตศาสตร์ หลักสูตรที่ใช้ และการจัดการศึกษาทางคณิตศาสตร์ จากโรงเรียนในสังกัดคณะกรรมการการอุดมศึกษา

จากผลการวิเคราะห์ข้อมูลตอนที่ 2 และ 3 สามารถสรุปเป็นตารางได้ดังต่อไปนี้

ตารางที่ 4.42 เปรียบเทียบผลการวิเคราะห์ข้อมูลจำลอง 6 ข้อ 15 ข้อ และ ข้อมูลเชิงประจักษ์ 15 ข้อ ในการจำแนกกลุ่มความสามารถแฝงของแต่ละบุคคลด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

รายละเอียด	**ผลการวิเคราะห์ข้อมูล**		
	ข้อมูลจำลอง 6 ข้อ	**ข้อมูลจำลอง 15 ข้อ**	**ข้อมูลเชิงประจักษ์ 15 ข้อ**
1. ลักษณะข้อมูล	ผลการตอบข้อสอบที่มีการให้คะแนนสองค่า โดยจำลองข้อมูลให้เป็น 6 กลุ่มแฝงที่มีความสามารถแตกต่างกัน	ผลการตอบข้อสอบที่มีการให้คะแนนสองค่า โดยจำลองข้อมูลให้เป็น 6 กลุ่มแฝงที่มีความสามารถแตกต่างกัน	ผลการตอบข้อสอบที่มีการให้คะแนนสองค่า จากการทดสอบระดับชาติขั้นพื้นฐาน (O-NET) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6
2. จำนวนกลุ่มตัวอย่าง	จำนวนโรงเรียน 24 โรงเรียน จำนวนนักเรียน 32,370 คน	จำนวนโรงเรียน 24 โรงเรียน จำนวนนักเรียน 32,370 คน	จำนวนโรงเรียน 400 โรงเรียน จำนวนนักเรียน 72,295 คน
3. การประมาณค่าพารามิเตอร์	ค่าอำนาจจำแนก มีค่า 1.020 ถึง 1.537 ค่าความยาก มีค่า -0.727 ถึง 0.591 ค่าการเดา มีค่า 0.00 ถึง 0.25 Log likelihood มีค่า -126318	ค่าอำนาจจำแนก มีค่า 1.127 ถึง 1.832 ค่าความยาก มีค่า -0.473 ถึง 0.575 ค่าการเดา มีค่า 0.071 ถึง 0.25 Log likelihood มีค่า -1304231	ค่าอำนาจจำแนก มีค่า 0.353 ถึง 4.162 ค่าความยาก มีค่า 0.889 ถึง 1.922 ค่าการเดา มีค่า 0.00 ถึง 0.25 Log likelihood มีค่า -1598121
4. จำนวนรูปแบบการตอบ	64	13,209	12,553
5. ค่าความสามารถ	-1.310 ถึง 1.154	-1.678 ถึง 1.674	0.453 ถึง 2.431
6. ผลการจำแนกกลุ่มความสามารถแฝงระดับบุคคล	8 กลุ่มแฝง (maximum BIC = 4884.123)	16 กลุ่มแฝง (maximum BIC = 3257869)	20 กลุ่มแฝง (maximum BIC = 3059008)
7. ผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน	6 กลุ่มแฝง (maximum BIC = 1017.884)	6 กลุ่มแฝง (maximum BIC = 2473.793)	9 กลุ่มแฝง (maximum BIC = 45543.49)
8. ผลการตรวจสอบความถูกต้องของการจำแนกกลุ่มความสามารถแฝง	มีความสอดคล้องกับข้อมูลจำลอง ทั้งสิ้น 24 ครั้งจากการวิเคราะห์ทั้งหมด 25 ครั้ง คิดเป็นร้อยละ 96 แสดงว่าโมเดลที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ถูกต้อง	มีความสอดคล้องกับข้อมูลจำลอง ทั้งสิ้น 25 ครั้งจากการวิเคราะห์ทั้งหมด 25 ครั้ง คิดเป็นร้อยละ 100 แสดงว่าโมเดลที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ถูกต้อง	-

บทที่ 5
สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์ 3 ข้อ คือ (1) เพื่อพัฒนาโมเดลการประมาณค่ากลุ่มความสามารถแฝงของบุคคลตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ (2) เพื่อตรวจสอบความถูกต้อง แม่นยำของการจำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคลโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลจำลอง และ (3) เพื่อวิเคราะห์จำแนกกลุ่มคุณลักษณะแฝงของแต่ละบุคคล โดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ จากข้อมูลผลการตอบข้อสอบวิชาคณิตศาสตร์ในการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT) ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ สำหรับจำแนกกลุ่มแฝงของบุคคล เป็นโมเดลใหม่ซึ่งพัฒนาโดยการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งเป็นการพัฒนาโมเดลต่อยอดจากโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับแบบผสมที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 1 และ 2 พารามิเตอร์ จึงมีความจำเป็นต้องตรวจสอบความถูกต้อง แม่นยำในการจำแนกกลุ่มแฝง ผู้วิจัยจึงสร้างข้อมูลจำลองสำหรับใช้ในการจำแนกกลุ่มแฝงระดับโรงเรียนโดยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ จากนั้นนำโมเดลดังกล่าวไปใช้กับข้อมูลจริง ในส่วนของการพัฒนาโมเดลงานวิจัยนี้ได้พัฒนาโดยใช้โปรแกรม R ซึ่งมีการนำแพคเกจ IRTOYS (Partchev, Partchev, & Suggests, 2015) ltm (Rizopoulos, 2006) และ Mclust (Fraley & Raftery, 2015) มาประยุกต์ใช้ โดยแพคเกจ IRTOYS ถูกนำมาใช้ในการจำลองข้อมูลผลตอบข้อสอบ แพคเกจ ltm ถูกนำมาใช้ในส่วนของการประมานค่าพารามิเตอร์ข้อสอบ และความสามารถแฝงระดับบุคคล และ แพคเกจ Mclust ถูกนำมาใช้ในส่วนของการจำแนกกลุ่มแฝง

5.1 สรุปผลการวิจัย

จากผลการวิเคราะห์ข้อมูลสามารถสรุปผลการวิจัย จำแนกตามวัตถุประสงค์ได้ดังนี้

5.1.1. ผลการพัฒนาโมเดลการประมาณค่าความสามารถตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ (MMixIRT)

สำหรับโมเดลการประมาณค่าความสามารถตามทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ ที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ โมเดลการวิเคราะห์กลุ่มแฝง และโมเดลพหุระดับ มีรายละเอียดโดยย่อดังนี้

สำหรับโมเดลการตอบสนองข้อสอบแบบผสมการวัดพหุระดับ (Multilevel Mixed-measurement IRT) แบบพหุมิติ (3PL model) ความน่าจะเป็นในการตอบข้อสอบถูก $P\left(\underset{\sim}{y_t} \right)$ ของการสังเกตกลุ่มของการตอบสนองในหน่วยลดหลั่น t ในกรณีนี้เป็นโรงเรียน แสดงดังนี้

$$P\left(\underset{\sim}{y_t} \right) = \sum_{k=1}^{K} \pi_k P\left(\underset{\sim}{y_t} | k \right) \quad \text{.....(5.1)}$$

เมื่อ k = 1,... , K เป็นระดับกลุ่ม (เช่น โรงเรียน t) ที่เป็นกลุ่มแฝง คล้ายกับใน LC

π_k เป็นความน่าจะเป็นของหน่วยลำดับขั้น (เช่น โรงเรียน) ในกลุ่มแฝง k

ความน่าจะเป็นแบบมีเงื่อนไข $P\left(\underset{\sim}{y_t} | \ k \right)$ ของการสังเกต กลุ่มของการตอบสนอง $\boldsymbol{n}_t$ ภายในแต่ละโรงเรียน k แสดงดังนี้

$$P\left(\underset{\sim}{y_t} | k \right) = \prod_{j=1}^{n_t} P\left(y_{tj} | k \right) \quad \text{.....(5.2)}$$

ความน่าจะเป็นของการสังเกต เวคเตอร์การตอบสนองแต่ละบุคคลในโรงเรียน t เป็นอิสระทางสถิติของการตอบสนองของแต่ละบุคคลในกลุ่มโรงเรียน k

ความน่าจะเป็นแบบมีเงื่อนไขนี้ถูกแสดงดังนี้

$$P\left(y_{tj} | k \right) = \sum_{g=1}^{G} \pi\left(g | k \right) P\left(\underset{\sim}{y_{tj}} \middle| \underset{\sim}{\theta_j}, g \right) \quad \text{.....(5.3)}$$

ในที่นี้จะเห็นว่าความน่าจะเป็นของกลุ่มบุคคล $\pi\left(g | k \right)$ มีความไม่แน่นอนบนกลุ่มโรงเรียน k อย่างไรก็ตาม ความน่าจะเป็นในการตอบสนองข้อสอบของแต่ละบุคคล j แสดงดังนี้

$$P\left(\underset{\sim}{y_{tj}} \middle| \underset{\sim}{\theta_j}, g \right) = \prod_{i=1}^{I} P\left(\underset{\sim}{y_{tji}} \middle| \underset{\sim}{\theta_j}, g \right) \quad \text{.....(5.4)}$$

j = 1,...,n_t ในโรงเรียน t ในกลุ่มของข้อ I ซึ่งขึ้นอยู่กับกลุ่มบุคคล k และคุณลักษณะของแต่ละบุคคล $\underset{\sim}{\theta_j}$ เท่านั้น ซึ่งไม่เป็นฟังก์ชันของกลุ่มโรงเรียน k ซึ่งบ่งบอกว่าแต่ละกลุ่มบุคคล g เป็น

หนึ่งเดียวที่ถูกนิยามโดยโมเดลการวัด $P\left(y_{ti} \middle| \underset{\sim}{\theta_j}, g \right)$ และความไม่แปรเปลี่ยนข้ามโรงเรียน และกลุ่มโรงเรียน

นอกจากนี้มีข้อตกลงเบื้องต้นว่ามีความอิสระของการตอบข้อสอบ (local independence) ภายในแต่ละกลุ่มบุคคล และให้การตอบสนอง $\underset{\sim}{\theta_j}$ เป็นอิสระทางสถิติ

ความน่าจะเป็นแสดงดังนี้

$$\mathrm{P}\left(\mathrm{y_{tji}} \middle| \underset{\sim}{\theta_j}, \mathrm{g} \right) = \mathrm{c_i} + \frac{1-\mathrm{c_i}}{1+\exp(-\eta_{ji})} \quad \text{.....(5.5)}$$

เมื่อ η_{ji} มีค่าเท่ากับ $\beta_i + \lambda_i \theta_j$ โดยที่ β_i , λ_i , c_i แทนพารามิเตอร์ความยาก อำนาจจำแนก และการเดา ตามลำดับ

จากโมเดลที่เสนอข้างต้นนี้ผู้วิจัยสามารถสรุปเป็นขั้นตอนในการวิเคราะห์ ดังนี้

ขั้นตอนที่ 1 การประมาณค่าพารามิเตอร์ของข้อมูล จากรูปแบบการตอบข้อสอบทั้งหมด โดยพารามิเตอร์ที่ประมาณค่าได้แก่ discriminant (a), difficulty(b), guessing(c) และ ability(θ) และคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ของแต่ละคนที่มีรูปแบบการตอบเหมือนกันจะตอบถูกในแต่ละข้อ

ขั้นตอนที่ 2 จำแนกกลุ่มความสามารถแฝงระดับบุคคล (Individual Class) โดยนำความน่าจะเป็นในขั้นตอนที่ 1 มาวิเคราะห์กลุ่มแฝง

ขั้นตอนที่ 3 คำนวณหาความสามารถของโรงเรียนโดยใช้สัดส่วนความน่าจะเป็นของจำนวนกลุ่มแฝงระดับบุคคลในแต่ละโรงเรียน

ขั้นตอนที่ 4 จำแนกกลุ่มความสามารถแฝงระดับโรงเรียน (School Class) โดยนำความสามารถของโรงเรียนในขั้นตอนที่ 3 มาวิเคราะห์กลุ่มแฝง

ขั้นตอนที่ 5 คำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน โดยคำนวณจากผลคูณของสัดส่วนของจำนวนคนในกลุ่มแฝงต่อจำนวนคนทั้งหมดกับ โอกาสในการตอบถูกของแต่ละกลุ่มแฝงในแต่ละข้อ โดยใช้สูตร $\sum_{k=1}^{K} \pi_k P(\theta_k)$ เมื่อ k = 1,... , K เป็นระดับกลุ่ม (เช่น โรงเรียน t) ที่เป็นกลุ่มแฝง และ π_k เป็นความน่าจะเป็นของหน่วยระดับโรงเรียนในกลุ่มแฝง k

5.1.2 ผลการทดสอบความถูกต้องแม่นยำในการจำแนกกลุ่มแฝงโดยใช้โมเดลที่พัฒนาขึ้นกับข้อมูลจำลอง

การจำลองกลุ่มแฝงของโรงเรียน ผู้วิจัยจำลองรูปแบบการตอบของแต่ละโรงเรียนตามความสามารถของนักเรียนในแต่ละโรงเรียนที่มีความเชี่ยวชาญความรู้ในด้านที่ต่างกัน รวมทั้งขนาด

ของโรงเรียนที่ต่างกัน โดยผู้วิจัยจำลองกลุ่มแฝงของโรงเรียนออกเป็น 6 กลุ่มแฝงตามความสามารถในการตอบข้อสอบ 3 สาระ ได้แก่

กลุ่มแฝงที่ 1 ซึ่งเป็นกลุ่มคละความสามารถ มีความสามารถสูงในสาระที่ 1 มีความสามารถปานกลางในสาระที่ 2 มีความสามารถระดับต่ำในสาระที่ 3

กลุ่มแฝงที่ 2 ซึ่งเป็นกลุ่มคละความสามารถ มีความสามารถปานกลางในสาระที่ 1 มีความสามารถต่ำในสาระที่ 2 มีความสามารถระดับสูงในสาระที่ 3

กลุ่มแฝงที่ 3 ซึ่งเป็นกลุ่มคละความสามารถ มีความสามารถต่ำในสาระที่ 1 มีความสามารถสูงในสาระที่ 2 มีความสามารถระดับปานกลางในสาระที่ 3

กลุ่มแฝงที่ 4 กลุ่มความสามารถปานกลางในทั้ง 3 สาระ

กลุ่มแฝงที่ 5 กลุ่มความสามารถระดับสูงในทั้ง 3 สาระ

กลุ่มแฝงที่ 6 กลุ่มความสามารถระดับต่ำทั้ง 3 สาระ

ในการจำลองข้อมูลนี้แต่ละกลุ่มแฝงจะจำลองโรงเรียนเป็น 4 ขนาด แตกต่างกัน รวม 32,370 คน ดังนั้นในงานวิจัยนี้จำลองข้อมูลทั้งหมด 24 โรงเรียน นอกจากนี้ผู้วิจัยจำลองผลการตอบข้อสอบใน 2 สถานการณ์ ได้แก่ ผลการตอบข้อสอบจำนวน 6 ข้อ ซึ่งแต่ละด้านถูกวัดด้วยข้อสอบ 2 ข้อ และผลการตอบข้อสอบจำนวน 15 ข้อ แต่ละด้านถูกวัดด้วยข้อสอบ 5 ข้อ

สำหรับการกำหนดค่าเริ่มต้นสำหรับการจำลองผลการตอบของนักเรียนที่มีความสามารถแตกต่างกัน ในการจำลองข้อมูล ในที่นี้ผู้วิจัยกำหนดค่าเริ่มต้น ซึ่งได้แก่ค่าความยากของข้อสอบ อำนาจจำแนก พารามิเตอร์การเดา สำหรับแต่ละกลุ่มแฝง และจำลองความสามารถภายใต้โค้งปกติ N(0,1) สำหรับแต่ละกลุ่มแฝง เช่นเดียวกับ Li, F., Cohen, A. S., Kim, S. H. and Cho, S. J. (2009).

ผลการวิเคราะห์ข้อมูลจำลองมีรายละเอียด ดังนี้

สถานการณ์ที่ 1 ผลการวิจัยที่ได้จากการวิเคราะห์ข้อมูลจำลองจำนวน 6 ข้อ พบว่าค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.00 ถึง 0.250 ค่าความยาก มีค่าตั้งแต่ -0.727 ถึง 0.591 และ ค่าอำนาจจำแนก มีค่าตั้งแต่ 1.020 ถึง 1.537 และมีค่า maximum log likelihood เท่ากับ -126,318 สำหรับผลการตอบข้อสอบของข้อมูลจำลองในระดับบุคคล จำนวน 32,370 คน มีทั้งหมด 64 รูปแบบ ซึ่งมีค่าความสามารถ (θ) อยู่ระหว่าง -1.310 ถึง 1.154 จากนั้นหาความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน และสัดส่วนของความสามารถของโรงเรียนจำแนกตามกลุ่มแฝงระดับบุคคล เพื่อใช้ในการจำแนกกลุ่มความสามารถแฝง พบว่า สามารถจำแนกกลุ่มแฝงระดับบุคคลได้ 16 กลุ่ม และระดับโรงเรียนได้ 6 กลุ่ม โดยมีค่า log likelihood เท่ากับ 2,832.997 และ 613.8179 ค่า BIC อยู่ที่ 4,884.12 และ 1,017.884 ตามลำดับ ซึ่งค่า

ดังกล่าวเป็นค่าที่มากที่สุด สำหรับใช้กำหนดจำนวนกลุ่มแฝงในระดับบุคคลและโรงเรียน นอกจากนี้ ผลการตรวจสอบความถูกต้องแม่นยำในการจำแนกกลุ่มแฝง พบว่าผลการจำแนกกลุ่มแฝงด้วยโมเดลที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ถูกต้องทั้งสิ้น 24 ครั้ง จากการวิเคราะห์ซ้ำ 25 ครั้ง คิดเป็นร้อยละ 96

สถานการณ์ที่ 2 ผลการวิจัยที่ได้จากการวิเคราะห์ข้อมูลจำลองจำนวน 15 ข้อ พบว่า ค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.071 ถึง 0.25 ค่าความยาก มีค่าตั้งแต่ -0.473 ถึง 0.575 และ ค่าอำนาจจำแนก มีค่าตั้งแต่ 1.127 ถึง 1.832 และมีค่า maximum log likelihood เท่ากับ -304231 สำหรับผลการตอบข้อสอบของข้อมูลจำลองจำนวน 32,370 คน มีทั้งหมด 13,209 รูปแบบ ซึ่งมีค่าความสามารถ (θ) อยู่ระหว่าง -1.678 ถึง 1.674 จากนั้นหาความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน และสัดส่วนของความสามารถของโรงเรียนจำแนกตามกลุ่มแฝงระดับบุคคล เพื่อใช้ในการจำแนกกลุ่มความสามารถแฝง พบว่า สามารถจำแนกกลุ่มแฝงทั้งระดับบุคคลได้ 16 กลุ่ม และระดับโรงเรียนได้ 6 กลุ่ม โดยมีค่า log likelihood เท่ากับ 163,9253 และ 1,430.758 ค่า BIC อยู่ที่ 3,257,869 และ 2,473.793 ตามลำดับ ซึ่งค่าดังกล่าวเป็นค่าที่มากที่สุด สำหรับใช้กำหนดจำนวนกลุ่มแฝงในระดับบุคคลและโรงเรียน นอกจากนี้ผลการตรวจสอบความถูกต้องแม่นยำในการจำแนกกลุ่มแฝง พบว่าผลการจำแนกกลุ่มแฝงด้วยโมเดลที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ถูกต้องทั้ง 25 ครั้ง คิดเป็นร้อยละ 100

ในการตรวจสอบความถูกต้องในการจำแนกกลุ่มเทียบกับข้อมูลที่จำลองแบบ 6 ข้อ และ 15 ข้อ พบว่าผลการจำแนกกลุ่มแฝงระดับโรงเรียนโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมที่พัฒนาขึ้นสามารถจำแนกกลุ่มความสามารถแฝงได้ตรงตามข้อมูลที่ผู้วิจัยกำหนดในเงื่อนไขการจำลองข้อมูลในเริ่มต้นเช่นเดียวกัน โดยจำแนกกลุ่มความสามารถแฝงของข้อมูลจำลอง 6 ข้อ ด้วยโมเดลที่พัฒนาขึ้นได้ถูกต้องทั้งสิ้น 24 ครั้ง จากการวิเคราะห์ซ้ำ 25 ครั้ง คิดเป็นร้อยละ 96 และสามารถจำแนกกลุ่มความสามารถแฝงสำหรับข้อมูลจำลอง 15 ข้อได้ถูกต้องทุกครั้ง คิดเป็นร้อยละ 100 ซึ่งแสดงให้เห็นว่าโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น สามารถใช้ในการจำแนกกลุ่มความสามารถแฝงของบุคคลจากผลการตอบข้อสอบแบบจำนวนข้อสอบสั้น และปานกลางได้ แต่สำหรับจำนวนข้อสอบที่มากกว่านี้ ผู้วิจัยไม่ได้ดำเนินการวิเคราะห์เนื่องจากมีข้อจำกัดทางด้านฮาร์ดแวร์ที่ใช้ในการวิเคราะห์ข้อมูล

5.1.3 ผลการจำแนกกลุ่มแฝงโดยใช้โมเดลที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์

ผลการวิจัยที่ได้จากการวิเคราะห์ข้อมูลเชิงประจักษ์ผลการตอบข้อสอบจำนวน 15 ข้อ โดยศึกษากับกลุ่มตัวอย่างจำนวน 72,295 คน มีทั้งหมด 12,553 รูปแบบ พบว่า ค่าพารามิเตอร์การเดา มีค่าตั้งแต่ 0.000 ถึง 0.250 ค่าความยาก มีค่าตั้งแต่ 0.889 ถึง 1.922 และ ค่าอำนาจจำแนก มีค่าตั้งแต่ 0.353 ถึง 4.162 และมีค่า maximum log likelihood เท่ากับ –598,121 ซึ่งมีค่าความสามารถ (θ) อยู่ระหว่าง -0.453 ถึง 2.431 จากนั้นหาความน่าจะเป็นแบบมีเงื่อนไขของการตอบข้อสอบถูกของนักเรียนที่มีความสามารถต่าง ๆ กัน และสัดส่วนของความสามารถของโรงเรียนจำแนกตามกลุ่มแฝงระดับบุคคล เพื่อใช้ในการจำแนกกลุ่มความสามารถแฝงระดับพบว่า สามารถจำแนกกลุ่มแฝงทั้งระดับบุคคลได้ 20 กลุ่มที่มีลักษณะความสามารถแตกต่างกัน ทั้งนี้ผู้วิจัยตั้งชื่อกลุ่มแฝงตามโอกาสการตอบข้อสอบถูกในแต่ละข้อ โดยกำหนดชื่อกลุ่มแฝงด้วยรหัสตัวเลขแสดงจำนวนข้อตามด้วยตัวอักษรภาษาอังกฤษ แสดงระดับของโอกาสการตอบข้อสอบถูก โดยกำหนดให้ ระดับมากมีค่ามากกว่า 0.6 แทนด้วย A ระดับปานกลาง มีค่า 0.4 ถึง 0.6 แทนด้วย B และ ระดับต่ำ มีค่าน้อยกว่า 0.4 แทนด้วย C สำหรับกลุ่มที่มีชื่อเดียวกัน ผู้วิจัยจะเขียนรหัสต่อท้ายด้วยตัวเลข 1 ถึง 4 เพื่อแสดงถึงความน่าจะเป็นในการตอบข้อสอบถูกเรียงลำดับเป็นสูงไปต่ำ

สรุปผลการจำแนกกลุ่มแฝงระดับบุคคลดังนี้

กลุ่มแฝงที่ 1 0A 0B 15C 4 มีจำนวน 1,377 คน คิดเป็นร้อยละ 10.969,

กลุ่มแฝงที่ 2 0A 0B 15C 3 มีจำนวน 796 คน คิดเป็นร้อยละ 6.341,

กลุ่มแฝงที่ 3 0A 0B 15C 2 มีจำนวน 549 คน คิดเป็นร้อยละ 4.373,

กลุ่มแฝงที่ 4 0A 0B 15C 1 มีจำนวน 592 คน คิดเป็นร้อยละ 4.716,

กลุ่มแฝงที่ 5 0A 1B 14C มีจำนวน 591 คน คิดเป็นร้อยละ 4.708,

กลุ่มแฝงที่ 6 0A 2B 13C มีจำนวน 369 คน คิดเป็นร้อยละ 2.940,

กลุ่มแฝงที่ 7 0A 5B 10C มีจำนวน 563 คน คิดเป็นร้อยละ 4.485,

กลุ่มแฝงที่ 8 0A 3B 12C มีจำนวน 598 คน คิดเป็นร้อยละ 4.764,

กลุ่มแฝงที่ 9 0A 4B 11C มีจำนวน 663 คน คิดเป็นร้อยละ 5.282,

กลุ่มแฝงที่ 10 1A 5B 9C มีจำนวน 657 คน คิดเป็นร้อยละ 5.234,

กลุ่มแฝงที่ 11 2A 5B 8C มีจำนวน 649 คน คิดเป็นร้อยละ 5.170,

กลุ่มแฝงที่ 12 2A 6B 7C มีจำนวน 490 คน คิดเป็นร้อยละ 3.903,

กลุ่มแฝงที่ 13 2A 7B 6C 1 มีจำนวน 568 คน คิดเป็นร้อยละ 4.525,

กลุ่มแฝงที่ 14 2A 7B 6C 2 มีจำนวน 594 คน คิดเป็นร้อยละ 4.732,

กลุ่มแฝงที่ 15 5A 4B 6C มีจำนวน 719 คน คิดเป็นร้อยละ 5.728,

กลุ่มแฝงที่ 16 6A 4B 5C มีจำนวน 653 คน คิดเป็นร้อยละ 5.202,
กลุ่มแฝงที่ 17 7A 5B 3C 1 มีจำนวน 551 คน คิดเป็นร้อยละ 4.389,
กลุ่มแฝงที่ 18 7A 5B 3C 2 มีจำนวน 589 คน คิดเป็นร้อยละ 4.692,
กลุ่มแฝงที่ 19 9A 6B 0C มีจำนวน 586 คน คิดเป็นร้อยละ 4.668,
กลุ่มแฝงที่ 20 11A 4B 0C มีจำนวน 399 คน คิดเป็นร้อยละ 3.179

สรุปผลการจำแนกกลุ่มแฝงระดับโรงเรียนได้ 9 กลุ่มแฝง โดยตั้งชื่อกลุ่มแฝงในลักษณะเดียวกับการตั้งชื่อในระดับบุคคล สำหรับกลุ่ม 0A 0B 15C มีชื่อกลุ่มซ้ำกัน 5 กลุ่ม ผู้วิจัยจึงกำหนดรหัสต่อท้ายด้วยตัวเลข 1 หลัก คือเลข 1 ถึง 5 เพื่อแสดงถึงค่าเฉลี่ยความน่าจะเป็นในการตอบข้อสอบถูกเรียงตามลำดับจากสูงไปต่ำ ผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนโรงเรียนที่ถูกจำแนกอยู่ในกลุ่มแฝงระดับโรงเรียน

กลุ่มแฝงที่ 1 0A 0B 15C 2 มีจำนวน 81 โรงเรียน คิดเป็นร้อยละ 20.25,
กลุ่มแฝงที่ 2 0A 0B 15C 3 จำนวน 90 โรงเรียน คิดเป็นร้อยละ 22.5,
กลุ่มแฝงที่ 3 0A 5B 10C จำนวน 28 โรงเรียนคิดเป็นร้อยละ 7,
กลุ่มแฝงที่ 4 0A 0B 15C 5 จำนวน 26 โรงเรียน คิดเป็นร้อยละ 6.5,
กลุ่มแฝงที่ 5 0A 0B 15C 1 จำนวน 47 โรงเรียน คิดเป็นร้อยละ 11.75,
กลุ่มแฝงที่ 6 0A 0B 15C 4 จำนวน 91 โรงเรียน คิดเป็นร้อยละ 22.75,
กลุ่มแฝงที่ 7 0A 2B 13C จำนวน 23 โรงเรียน คิดเป็นร้อยละ 5.75,
กลุ่มแฝงที่ 8 2A 8B 5C จำนวน 12 โรงเรียน คิดเป็นร้อยละ 3,
กลุ่มแฝงที่ 9 10A 5B 0C จำนวน 2 โรงเรียน คิดเป็นร้อยละ 0.5

ผลการคำนวณค่า log likelihood และ BIC สำหรับการจำแนกกลุ่มระดับบุคคลและโรงเรียนโดยมีค่า log likelihood เท่ากับ 1,542,335 และ 23,418.83 ค่า BIC อยู่ที่ 3,059,008 และ 45,543.49 ตามลำดับ ซึ่งค่าดังกล่าวเป็นค่าที่มากที่สุด สำหรับใช้กำหนดจำนวนกลุ่มแฝงในระดับบุคคลและโรงเรียน

ทั้งนี้จะเห็นได้ว่าเมื่อจำแนกกลุ่มความสามารถแฝงระดับโรงเรียน กลุ่มแฝงที่ 9 เป็นกลุ่มที่มีความสามารถสูงที่สุด ซึ่งมีจำนวนเพียง 2 โรงเรียน และกลุ่มแฝงที่ 8 เป็นกลุ่มที่มีความสามารถสูงรองลงมา เมื่อเปรียบเทียบผลการจำแนกกลุ่มความสามารถแฝงนี้กับลักษณะโรงเรียนซึ่งเป็นตัวแปรสังเกตได้พบว่า โรงเรียนในกลุ่มแฝงที่ 9 เป็นโรงเรียนที่มีการจัดห้องเรียนแบบจัดอันดับความสามารถ มีการสอบคัดเลือกนักเรียนเข้าเรียน เป็นโรงเรียนขนาดใหญ่ และขนาดกลาง ในกรุงเทพฯ โรงเรียนในกลุ่มแฝงที่ 8 เป็นโรงเรียนในสังกัดอุดมศึกษา และ สังกัด สพฐ. ที่มีขนาดใหญ่ โดยในกลุ่มแฝงที่ 8 มีทั้งโรงเรียนที่ใช้หลักสูตรปกติ และหลักสูตรทางเลือก และมีทั้งโรงเรียนขนาดกลาง ขนาดใหญ่ และ

ขนาดใหญ่พิเศษ กลุ่มแฝงที่ 4 เป็นกลุ่มที่มีความสามารถต่ำที่สุดเป็นโรงเรียนที่มีขนาดเล็กในต่างจังหวัด สังกัดกรมส่งเสริมการปกครองส่วนท้องถิ่น

5.2 อภิปรายผลการวิจัย

จากผลการวิเคราะห์ข้อมูลและสรุปผลการวิจัย ผู้วิจัยนำเสนอประเด็นอภิปรายตามวัตถุประสงค์การวิจัยดังนี้

ประเด็นที่ 1 อภิปรายในส่วนของการพัฒนาโมเดล

โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ เป็นโมเดลที่บูรณาการโมเดลพหุระดับ โมเดลการตอบสนองข้อสอบ และโมเดลการวิเคราะห์กลุ่มแฝง สำหรับการจำแนกกลุ่มความสามารถแฝงโดยใช้ข้อมูลแบบต่อเนื่องที่มีลักษณะข้อมูลเป็นพหุระดับ การพัฒนาโมเดลนี้เป็นการพัฒนาต่อยอดจากโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการผสมโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Tay, Diener, et al., 2011) และ แบบ 1 พารามิเตอร์ (Cho & Cohen, 2010) ที่ผ่านมามีงานวิจัยเกี่ยวกับ โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับเพียง 3 เรื่อง ได้แก่งานวิจัยของ Cho และ Cohen (2010) พัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการผสมโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ และนำโมเดลไปใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ งานวิจัยของTay และคณะ (2011) พัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการผสมโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ และนำโมเดลไปประยุกต์ใช้ตรวจสอบการรายงานอารมณ์ของบุคคลในองค์กร แบบข้ามระดับประเทศ และงานวิจัยของ De Jong และ Steenkamp (2010) เสนอโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่เป็นการผสมกับโมเดลพหุมิติ และนำไปใช้กับการศึกษาข้ามวัฒนธรรม

ดังนั้นงานวิจัยนี้จึงเป็นการวิจัยที่ใหม่โดยเป็นการพัฒนาโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งจะเหมาะสมกับการนำไปใช้กับบริบททางการศึกษา เนื่องจากแบบวัดในปัจจุบันโดยส่วนใหญ่เป็นแบบวัดชนิดเลือกตอบ ซึ่งการเดาส่งผลต่อโมเดลการวัดของแต่ละบุคคล ทั้งนี้ในการวิเคราะห์ข้อมูลตามโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับสามารถวิเคราะห์ด้วยโปรแกรมสำเร็จรูป Latent gold (Vermunt & Magidson, 2008) แต่โปรแกรมดังกล่าวรองรับโมเดลที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 1 และ 2 พารามิเตอร์เท่านั้น งานวิจัยนี้จึงไม่สามารถใช้โปรแกรมสำเร็จรูปในการวิเคราะห์ข้อมูลได้ ดังนั้นในการพัฒนาโมเดลเพื่อใช้ในงานวิจัยนี้ ผู้วิจัยศึกษาจากโมเดลสมการทางคณิตศาสตร์ แล้วสรุปเป็นขั้นตอนในการวิเคราะห์ข้อมูลทั้งสิ้น 5 ขั้นตอน และเขียนคำสั่งในโปรแกรม R ดังรายละเอียดในบทที่ 4

ดังนั้นผลการวิจัยนี้ทำให้ได้โมเดลการวิเคราะห์ และได้ชุดคำสั่งสำหรับการวิเคราะห์จำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

ประเด็นที่ 2 อภิปรายในส่วนของการนำโมเดลไปใช้กับข้อมูลจำลอง

จากผลการตรวจสอบความถูกต้องของการจำแนกกลุ่มความสามารถแฝงของข้อมูลจำลองโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น ซึ่งงานวิจัยนี้ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลแต่ละเงื่อนไขจำนวน 25 ครั้ง พบว่าสามารถจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนได้ถูกต้องตามข้อมูลที่จำลองทั้ง 2 สถานการณ์ ได้แก่ ผลการตอบข้อสอบ 6 ข้อ และ 15 ข้อ สำหรับผลการตรวจสอบความถูกต้องของการจำแนกกลุ่มความสามารถแฝงของผลการตอบข้อสอบ 6 ข้อ จำแนกได้ถูกต้องทั้งสิ้น 24 ครั้ง คิดเป็นร้อยละ 96 และสำหรับผลการตอบข้อสอบ 15 ข้อ จำแนกได้ถูกต้องทั้งสิ้น 25 ครั้ง คิดเป็นร้อยละ 100 แสดงให้เห็นว่าโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้นสามารถจำแนกได้ถูกต้องแม่นยำมากขึ้น เมื่อใช้กับข้อสอบที่มีจำนวนข้อมากขึ้น

ประเด็นที่ 3 อภิปรายในส่วนของการนำโมเดลไปใช้ในการจำแนกกลุ่มความสามารถแฝงของบุคคลจากข้อมูลเชิงประจักษ์

ในการจำแนกกลุ่มความสามารถแฝงของบุคคลจากข้อมูลเชิงประจักษ์ ในที่นี้ผู้วิจัยเลือกใช้ผลการตอบข้อสอบ O-NET วิชาคณิตศาสตร์ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เนื่องจากการทดสอบ O-NET เป็นการทดสอบมาตรฐานระดับชาติ และเป็นการทดสอบที่มีมาตรฐานในการสอบซึ่งนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ทุกคนจะเข้ารับการทดสอบ และรายวิชาคณิตศาสตร์เป็นรายวิชาที่มีความสำคัญ นักเรียนทุกคนจะต้องนำผลคะแนนจากรายวิชานี้ในการเข้าศึกษาต่อระดับอุดมศึกษา แต่อย่างไรก็ตาม ลักษณะข้อสอบเป็นแบบเลือกตอบจึงอาจมีความคลาดเคลื่อนในการวัดซึ่งเกิดจากการเดาข้อสอบ เนื่องจากบริบททางการศึกษาเป็นบริบทที่มีโครงสร้างแบบพหุระดับที่มีนักเรียนสอดแทรกอยู่ในโรงเรียน โรงเรียนสอดแทรกอยู่ในสังกัด ดังนั้นการเลือกใช้ผลการตอบข้อสอบนี้จึงเหมาะสมสำหรับการนำมาจำแนกกลุ่มความสามารถแฝงของบุคคลด้วยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้น เนื่องจากโมเดลนี้จะไม่ละเลยความคลาดเคลื่อนนี้ทำให้สามารถวัดได้ตรงตามสิ่งที่ต้องการ นอกจากนั้นงานวิจัยนี้ผู้วิจัยใช้การสุ่มตัวอย่างสำหรับวิเคราะห์ข้อมูล เนื่องจากไม่สามารถวิเคราะห์ข้อมูลกับประชากรได้ ด้วยข้อจำกัดทางด้านฮาร์ดแวร์ แต่อย่างไรก็ตาม การสุ่มตัวอย่างในที่นี้เป็นไปตามหลักความน่าจะเป็น ผู้วิจัยใช้การสุ่มตัวอย่างอย่างง่าย โดยกระบวนการสุ่มโรงเรียน ใช้การสุ่มแบบชั้นภูมิ (Stratified random sampling) และทุกสังกัดโรงเรียนมีการสุ่มโรงเรียนตามขนาดของโรงเรียน ดังนั้นกลุ่มตัวอย่างที่ศึกษาจึงเป็นตัวแทนที่ดีของประชากร

ผลการนำโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้นไปใช้กับข้อมูลเชิงประจักษ์ทำให้ทราบว่าการจำแนกกลุ่มความสามารถแฝงด้วยโมเดลการตอบสนองข้อสอบแบบผสมที่พัฒนาขึ้นทำให้ได้สารสนเทศที่มากกว่าการจำแนกกลุ่มแบบดั้งเดิม ดังนี้

สำหรับสารสนเทศจากการจำแนกกลุ่มความสามารถแฝงจากข้อมูลเชิงประจักษ์นี้ ได้สารสนเทศทั้งระดับโรงเรียนและระดับนักเรียน ผลการจำแนกกลุ่มความสามารถแฝงระดับโรงเรียนทำให้ทราบว่ามีโรงเรียนใดบ้างอยู่ในกลุ่มเดียวกัน และโรงเรียนที่อยู่ในกลุ่มเดียวกันมีลักษณะใดสอดคล้องกัน และโรงเรียนที่อยู่ในกลุ่มแฝงที่มีความสามารถสูงมีลักษณะอย่างไร กลุ่มแฝงที่มีความสามารถต่ำมีลักษณะอย่างไร และผลการจำแนกกลุ่มความสามารถแฝงระดับนักเรียน ทำให้ทราบว่านักเรียนที่มีลักษณะการตอบแบบใดอยู่ในกลุ่มแฝงเดียวกัน และกลุ่มแฝงนั้นมีลักษณะเด่นอย่างไร

จากผลการจำแนกกลุ่มความสามารถแฝงเมื่อเปรียบเทียบกับลักษณะโรงเรียนซึ่งเป็นตัวแปรที่สังเกตได้ ทำให้ทราบว่าควรเร่งพัฒนาการศึกษาทางคณิตศาสตร์ให้กับโรงเรียนในสังกัดกรมส่งเสริมการปกครองส่วนท้องถิ่นที่มีขนาดเล็กเป็นลำดับแรก และควรพัฒนาในสาระจำนวนและการดำเนินการ พีชคณิต การวัด โดยศึกษาแนวทางการพัฒนาความสามารถทางคณิตศาสตร์ หลักสูตรที่ใช้ และการจัดการศึกษาทางคณิตศาสตร์ จากโรงเรียนที่ถูกจำแนกอยู่ในกลุ่มความสามารถแฝงที่มีความสามารถสูงที่สุด ซึ่งมีจำนวน 2 โรงเรียน โดยทั้งสองโรงเรียนนี้เป็นโรงเรียนที่มีการจัดห้องเรียนแบบจัดอันดับความสามารถ และมีการสอบคัดเลือกนักเรียนเข้าเรียนในโรงเรียน นอกจากนี้อาจศึกษาแนวทางการพัฒนาความสามารถทางคณิตศาสตร์จากโรงเรียนที่ถูกจำแนกอยู่ในกลุ่มความสามารถแฝงที่มีความสามารถเป็นอันดับสอง ซึ่งในกลุ่มนี้จะเป็นโรงเรียนในสังกัดอุดมศึกษา และโรงเรียนในสังกัด สพฐ. ที่มีขนาดใหญ่ และโรงเรียนที่ใช้หลักสูตรทางเลือก นอกจากนี้ผลการศึกษานี้ทำให้ทราบว่าโรงเรียนที่ใช้หลักสูตรทางเลือกมีความสามารถทางด้านคณิตศาสตร์ไม่แตกต่างจากโรงเรียนที่ใช้หลักสูตรปกติ และพบว่าโรงเรียนขนาดใหญ่และขนาดใหญ่พิเศษมีนักเรียนที่มีความสามารถคณิตศาสตร์สูงกว่าโรงเรียนขนาดเล็กและปานกลาง

จะเห็นได้ว่าผลจากการจำแนกกลุ่มความสามารถแฝงนี้ทำให้ผู้บริหารของโรงเรียนรู้ว่าโรงเรียนของตนจำเป็นต้องพัฒนาด้านไหน สามารถกำหนดทิศทางการบริหารโรงเรียนเพื่อให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนที่สูงขึ้น และครูสามารถกำหนดแนวทางในการพัฒนานักเรียนต่อไป

5.3 ข้อเสนอแนะในการนำผลการวิจัยไปใช้

ข้อเสนอแนะในการนำผลการวิจัยไปใช้ สามารถสรุปเป็น 2 ประเด็น

ประเด็นที่ 1 โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้นเป็นประโยชน์กับหน่วยงานที่เกี่ยวข้องกับการจัดระบบการศึกษา และองค์กรทางการศึกษา เช่น สำนักงานการศึกษาขั้นพื้นฐาน สามารถนำโมเดลการวิเคราะห์หรือชุดคำสั่งในการวิเคราะห์ตามโมเดลนี้ ไปประยุกต์ใช้กับการจัดกลุ่มความสามารถของนักเรียน โดยทำการศึกษากับประชากรทั้งหมด เพื่อตัดสินใจว่าโรงเรียนใดบ้างที่มีความจำเป็นเร่งด่วนในการพัฒนา และควรให้ความสำคัญกับด้านใดเป็นอันดับแรก องค์กรทางการศึกษาอื่น ๆ เช่น สถาบันราชภัฏ สถาบันเทคโนโลยีราชมงคล หรือองค์กรระดับมหาวิทยาลัย สามารถนำโมเดลการวิเคราะห์หรือชุดคำสั่งในการวิเคราะห์ตามโมเดลนี้ไปประยุกต์ใช้กับการจัดกลุ่มคณะ หรือหลักสูตรต่างๆ ได้ สถาบันส่งเสริมการสอนวิทยาศาสตร์ และเทคโนโลยี หรือหน่วยงานที่เกี่ยวข้องสามารถประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้กับการทดสอบระดับนานาชาติ เช่น PISA, TIMSS และจำแนกกลุ่มบุคคลตามคุณลักษณะแฝงแบบข้ามระดับประเทศ เป็นประโยชน์ในการวินิจฉัยความสามารถที่แท้จริงของผู้สอบรายบุคคลแบบข้ามระดับประเทศ เพื่อออกแบบการพัฒนาคุณภาพการจัดการศึกษา และเป็นแนวทางในการพัฒนาการศึกษาระดับชาติให้ทัดเทียมกับนานาชาติต่อไป

นอกจากนี้สามารถประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ในการวินิจฉัยผู้เรียนโดยการเปรียบเทียบข้ามระดับกลุ่มหรือระดับโรงเรียนเพื่อให้เกิดผลที่ชัดเจนขึ้น โดยนำโมเดลนี้ใช้กับข้อสอบแบบเลือกตอบที่มีการสร้างข้อคำถาม และตัวลวงที่ดี ซึ่งผลการวิจัยนี้เป็นประโยชน์กับครู ทำให้ทราบว่าผู้เรียนมีการตอบผิดในลักษณะใด เป็นประโยชน์ในการวินิจฉัยผู้เรียนอย่างลึกซึ้ง เช่นเดียวกับงานวิจัยของ Bolt, Cohen & Wollack (2001) ที่ประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบผสม ในการวินิจฉัยผู้เรียนในวิชาภาษาอังกฤษ

ประเด็นที่ 2 ผลการจำแนกกลุ่มความสามารถแฝงกับข้อมูลเชิงประจักษ์นี้ หน่วยงานหรือองค์กรทางการศึกษา สามารถนำไปใช้ในการตัดสินใจกำหนดนโยบาย จัดสรรงบประมาณในการพัฒนาโรงเรียนต่อไป ผู้บริหารโรงเรียนสามารถนำไปใช้ในการกำหนดนโยบาย และทิศทางในการพัฒนาโรงเรียนให้มีความสามารถด้านคณิตศาสตร์สูงขึ้น ครูหรือนักการศึกษาที่สนใจสามารถนำผลการจำแนกกลุ่มความสามารถแฝงไปวินิจฉัยได้ว่านักเรียนควรจะพัฒนาด้านใดเพิ่มเติม โดยเปรียบเทียบความสามารถของผู้ตอบกับลักษณะข้อสอบรายข้อ จะทำให้ทราบว่านักเรียนในแต่ละกลุ่มแฝงไม่เข้าใจเนื้อหาเรื่องใด เพื่อเป็นแนวทางในการออกแบบการจัดการเรียนการสอนให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนที่สูงขึ้น

5.4 ข้อเสนอแนะในการทำวิจัยต่อไป

ข้อเสนอแนะในการทำวิจัย เพื่อขยายองค์ความรู้ในด้านการพัฒนาโมเดลสามารถสรุปได้ 3 ข้อ ดังนี้

1. ในการจำแนกกลุ่มความสามารถแฝงโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับที่พัฒนาขึ้นนี้ เป็นการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งเป็นการพัฒนาโมเดลต่อยอดจากโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ ที่มีการบูรณาการกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับแบบ 1 และ 2 พารามิเตอร์ ซึ่งผลการจำแนกกลุ่มความสามารถแฝงโดยใช้โมเดลที่พัฒนาขึ้นนี้ อาจเหมือนหรือแตกต่างกับการใช้โมเดลแบบ 1 และ 2 พารามิเตอร์

ดังนั้นจึงควรมีการศึกษาเปรียบเทียบผลการจำแนกกลุ่มความสามารถแฝงโดยโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับแบบ 1 พารามิเตอร์ แบบ 2 พารามิเตอร์ และ แบบ 3 พารามิเตอร์ ต่อไป

2. สำหรับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับในงานวิจัยนี้ มีการบูรณาการโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ซึ่งเหมาะสมกับการประยุกต์ใช้กับข้อสอบแบบเลือกตอบที่มีการให้คะแนนแบบสองค่า (dichotomous scoring) ซึ่งมีอิทธิพลจากการเดาข้อสอบ แต่ทั้งนี้ ข้อสอบในปัจจุบัน และแบบวัดทางจิตวิทยา มีลักษณะการให้คะแนนแบบหลายค่า (polytomous scoring) เช่น แบบวัดที่เป็น rating scale หรือแบบสอบที่เป็น multiple weight choice

ดังนั้นการขยายความรู้เกี่ยวกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับโดยพัฒนาโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ ในการทดสอบ หรือการประเมินโดยใช้แบบสอบหรือแบบวัดที่มีการให้คะแนนแบบหลายค่าได้ โดยการรวม Latent Class Model Multilevel Model และ IRT Model ที่มีการให้คะแนนแบบหลายค่า (Polytomous) จะเป็นประโยชน์อย่างยิ่งในการจำแนกกลุ่มของแต่ละบุคคลว่าบุคคลใดมีความสามารถ หรือคุณลักษณะแฝงอยู่ในระดับใดเมื่อเปรียบเทียบกับข้อมูลในระดับประเทศ หรือเปรียบเทียบระดับบุคคล โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับนี้ ในขอบเขตของข้อสอบที่มีการให้คะแนนแบบหลายค่า

3. จากผลการประเมินการศึกษาระดับนานาชาติ และระดับชาติ เช่น PISA, TIMSS, PAT, O-NET, NT มีการเก็บข้อมูลที่ต่อเนื่องกันในทุก 3 ปี หรือทุกปี และจากที่ประเทศไทยได้นำผลการประเมินมาเป็นพื้นฐานในการออกแบบการพัฒนาคุณภาพการศึกษามาโดยตลอด แต่ยังไม่มีการศึกษาเปรียบเทียบว่าผลจากการพัฒนานั้นมีแนวโน้มเป็นอย่างไร นักเรียนมีการพัฒนาขึ้นหรือไม่ เมื่อจำแนกนักเรียนตามคุณลักษณะแฝงแบบพหุระดับ ทั้งนี้การศึกษาแนวโน้มว่าในการพัฒนาคุณภาพ

การศึกษานั้นดีขึ้นหรือไม่ และถ้าไม่ดีขึ้นนั้นมีสาเหตุมาจากอะไร และนักเรียนกลุ่มที่สามารถพัฒนาได้นั้นมีคุณลักษณะแฝงเป็นอย่างไร จะเป็นประโยชน์ในการวินิจฉัยได้ตรงประเด็นมากขึ้น

ดังนั้นการขยายความรู้เกี่ยวกับโมเดลการตอบสนองข้อสอบแบบผสมพหุระดับโดยพัฒนาโมเดลทฤษฎีการตอบสนองข้อสอบแบบผสมพหุระดับ ในการทดสอบ โดยการรวม Latent Class Model Multilevel Model IRT Model และ Growth model หรือ Latent Transition model เพื่อการจำแนกกลุ่มของแต่ละบุคคลตามคุณลักษณะแฝงแบบข้ามระดับกับข้อมูลระยะยาวซึ่งจะเป็นประโยชน์ทางการศึกษาต่อไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายการอ้างอิง

ภาษาไทย

เกศมณี พยัคฆ์. (2543). *การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบทดสอบความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกันที่วิเคราะห์ด้วยราส์ชโมเดล*. วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาวิชาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

คณาจารย์ภาควิชาคณิตศาสตร์คณะวิทยาศาสตร์จุฬาลงกรณ์มหาวิทยาลัย. (2551). *ความน่าจะเป็นและสถิติ*. กรุงเทพมหานคร: โรงพิมพ์ห้างหุ้นส่วนจำกัด พิทักษ์การพิมพ์.

ชนะศึก นิชานนท์. (2553). *ประสิทธิภาพของการประมาณค่าพารามิเตอร์แบบเบส์โดยใช้การสรุปอ้างอิงความน่าเชื่อถือของโมเดลการตอบสนองข้อสอบ*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

รัตนา ศรีเหรัญ. (2539). *การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติค 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัม ไลค์ลิฮูด วิธีฮิวริสติคและวิธีของเบส์เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกันด้วยเทคนิคมอนติคาร์โล*. วิทยานิพนธ์ปริญาดุษฎีบัณฑิต. สาขาวิชาการทดสอบและวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

วิชุดา บัวคง. (2532). *การเปรียบเทียบประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ของแบบจำลอง โลจิสติก 3 พารามิเตอร์ระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีฮิวริสติก และวิธีของเบส์ในแบบสอบวัดผลสัมฤทธิ์*. วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2550). *การวิเคราะห์พหุระดับ*. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่*. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ศิริรัตน์ วงศ์ประกรณ์กุล. (2539). การจำลองโดยใช้เทคนิคมอนติคาร์โล. *วารสารวิทยาศาสตร์มหาวิทยาลัยขอนแก่น, 24*(4), 240-246.

ศึกษาธิการ, กระทรวง. สำนักงานการศึกษาขั้นพื้นฐาน. (2552). *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551.* กรุงเทพมหานคร: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

สถาบันทดสอบทางการศึกษา. (2555). รูปแบบข้อสอบ O-NET มัธยมศึกษาปีที่ 6 แยกตามรูปแบบของข้อสอบ ประจำปีการศึกษา 2555. [ออนไลน์]. Retrieved 25 กุมภาพันธ์ 2556, 2556, from http://www.niets.or.th/index.php/system_niest/index/3

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2554). *ผลการประเมิน* PISA 2009 *คณิตศาสตร์ การอ่าน และวิทยาศาสตร์ (ฉบับสมบูรณ์)*. กรุงเทพมหานคร: อรุณการพิมพ์.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2557). *ผลการประเมิน* PISA 2012 *คณิตศาสตร์ การอ่าน และวิทยาศาสตร์ (ฉบับสมบูรณ์)*. กรุงเทพมหานคร: อรุณการพิมพ์.

ภาษาอังกฤษ

Agresti, Alan, Booth, James G., Hobert, James P., & Caffo, Brian. (2000). Random-Effects Modeling of Categorical Response Data. *Sociological Methodology, 30*(1), 27-80.

Aitkin, Murray. (1999). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics, 55*(1), 117-128.

Baum, Leonard E., Petrie, Ted, Soules, George, & Weiss, Norman. (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics, 41*(1), 164-171.

Bergman, Lars R, & Magnusson, David. (1997). A person-oriented approach in research on developmental psychopathology. *Development and psychopathology, 9*(02), 291-319.

Bilir, Mustafa Kuzey. (2011). *Mixture Item Response Theory-MIMIC Model: Simultaneous Estimation of Differential Item Functioning for Manifest Groups and Latent Classes*. Charleston SC: Proquest, Umi Dissertation Publishing.

Birnbaum, Allan. (1968). Some latent trait models and their use in inferring an examinee's ability. *Statistical theories of mental test scores*.

Bock, R Darrell, & Aitkin, Murray. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*(4), 443-459.

Bolt, Daniel M, Cohen, Allan S, & Wollack, James A. (2001). A mixture item response model for multiple-choice data. *Journal of Educational and Behavioral Statistics, 26*(4), 381-409.

Bolt, Daniel M, Cohen, Allan S, & Wollack, James A. (2002). Item parameter estimation under conditions of test speededness: Application of a mixture Rasch model with ordinal constraints. *Journal of Educational Measurement, 39*(4), 331-348.

Bryk, A, & Stephen, W. (2002). *Hierarchical linear models for social and behavioral research: applications and data analysis methods*. Thousand Oaks, CA: Sage Publication, Inc.

Carter, Nathan T, Dalal, Dev K, Lake, Christopher J, Lin, Bing C, & Zickar, Michael J. (2011). Using mixed-model item response theory to analyze organizational survey responses: An illustration using the Job Descriptive Index. *Organizational Research Methods, 14*(1), 116-146.

Cho, Sun-Joo, & Cohen, Allan S. (2010). A multilevel mixture IRT model with an application to DIF. *Journal of Educational and Behavioral Statistics, 35*(3), 336-370.

Cho, Sun-Joo, Cohen, Allan S, Kim, Seock-Ho, & Bottge, Brian. (2010). Latent transition analysis with a mixture item response theory measurement model. *Applied Psychological Measurement, 34*(7), 483-504.

Cho, Sun-Joo, & Cohen, AS. (2007). *A multilevel mixture IRT model for DIF analysis.* Paper presented at the International Meeting of the Psychometric Society: The 72nd annual meeting of the Psychometric Society, Tokyo, Japan.

Chris Fraley, Adrian E. Raftery, Luca Scrucca, Thomas Brendan Murphy, & Fop, Michael. (2015). Package 'mclust'. Retrieved July 24, 2016, from R project: https://cran.r-project.org/web/packages/mclust/mclust.pdf.

Clogg, Clifford C. (1988). Latent class models for measuring *Latent trait and latent class models* (pp. 173-205). Heidelberg: Springer.

Cohen, AS, Cho, SJ, & Kim, SH. (2005). *A mixture testlet model for educational tests.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Collins, Linda M, & Lanza, Stephanie T. (2010). Latent class and latent transition analysis. New Jersey: Wiley & Sons, Inc.

Congdon, Peter. (2005). *Bayesian models for categorical data*. New York: John Wiley & Sons.

Cronbach, Lee J, & Webb, Noreen. (1975). Between-class and within-class effects in a reported aptitude* treatment interaction: Reanalysis of a study by GL Anderson. *Journal of Educational Psychology, 67*(6), 717-724.

De Jong, Martijn G, & Steenkamp, Jan-Benedict EM. (2010). Finite mixture multilevel multidimensional ordinal IRT models for large scale cross-cultural research. *Psychometrika, 75*(1), 3-32.

Dempster, Arthur P, Laird, Nan M, & Rubin, Donald B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.

Egberink, Iris JL, Meijer, Rob R, & Veldkamp, Bernard P. (2010). Conscientiousness in the workplace: Applying mixture IRT to investigate scalability and predictive validity. *Journal of Research in Personality, 44*(2), 232-244.

Eid, Michael, & Rauber, Michael. (2000). Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment, 16*(1), 20.

Finch, W Holmes, & Pierson, Eric E. (2011). A mixture IRT analysis of risky youth behavior. *Frontiers in psychology, 2*.

Formann, Anton K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association, 87*(418), 476-486.

Fox, J-P. (2005). Multilevel IRT using dichotomous and polytomous response data. *British Journal of Mathematical and Statistical Psychology, 58*(1), 145-172.

Fox, Jean-Paul. (2007). Multilevel IRT modeling in practice with the package mlirt. *Journal of Statistical Software, 20*(5), 1-16.

Fox, Jean-Paul, & Glas, Cees AW. (2001). Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika, 66*(2), 271-288.

Fraley, Chris, & Raftery, Adrian E. (2015). MCLUST version 5: an R package for Normal Mixture Modelling for Model-Based Clustering, Classification, and Density Estimation: DTIC Document.

Frühwirth-Schnatter, Sylvia. (2006). *Finite mixture and Markov switching models*. New York: Springer Science & Business Media.

Geisser, Seymour, & Eddy, William F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association, 74*(365), 153-160.

Gelfand, Alan E, & Dey, Dipak K. (1994). Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 501-514.

Gelman, A, Carlin, JB, Stern, HS, & Rubin, DB. (1996). Bayesian Data. *Analysis. London: Chapman and Hall*.

Goodman, Leo A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika, 61*(2), 215-231.

Hambleton, Ronald K. (1991). *Fundamentals of item response theory* (Vol. 2). Thousand Oaks, CA: Sage Publications, Inc.

Hambleton, Ronald K, & Cook, Linda L. (1977). Latent trait models and their use in the analysis of educational test data1, 2, 3. *Journal of educational measurement, 14*(2), 75-96.

Hambleton, Ronald K, & Swaminathan, Hariharan. (1985). *Item response theory: Principles and applications* (Vol. 7). New York: Springer Science & Business Media.

Harwell, Michael R, & Janosky, Janine E. (1991). An empirical study of the effects of small datasets and varying prior variances on item parameter estimation in BILOG. *Applied Psychological Measurement, 15*(3), 279-291.

Harwell, Michael, Stone, Clement A, Hsu, Tse-Chi, & Kirisci, Levent. (1996). Monte Carlo studies in item response theory. *Applied psychological measurement, 20*(2), 101-125.

Haughton, Dominique, Legrand, Pascal, & Woolford, Sam. (2012). Review of three latent class cluster analysis packages: Latent Gold, poLCA, and MCLUST. *The American Statistician, 63*(1), 81-91.

Heinen, Ton. (1996). *Latent class and discrete latent trait models: Similarities and differences*. Thousand Oaks, CA: Sage Publications, Inc.

Hernández, Ana, Drasgow, Fritz, & González-Romá, Vicente. (2004). Investigating the functioning of a middle category by means of a mixed-measurement model. *Journal of Applied Psychology, 89*(4), 687.

Jordan, Michael I, & Jacobs, Robert A. (1992). *Hierarchies of adaptive experts.* Paper presented at the Advances in neural information processing systems.

Juang, Biing Hwang, & Rabiner, Laurence R. (1991). Hidden Markov models for speech recognition. *Technometrics, 33*(3), 251-272.

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

Kozlowski, Steve WJ, & Klein, Katherine J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.

Laird, Nan. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association, 73*(364), 805-811.

Lanza, Stephanie T, Collins, Linda M, Lemmon, David R, & Schafer, Joseph L. (2007). PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling, 14*(4), 671-694.

Lazarsfeld, Paul Felix, Henry, Neil W, & Anderson, Theodore Wilbur. (1968). *Latent structure analysis*. Boston: Houghton Mifflin Boston.

Lesaffre, Emmanuel, & Spiessens, Bart. (2001). On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 50*(3), 325-335.

Li, Feiming, Cohen, Allan S, Kim, Seock-Ho, & Cho, Sun-Joo. (2009). Model Selection Methods for Mixture Dichotomous IRT Models. *Applied Psychological Measurement, 33*(5), 353-373.

Lin, Ting Hsiang, & Dayton, C Mitchell. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics, 22*(3), 249-264.

Lord, Frederic M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement, 23*(2), 157-162.

Lord, Frederic M. (2009). *Applications of Item Response Theory to Practical Testing Problems*. New York: Routledge Taylor and Francis Group.

Lubke, Gitta H, & Muthén, Bengt. (2005). Investigating population heterogeneity with factor mixture models. *Psychological methods, 10*(1), 21.

Magidson, Jay, & Vermunt, Jeroen K. (2001). Latent Class Factor and Cluster Models, Bi-Plots, and Related Graphical Displays. *Sociological methodology, 31*(1), 223-264.

Magidson, Jay, & Vermunt, Jeroen K. (2004). Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, 175-198.

Maij-de Meij, Annette M, Kelderman, Henk, & van der Flier, Henk. (2008). Fitting a Mixture Item Response Theory Model to Personality Questionnaire Data: Characterizing Latent Classes and Investigating Possibilities for Improving Prediction. *Applied Psychological Measurement, 32*(8), 611-631.

Maij-de Meij, Annette M, Kelderman, Henk, & van der Flier, Henk. (2010). Improvement in detection of differential item functioning using a mixture item response theory model. *Multivariate Behavioral Research, 45*(6), 975-999.

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

McCutcheon, Allan L. (1987). *Latent class analysis*. Thousand Oaks, CA: Sage Publications, Inc.

McLachlan, Geoffrey, & Peel, David. (2004). *Finite mixture models*. New York: John Wiley & Sons.

Meehl, Paul E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60*(1), 117-174.

Mislevy, Robert, & Huang, Chun-Wei. (2007). *Measurement models as narrative structures*. Heidelberg: Springer.

Mislevy, Robert J, & Verhelst, Norman. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika, 55*(2), 195-215.

Chulalongkorn University

Muraki, Eiji. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement, 14*(1), 59-71.

Muraki, Eiji. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement, 16*(2), 159-176.

Muthén, Bengt. (2004). Latent variable analysis. *The Sage handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage Publications*, 345-368.

Muthen, Bengt, & Asparouhov, Tihomir. (2006). Item response mixture modeling: Application to tobacco dependence criteria. *Addictive behaviors, 31*(6), 1050-1066.

Muthén, Linda K, & Muthén, Bengt O. (2008). Mplus (Version 5.1). *Los Angeles, CA: Muthén & Muthén*.

Newman, DA, Hanges, Paul J, Duan, L, & Ramesh, A. (2008). A network model of organizational climate: Friendship, clusters, subgroup agreement, and climate schemas. *The people make the place: Dynamic linkages between individuals and organizations*, 101-126.

Partchev, Ivailo, Partchev, Maintainer Ivailo, & Suggests, MASS. (2015). Package 'irtoys'. Retrieved July 24, 2016, from R project: https://cran.r-project.org/web/packages/irtoys/irtoys.pdf.

Pearl, Judea, & Russell, Stuart. (1998). *Bayesian networks*. Los Angeles: Computer Science Department, University of California.

Rasch, Georg. (1960). Probabilistic models for some intelligence and achievement tests. *Copenhagen: Danish Institute for Educational Research*.

Raudenbush, Stephen, & Bryk, Anthony S. (1986). A hierarchical model for studying school effects. *Sociology of education*, 1-17.

Raudenbush, Stephen W, & Bryk, Anthony S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage Publication, Inc.

Reise, Steven P, & Gomel, Jessica N. (1995). Modeling qualitative variation within latent trait dimensions: Application of mixed-measurement to personality assessment. *Multivariate Behavioral Research, 30*(3), 341-358.

Rijmen, Frank, & De Boeck, Paul. (2003). A latent class model for individual differences in the interpretation of conditionals. *Psychological Research, 67*(3), 219-231.

Ripley, Brian D. (1988). Uses and abuses of statistical simulation. *Mathematical Programming, 42*(1-3), 53-68.

Rizopoulos, Dimitris. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of statistical software, 17*(5), 1-25.

Rizopoulos, Dimitris. (2013). Package 'ltm'. Retrieved July 24, 2016, from R project: https://cran.r-project.org/web/packages/ltm/ltm.pdf.

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

Rost, Jürgen. (1991). A logistic mixture distribution model for polychotomous item responses. *British Journal of Mathematical and Statistical Psychology, 44*(1), 75-92.

Rost, Jürgen. (1997). Logistic mixture models *Handbook of modern item response theory* (pp. 449-463). Heidelberg: Springer.

Rothenberg, Thomas J. (1971). Identification in parametric models. *Econometrica: Journal of the Econometric Society*, 577-591.

Skrondal, Anders, & Rabe-Hesketh, Sophia. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. London: Crc Press.

Smit, Arnold, Kelderman, Henk, & van der Flier, Henk. (2000). The mixed Birnbaum model: Estimation using collateral information. *Methods of Psychological Research Online, 5*(4), 31-43.

Spiegelhalter, David J, Best, Nicola G, Carlin, Bradley P, & Van der Linde, A. (1998). Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models: Research Report, 98-009.

Stephens, Matthew. (2000). Dealing with label switching in mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 62*(4), 795-809.

Stroud, Arthur H, & Secrest, Don. (1966). *Gaussian quadrature formulas* (Vol. 39). Englewood Cliffs, NJ: Prentice-Hall

Swaminathan, Hariharan, & Gifford, Janice A. (1982). Bayesian estimation in the Rasch model. *Journal of Educational and Behavioral Statistics, 7*(3), 175-191.

Swaminathan, Hariharan, & Gifford, Janice A. (1985). Bayesian estimation in the two-parameter logistic model. *Psychometrika, 50*(3), 349-364.

Swaminathan, Hariharan, & Gifford, Janice A. (1986). Bayesian estimation in the three-parameter logistic model. *Psychometrika, 51*(4), 589-601.

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

Tay, Louis, Newman, Daniel A, & Vermunt, Jeroen K. (2011). Using mixed-measurement item response theory with covariates (MM-IRT-C) to ascertain observed and unobserved measurement equivalence. *Organizational Research Methods, 14*(1), 147-176.

Thissen, D, Chen, WH, & Bock, RD. (2003). Multilog (version 7.0)[computer program]. *Mooresville, IN: Scientific Software*.

Thissen, David, & Steinberg, Lynne. (1986). A taxonomy of item response models. *Psychometrika, 51*(4), 567-577.

Van Nijlen, Daniël, & Janssen, Rianne. (2008). *Spelling in primary school: Mixture IRT-models as a diagnostic tool*. Paper presented at the EARLI Jure Conference Edition: 11, Germany.

Vandenberg, Robert J, & Lance, Charles E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods, 3*(1), 4-70.

Vermunt, Jeroen K. (2004). An EM algorithm for the estimation of parametric and nonparametric hierarchical nonlinear models. *Statistica Neerlandica, 58*(2), 220-233.

Vermunt, Jeroen K. (2007a). A hierarchical mixture model for clustering three-way data sets. *Computational Statistics & Data Analysis, 51*(11), 5368-5376.

Vermunt, Jeroen K. (2007b). Multilevel mixture item response theory models: an application in education testing. *Proceedings of the 56th session of the International Statistical Institute. Lisbon, Portugal, 2228*.

Vermunt, Jeroen K. (2008a). Latent class and finite mixture models for multilevel data sets. *Statistical Methods in Medical Research, 17*(1), 33-51.

Vermunt, Jeroen K. (2008b). Multilevel latent variable modeling: An application in education testing. *Austrian Journal of Statistics, 37*(3-4), 285-299.

Vermunt, Jeroen K, & Magidson, Jay. (2002). Latent class cluster analysis. *Applied latent class analysis, 11*, 89-106.

Vermunt, Jeroen K, & Magidson, Jay. (2003). Latent class models for classification. *Computational Statistics & Data Analysis, 41*(3), 531-537.

Vermunt, Jeroen K, & Magidson, Jay. (2005). Technical Guide for Latent GOLD 4.0: Basic and Advanced. Belmont, MA: Statistical Innovations.

Vermunt, Jeroen K, & Magidson, Jay. (2008). *LG-syntax User's Guide: Manual for Latent GOLD 4.5 syntax module*. Belmont, MA: Statistical Innovations Inc.

Vermunt, Jeroen K, & Van Dijk, Liesbet. (2001). A nonparametric random-coefficients approach: The latent class regression model. *Multilevel Modelling Newsletter, 13*(2), 6-13.

Von Davier, Matthias, & Yamamoto, Kentaro. (2004). Partially observed mixtures of IRT models: An extension of the generalized partial-credit model. *Applied Psychological Measurement, 28*(6), 389-406.

Waller, Niels G, & Meehl, Paul E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA: Sage Publications, Inc.

Wilson, Mark. (1989). Saltus: A psychometric model of discontinuity in cognitive development. *Psychological Bulletin, 105*(2), 276.

Wright, BD, & Panchapakesan, Nargis. (1969). A procedure for sample-free item analysis. *Educational and Psychological measurement, 29*(1), 23-48.

Yamamoto, Kentaro, & Everson, Howard. (1997). Modeling the effects of test length and test time on parameter estimation using the HYBRID model. *Applications of latent trait and latent class models in the social sciences*, 89-98.

Zhang, Nevin L. (2004). Hierarchical latent class models for cluster analysis. *The Journal of Machine Learning Research, 5*, 697-723.

Zickar, Michael J, Gibby, Robert E, & Robie, Chet. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods, 7*(2), 168-190.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก

ภาคผนวก ก คำสั่งที่ใช้ในการศึกษาด้วยโปรแกรม R

ภาคผนวก ข รายชื่อโรงเรียนในแต่ละกลุ่มแฝงที่จำแนกโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ก
คำสั่งที่ใช้ในการศึกษาด้วยโปรแกรม R

1. คำสั่ง สำหรับนำเข้า Library
2. คำสั่ง สำหรับการจำลองข้อมูล
3. คำสั่ง สำหรับการประมาณค่าพารามิเตอร์
4. คำสั่ง สำหรับการประมาณค่าความสามารถของแต่ละรูปแบบการตอบ
5. คำสั่ง สำหรับการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก
6. คำสั่ง สำหรับการแบ่งกลุ่มระดับบุคคล g
7. คำสั่ง สำหรับคำนวณหาความสามารถของโรงเรียน
8. คำสั่ง สำหรับการจำแนกกลุ่มระดับโรงเรียน
9. คำสั่ง สำหรับคำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1. คำสั่ง สำหรับนำเข้า Library

```
rm(list=ls())


library(foreign)
library(ltm)
library(matrixcalc)
```

2. คำสั่ง สำหรับการจำลองข้อมูล (ตัวอย่างคำสั่งสำหรับจำลองผลการตอบข้อสอบจำนวน 6 ข้อ)

```
##############################################
### simulation data sim(a,b,c , ability) ###
##############################################

library(irtoys)

All_school <- NULL

################
### Group 1 ###
################

pa1 <- cbind(rbind(2,2,1,1,1,1), rbind(-1.5,-1.5,0,0,1.5,1.5),
rbind(0.1,0.1,0.2,0.2,0.25,0.25))
pa1

All_school[[1]] <- sim(ip=pa1, x = rnorm(120,mean = 0,sd = 0.5))
colnames(All_school[[1]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[2]] <- sim(ip=pa1, x = rnorm(400,mean = 0,sd = 0.5))
colnames(All_school[[2]]) <- c("item1","item2","item3","item4","item5","item6")
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
All_school[[3]] <- sim(ip=pa1, x = rnorm(800,mean = 0,sd = 0.5))
colnames(All_school[[3]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[4]] <- sim(ip=pa1, x = rnorm(2000,mean = 0,sd = 0.5))
colnames(All_school[[4]]) <- c("item1","item2","item3","item4","item5","item6")

###############
### Group 2 ###
###############

pa2 <- cbind(rbind(1,1,1,1,2,2), rbind(0,0,1.5,1.5,-1.5,-1.5),
rbind(0.2,0.2,0.25,0.25,0.1,0.1))
pa2



All_school[[5]] <- sim(ip=pa2, x = rnorm(120,mean = 0,sd = 0.5))
colnames(All_school[[5]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[6]] <- sim(ip=pa2, x = rnorm(400,mean = 0,sd = 0.5))
colnames(All_school[[6]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[7]] <- sim(ip=pa2, x = rnorm(800,mean = 0,sd = 0.5))
colnames(All_school[[7]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[8]] <- sim(ip=pa2, x = rnorm(2000,mean = 0,sd = 0.5))
colnames(All_school[[8]]) <- c("item1","item2","item3","item4","item5","item6")

###############
### Group 3 ###
###############
```

```
pa3 <- cbind(rbind(1,1,2,2,1,1), rbind(1.5,1.5,-1.5,-1.5,0,0),
rbind(0.25,0.25,0.1,0.1,0.2,0.2))
pa3



All_school[[9]] <- sim(ip=pa3, x = rnorm(360,mean = 0,sd = 0.5))
colnames(All_school[[9]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[10]] <- sim(ip=pa3, x = rnorm(1200,mean = 0,sd = 0.5))
colnames(All_school[[10]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[11]] <- sim(ip=pa3, x = rnorm(2400,mean = 0,sd = 0.5))
colnames(All_school[[11]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[12]] <- sim(ip=pa3, x = rnorm(6000,mean = 0,sd = 0.5))
colnames(All_school[[12]]) <- c("item1","item2","item3","item4","item5","item6")



###############
### Group 4 ###
###############

pa4 <- cbind(rbind(1,1,1,1,1,1), rbind(0.25,0.25,-0.25,-0.25,0,0),
rbind(0.2,0.2,0.2,0.2,0.2,0.2))
pa4

All_school[[13]] <- sim(ip=pa4, x = rnorm(150,mean = 0,sd = 0.5))
colnames(All_school[[13]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[14]] <- sim(ip=pa4, x = rnorm(500,mean = 0,sd = 0.5))
colnames(All_school[[14]]) <- c("item1","item2","item3","item4","item5","item6")
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
All_school[[15]] <- sim(ip=pa4, x = rnorm(1000,mean = 0,sd = 0.5))
colnames(All_school[[15]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[16]] <- sim(ip=pa4, x = rnorm(2500,mean = 0,sd = 0.5))
colnames(All_school[[16]]) <- c("item1","item2","item3","item4","item5","item6")

###############
### Group 5 ###
###############

pa5 <- cbind(rbind(2,2,2,2,2,2), rbind(-2.5,-2,-2.5,-2,-2.5,-2), rbind(0.1,0.1,0.1,0.1,0.1,0.1))
pa5



All_school[[17]] <- sim(ip=pa5, x = rnorm(120,mean = 0,sd = 0.5))
colnames(All_school[[17]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[18]] <- sim(ip=pa5, x = rnorm(400,mean = 0,sd = 0.5))
colnames(All_school[[18]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[19]] <- sim(ip=pa5, x = rnorm(800,mean = 0,sd = 0.5))
colnames(All_school[[19]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[20]] <- sim(ip=pa5, x = rnorm(2000,mean = 0,sd = 0.5))
colnames(All_school[[20]]) <- c("item1","item2","item3","item4","item5","item6")

###############
### Group 6 ###
###############
```

```
pa6 <- cbind(rbind(1,1,1,1,1,1), rbind(2,2.5,2,2.5,2,2.5),
rbind(0.25,0.25,0.25,0.25,0.25,0.25))
pa6



All_school[[21]] <- sim(ip=pa6, x = rnorm(300,mean = 0,sd = 0.5))
colnames(All_school[[21]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[22]] <- sim(ip=pa6, x = rnorm(1000,mean = 0,sd = 0.5))
colnames(All_school[[22]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[23]] <- sim(ip=pa6, x = rnorm(2000,mean = 0,sd = 0.5))
colnames(All_school[[23]]) <- c("item1","item2","item3","item4","item5","item6")

All_school[[24]] <- sim(ip=pa6, x = rnorm(5000,mean = 0,sd = 0.5))
colnames(All_school[[24]]) <- c("item1","item2","item3","item4","item5","item6")



alldata <- rbind(       All_school[[1]],All_school[[2]],All_school[[3]],All_school[[4]],
                        All_school[[5]],All_school[[6]],All_school[[7]],All_school[[8]],
                        All_school[[9]],All_school[[10]],All_school[[11]],All_school[[12]],

        All_school[[13]],All_school[[14]],All_school[[15]],All_school[[16]],

        All_school[[17]],All_school[[18]],All_school[[19]],All_school[[20]],

        All_school[[21]],All_school[[22]],All_school[[23]],All_school[[24]])
detach("package:irtoys",unload=TRUE)

num_of_item <- ncol(alldata)
num_of_item
```

```
##### items mean ####

items_mean <- NULL

for(n in 1:num_of_item)
{

items_mean[[n]] <- mean(alldata[,n])

}
items_mean
```

3. คำสั่งสำหรับการประมาณค่าพารามิเตอร์

```
####################################
### Estimate 3pl IRT Parameters ###
####################################

test_model <- NULL

max_iteration <- 10
aaa = matrix(0,max_iteration,3)

for(n in 1:max_iteration)

{
H <- FALSE #is.positive.definite(temp$hessian)

while( H != TRUE )
{
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
        temp <- tpm(alldata, max.guessing = 0.25, start.val = "random",
control=list(iter.qn=100000))
        H <- is.positive.definite(temp$hessian)
        print(H)
        flush.console()
}

print(n)
flush.console()

aaa[n,1] <- n
aaa[n,2] <- temp$log.Lik
aaa[n,3] <- is.positive.definite(temp$hessian)

test_model[[n]] <- temp
}
aaa
max_loglik <- which.max(aaa[,2])

fit_mydata <-test_model[[max_loglik]] ### Maximum log likelihood
```

4. คำสั่ง สำหรับการประมาณค่าความสามารถของแต่ละรูปแบบการตอบ

```
model <- factor.scores(fit_mydata, method ="EB")
score <- model$score.dat
score
```

5. คำสั่ง สำหรับการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขในการตอบข้อสอบถูก

```
prob <- fitted(fit_mydata, type = c("conditional-probabilities"))
prob
model$coef
```

6. คำสั่ง สำหรับการจำแนกกลุ่มระดับบุคคล g

```
###################################
### LCA for Individual level k ###
###################################

library(mclust)

mod1 <- Mclust(prob,G=1:20, modelNames= mclust.options("emModelNames"))

AA <- summary(mod1, parameters =TRUE, classification = TRUE)
AA
ind_class <- cbind(AA$classification)
colnames(ind_class)<-c("class")
ind_class
BB <- cbind(score[,1:num_of_item],ind_class)

num_of_class <- max(AA$classification)
CC <- cbind(1:num_of_class,t(AA$mean))
colnames(CC) <- c("class","p1","p2","p3","p4","p5","p6") ## Input number of p as
number of item ##
ind_info <- merge(BB,CC,by = c("class"))
```

7. คำสั่ง สำหรับคำนวณหาความสามารถของโรงเรียน

```
###################################
### Calculate School Ability ###
###################################

num_of_school <- length(All_school)
School_ability = matrix(0,num_of_school,num_of_class)
School_ability
```

Chulalongkorn University

```
School_info <- NULL

for(d in 1:num_of_school)
{
	temp <- All_school[[d]]
	temp_num <-nrow(temp)

	temp <- merge(temp,ind_info,by =
c("item1","item2","item3","item4","item5","item6"))  ## Input number of item ##
	temp
	DD <- data.frame(table(temp[num_of_item+1]))
	colnames(DD) <- c("class","freq")
	School <- merge(CC,DD,by="class",all="true")
	School[is.na(School)] <- 0
	freq_location <- length(School)
	School_info[[d]] <- School

	for(e in 1:num_of_class)
	{
		School_ability[d,e] <- School[e,freq_location]/temp_num
	}
}
School_ability
#School_info
```

จุฬาลงกรณ์มหาวิทยาลัย

8. คำสั่ง สำหรับการจำแนกกลุ่มระดับโรงเรียน

```
#################################################################
### LCA for School level g ###
################################
mod2 <- Mclust(School_ability, modelNames= c("VII","VEI","VVI"))


EE <- summary(mod2, parameters =TRUE, classification = TRUE)
EE
t(EE$mean)
EE$classification
```

9. คำสั่ง สำหรับคำนวณหาความน่าจะเป็นในการตอบข้อสอบถูกของแต่ละโรงเรียน

```
################################
### Calculate Probability ###
################################
num_of_school_class <- max(EE$classification)
School_class_ability <- matrix(0,num_of_school_class,num_of_item)


for(m in 1:num_of_school_class)
{
        for(n in 1:num_of_item)
        {
                School_class_ability[m,n] <- sum(CC[,n+1]*EE$mean[,m])
        }
}
School_class_ability
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ข
รายชื่อโรงเรียนในแต่ละกลุ่มแฝง
ที่จำแนกโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายชื่อโรงเรียนในแต่ละกลุ่มแฝงที่จำแนกโดยใช้โมเดลการตอบสนองข้อสอบแบบผสมพหุระดับ

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
1	เตรียมอุดมศึกษาน้อมเกล้า กบินทร์บุรี	1
1	บ้านแก้งวิทยา	3
1	เขาฉกรรจ์วิทยาคม	4
1	สีคิ้ว "สวัสดิ์ผดุงวิทยา"	9
1	สาหร่ายวิทยาคม	12
1	โนนหันวิทยายน	25
1	ดงมูลวิทยาคม	28
1	ภูดินแดงวิทยา	31
1	มัธยมวัดมกุฏกษัตริย์	44
1	มัธยมวัดธาตุทอง	51
1	บดินทรเดชา (สิงห์ สิงหเสนี) ๔	52
1	สุขุมนวพันธ์อุปถัมภ์	56
1	ทวีธาภิเศก ๒	68
1	แจงร้อนวิทยา	70
1	วิมุตยารามพิทยากร	72
1	ราชวินิตบางแคปานขำ	73
1	วัดพุทธบูชา	75
1	หาดอมราอักษรลักษณ์วิทยา	76
1	นนทบุรีพิทยาคม	79
1	มัธยมวัดหัตถสารเกษตร	83
1	เตรียมอุดมศึกษาพัฒนาการลำลูกกา	88
1	ภาชี (สุนทรวิทยานุกูล)	92
1	วิเศษไชยชาญ (ตันติวิทยาภูมิ)	95
1	บ้านหมี่วิทยา	100
1	คุรุประชาสรรค์	102
1	มวกเหล็กวิทยา	106
1	ปลวกแดงพิทยาคม	113
1	ท่าใหม่ (พูลสวัสดิ์ราษฎร์นุกูล)	118
1	ขลุงรัชดาภิเษก	120
1	คิชฌกูฏวิทยา	122
1	เบญจมราชรังสฤษฎิ์ 2	124
1	มัธยมสิริวัณวรี 3 ฉะเชิงเทรา	126
1	กบินทร์วิทยา	128
1	บ้านนา (นายกพิทยากร)	132
1	วังสมบูรณ์วิทยาคม	134
1	บุญวัฒนา	138
1	ชุมพวงศึกษา	145
1	รมย์บุรีพิทยาคม รัชมังคลาภิเษก	149
1	กระเทียมวิทยา	169
1	ประสานมิตรวิทยา	178
1	นครศรีลำดวนวิทยา	180
1	ขุขันธ์	183
1	นารีนุกูล	189
1	เดชอุดม	198
1	กาญจนาภิเษกวิทยาลัย ชัยภูมิ	204
1	ภักดีชุมพลวิทยา	206
1	แวงใหญ่วิทยาคม	219
1	บัวใหญ่พิทยาคม	222
1	สร้างคอมวิทยา	226
1	ท่าลี่วิทยา	234
1	ศรีธวัชวิทยาลัย	240
1	สามชัย	245
1	ธรรมบวรวิทยา	250
1	นาแกพิทยาคม	252
1	ป่าซาง	270
1	ตรอนตรีสินธุ์	274

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
1	ทองแสนขันวิทยา	275
1	ฟากท่าวิทยา	276
1	ม่วงไข่พิทยาคม	280
1	วิไลเกียรติอุปถัมภ์	281
1	บ้านหลวง	282
1	ถ้ำปินวิทยาคม	288
1	ดงเจนวิทยาคม	289
1	ป่าแดดวิทยาคม	298
1	แม่เจดีย์วิทยาคม	299
1	ปล้องวิทยาคม	302
1	พญาเม็งราย	303
1	ขุนตาลวิทยาคม	305
1	ปายวิทยาคาร	306
1	ปางมะผ้าพิทยาสรรพ์	307
1	สวนเมี่ยงวิทยา	318
1	ผะดุงศิษย์พิทยา	346
1	เกษมพิทยา	349
1	บางกอกวิทยา (มูลนิธิ)	352
1	ประภามนตรี 2	355
1	อัสสัมชัญ นครราชสีมา	364
1	ประทีปศาสน์	381
1	ดาวรุ่งวิทยา	384
1	แสงประทีปวิทยามูลนิธิ	392
1	ตรังวิทยา	394
1	พัฒนาอิสลาม	398
2	ซับนกแก้วิทยา	2
2	วังไม้แดงพิทยาคม	11
2	กุดจิกวิทยา	13
2	กีฬาเทศบาลนครนครราชสีมา	17
2	เนินสง่าวิทยา	24
2	หัวเรือพิทยาคม	27

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
2	พัฒนาประชาอุปถัมภ์	33
2	สาธิตเทศบาลบ้านเชตวัน	34
2	ศรีอินทราทิตย์พิทยาคม	37
2	วัดราชาธิวาส	41
2	กุนนทีรุทธารามวิทยาคม	48
2	วัดรางบัว	69
2	วิสุทธิกษัตรี	77
2	บดินทรเดชา (สิงห์ สิงหเสนี) นนทบุรี	80
2	หนองเสือวิทยาคม	87
2	สาคลีวิทยา	94
2	ปากเพรียววิทยาคม[สวนกุหลาบวิทยาลัยสระบุรี]	103
2	หินกองวิทยาคม	105
2	ชำนาญสามัคคีวิทยา	114
2	นายายอามพิทยาคม	119
2	แหลมสิงห์วิทยาคม	121
2	ราชสาส์นวิทยา	125
2	วังตะเคียนวิทยาคม	129
2	องครักษ์	133
2	บุญเหลือวิทยานุสรณ์	139
2	มัธยมวชิราลงกรณวราราม	141
2	บ้านเหลื่อมพิทยาสรรพ์	144
2	บัวหลวงวิทยาคม	148
2	สิงหวิทยาคม	151
2	ร่วมจิตต์วิทยา	152
2	หนองหงส์พิทยาคม	153
2	นาโพธิ์พิทยาคม	156
2	กู่สวนแตงพิทยาคม	157
2	พญารามวิทยา	160
2	ท่าสว่างวิทยา	161
2	ห้วยจริงวิทยา	162

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
2	ลานทรายพิทยาคม	166
2	ราชประชานุเคราะห์ 29 จังหวัดศรีสะเกษ	174
2	ส้มป่อยพิทยาคม	181
2	โนนกระสังวิทยาคม	184
2	สหธาตุศึกษา	190
2	หกสิบพรรษาวิทยาคม อุบลราชธานี	191
2	พังเคนพิทยา	193
2	พะลานวิทยาคม	194
2	นาเยียศึกษา รัชมังคลาภิเษก	196
2	คำเขื่อนแก้ววิทยาคม	211
2	ศรีเจริญศึกษา	213
2	ดงมันพิทยาคม	221
2	ยิ่งยวดพิทยานุกูล	227
2	นาด้วงวิทยา	230
2	ธาตุพิทยาคม	231
2	วังม่วงพิทยาคม	236
2	โนนราษีวิทยา	238
2	เมยวดีพิทยาคม	243
2	โคกศรีเมือง	244
2	ห้วยผึ้งพิทยา	247
2	บ้านนาคูพัฒนา (กรป.กลาง อุปถัมภ์)	248
2	เต่างอยพัฒนศึกษา	249
2	นางัวราษฎร์รังสรรค์	254
2	ไชยปราการ	263
2	บ้านกาดวิทยาคม	264
2	ดอยเต่าวิทยาคม	267
2	ถิ่นโอภาสวิทยา	278
2	เชียงม่วนวิทยาคม	292
2	ปงรัชดาภิเษก	293

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
2	ห้วยสักวิทยาคม	294
2	บ้านแซววิทยาคม	300
2	สว่างอารมณ์วิทยาคม	313
2	วังเจ้าวิทยาคม	315
2	วังโป่งพิทยาคม	319
2	ชะอวดเคร่งธรรมวิทยา	322
2	ปากน้ำหลังสวนวิทยา	324
2	วิเชียรมาตุ 2	325
2	ธัญธารวิทยา	328
2	อิสลามสันติชน	353
2	ญามีอุ้ลอิควาน	356
2	อิสลามศรีอยุธยามูลนิธิ	358
2	อัสสัมชัญอุบลราชธานี	366
2	สหมิตรวิทยา	373
2	เมธีวุฒิกร	375
2	ลูกรักเชียงของ	378
2	อินทุภูติพิทยา	379
2	ตัรบียะห์อิสลามียะห์	382
2	ดาริสสลามวิทยา	383
2	รุ่งโรจน์วิทยา	386
2	ตัสดีกียะห์	387
2	บำรุงอิสลามมูลนิธิ	395
2	ศาสนูปถัมภ์ปากพะยูนมูลนิธิ	396
2	อิสลามศาสตร์ มูลนิธิ	397
2	มูฮัมมาดียะห์	399
3	เตรียมอุดมศึกษาน้อมเกล้า นครราชสีมา	6
3	โยธินบูรณะ	42
3	วัดสุทธิวราราม	49
3	บดินทรเดชา (สิงห์ สิงหเสนี) ๒	55
3	หอวัง	57

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
3	สตรีวิทยา ๒	60
3	เตรียมอุดมศึกษาน้อมเกล้า	62
3	มัธยมวัดนายโรง	67
3	สวนกุหลาบวิทยาลัย นนทบุรี	81
3	จุฬาภรณราชวิทยาลัย ปทุมธานี	85
3	จุฬาภรณราชวิทยาลัย ลพบุรี	99
3	ชลราษฎรอำรุง	107
3	ระยองวิทยาคม	112
3	เบญจมราชูทิศ จังหวัดจันทบุรี	116
3	เบญจมราชรังสฤษฎ์	123
3	เบ็ญจะมะมหาราช	188
3	ขอนแก่นวิทยายน	215
3	อุดรพิทยานุกูล	223
3	จุฬาภรณราชวิทยาลัย เชียงราย	296
3	สาธิต มหาวิทยาลัยเกษตรศาสตร์ บางเขน	331
3	สาธิตมหาวิทยาลัยราชภัฏ พระนคร	332
3	มัธยมสาธิตมหาวิทยาลัย ราชภัฏบ้านสมเด็จเจ้าพระยา	335
3	วชิราวุธวิทยาลัย	336
3	อัสสัมชัญคอนแวนต์	339
3	กรุงเทพคริสเตียนวิทยาลัย	340
3	มาแตร์เดอีวิทยาลัย	342
3	อัสสัมชัญสมุทรปราการ	354
3	มงฟอร์ตวิทยาลัย แผนกมัธยม	371
4	สองครพิทยาคม	8
4	พระทองคำวิทยา	10
4	หนองถ่มวิทยา	19
4	โพธิ์ธาตุประชาสรรค์	20

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
4	รัชตวิทยาคม	38
4	มณีเสวตรอุปถัมภ์	130
4	ลิ้นฟ้าพิทยาคม	175
4	บัวน้อยวิทยา	177
4	พอกพิทยาคม รัชมังคลาภิเษก	185
4	สำโรงวิทยาคาร	197
4	ท่าโพธิ์ศรีพิทยา	199
4	โคกนางามพิทยาสรรพ์	218
4	แวงน้อยศึกษา	220
4	ปากสวยพิทยาคม	237
4	ดงหลวงวิทยา	257
4	ประชาราชวิทยา	272
4	สันติสุขพิทยาคม	284
4	ราชประชานุเคราะห์ 21	309
4	หนองตูมวิทยา	316
4	นาสนุ่นวิทยาคม	321
4	ละอุ่นวิทยาคาร	323
4	ประชาบำรุง	326
4	พินิตประสาธน์	377
4	วชิรานุกูล	385
4	ดรุณวิทยา	388
4	ส่งเสริมศาสนา	390
5	ครบุรี	14
5	ไพรบึงวิทยาคม	18
5	หนองห้างพิทยา	29
5	เทศบาลเพชรวิทย์	35
5	เทศบาลเฉลิมพระเกียรติ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี	36
5	มัธยมวัดเบญจมบพิตร	39
5	วัดบวรนิเวศ	43

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
5	ไตรมิตรวิทยาลัย	46
5	มัธยมวัดบึงทองหลาง	53
5	วัดอินทาราม	65
5	มัธยมวัดดุสิตาราม	66
5	ทีปังกรวิทยาพัฒน์(ทวีวัฒนา) ในพระราชูปถัมภ์ฯ	74
5	ธรรมศาสตร์คลองหลวง วิทยาคม	84
5	สวนกุหลาบวิทยาลัย ชลบุรี	109
5	สัตหีบวิทยาคม	110
5	สิงห์สมุทร	111
5	สุนทรภู่พิทยา	115
5	ปราจิณราษฎรอำรุง	127
5	เฉลิมพระเกียรติสมเด็จพระศรีนครินทร์ ศรีสะเกษ	173
5	กุดข้าวปุ้นวิทยา	192
5	เมืองกลางประชานุกูล	203
5	โรงเรียนอำนาจเจริญ	209
5	กัลยาณวัตร	216
5	สตรีราชินูทิศ	224
5	จุฬาภรณราชวิทยาลัย เลย	232
5	สตรีศึกษา	239
5	ดอยสะเก็ดวิทยาคม	260
5	สันป่ายางวิทยาคม	261
5	กิ่วลมวิทยา	271
5	แม่สันวิทยา	273
5	เวียงเทพวิทยา	279
5	เมืองยมวิทยาคาร	285
5	ฟากกว้านวิทยาคม	287
5	เฉลิมพระเกียรติสมเด็จพระศรีนครินทร์ พะเยา	290
5	ทัพทันอนุสรณ์	312

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
5	บ้านทุ่งนาวิทยา	314
5	วังทองวิทยา	317
5	ศรีจันทร์วิทยาคม รัชมังคลาภิเษก	320
5	ควนพระสาครินทร์	327
5	ทิวไผ่งาม	344
5	เซนต์จอห์น	347
5	อัสสัมชัญคอนแวนต์ลำนารายณ์	359
5	อัสสัมชัญศรีราชา	360
5	มารีวิทยากบินทร์บุรี	362
5	วชิราลัย	372
5	วชิรวิทย์ ฝ่ายมัธยม	374
5	บัณฑิตศึกษาลัย	380
6	บัวใหญ่	5
6	สูงเนิน	7
6	วัดประชานิมิตร	15
6	ประทาย	16
6	ปรือใหญ่วิทยบัลลังก์	21
6	เบ็ตตี้ดูเมน 2 ช่องเม็ก	22
6	โนนกอกวิทยา	23
6	นาข่าวิทยาคม	26
6	สกลทวาปี	30
6	แม่อายวิทยาคม	32
6	จันทร์หุ่นบำเพ็ญ	54
6	วัดบวรมงคล	71
6	หอวัง ปทุมธานี	82
6	ปทุมวิไลวิทยาคม[เตรียมอุดมศึกษาน้อมเกล้า ปทุมธานี]	86
6	เทพศิรินทร์คลองสิบสาม ปทุมธานี	89
6	มัธยมสังคีตวิทยา กรุงเทพมหานคร	90

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
6	ท่าเรือ"นิตยานุกูล"	91
6	เสนา"เสนาประสิทธิ์"	93
6	โคกสำโรงวิทยา	96
6	โคกกะเทียมวิทยาลัย	98
6	พัฒนานิคม	101
6	หนองแค (สรกิจพิทยา)	104
6	ประจันตราษฎร์บำรุง	131
6	ตาพระยา	135
6	โนนสมบูรณ์วิทยา	140
6	มัธยมด่านขุนทด	142
6	ขามสะแกแสง	143
6	ภู่วิทยา	146
6	พลับพลาชัยพิทยาคม	150
6	ตูมใหญ่วิทยา	154
6	สำโรงทาบวิทยาคม	163
6	สินรินทร์วิทยา	164
6	ชุมพลวิทยาสรรค์	165
6	โนนแท่นพิทยาคม	167
6	กาบเชิงวิทยา	168
6	ศรีณรงค์พิทยาลัย	170
6	สตรีสิริเกศ	172
6	กันทรารมณ์	176
6	เขื่อนช้างวิทยาคาร	179
6	ทุ่งสิมวิทยาคม	182
6	กันทรลักษ์วิทยา	186
6	เบญจลักษ์พิทยา	187
6	อ่างศิลา	195
6	ทุ่งเทิงยิ่งวัฒนา	200
6	ราชประชานุเคราะห์ 28 จังหวัดยโสธร	201
6	ค้อวังวิทยาคม	202

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
6	บ้านเขว้าวิทยายน	205
6	เจียงทองพิทยาคม	207
6	หนองบัวระเหวิทยาคาร	208
6	ชานุมานวิทยาคม	210
6	เสนางคนิคม	212
6	ลืออำนาจวิทยาคม	214
6	สุมเส้าพิทยาคาร	225
6	เลยอนุกูลวิทยา	229
6	ศรีสองรักษ์วิทยา	233
6	สังคมวิทยา	235
6	ศรีสมเด็จพิมพ์พัฒนาวิทยา	241
6	ทรายทองวิทยา	242
6	ดอนจานวิทยาคม	246
6	อุ่มเหม้าประชาสรรค์	251
6	บ้านแพงพิทยาคม	253
6	ผึ่งแดดวิทยาคาร	255
6	มุกดาวิทยานุกูล	256
6	คำชะอีวิทยาคาร	258
6	หนองสูงสามัคคีวิทยา	259
6	สะเมิงพิทยาคม	262
6	สองแควิทยาคม	265
6	ฮอดพิทยาคม	266
6	อมก๋อยวิทยาคม	268
6	ราชประชานุเคราะห์ 26	269
6	ท่าปลาประชาอุทิศ	277
6	นาหมื่นพิทยาคม	283
6	ไตรเขตประชาสามัคคีรัชมังคลาภิเษก	286
6	ฝายกวางวิทยาคม	291
6	เตรียมอุดมศึกษาพัฒนาการเชียงราย	295
6	พานพิเศษพิทยา	297

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
6	ห้วยซ้อวิทยาคมรัชมังคลาภิเษก	301
6	เวียงแก่นวิทยาคม	304
6	เฉลิมรัชวิทยาคม	308
6	พยุหะพิทยาคม	310
6	บรรพตพิสัยพิทยาคม	311
6	สวนพระยาวิทยา	329
6	เบญจมินทร์	345
6	ท่าอิฐศึกษา	357
6	ลาซาลจันทบุรี (มารดาพิทักษ์)	361
6	ขุขันธ์ราษฎร์บำรุง	365
6	อาเวมารีอา	367
6	มัธยมวัดกลางโกสุม	368
6	อรุโณทัย	376
6	แสงธรรมวิทยามูลนิธิ	391
6	สามัคคีอิสลามวิทยา	393
7	สุรศักดิ์มนตรี	47
7	ราชนันทาจารย์สามเสน วิทยาลัย 2	50
7	สารวิทยา	58
7	ฤทธิยะวรรณาลัย	61
7	ราชวินิตบางแก้ว ในพระบรมราชูปถัมภ์	78
7	พิบูลวิทยาลัย	97
7	ชลกันยานุกูล	108
7	ศรียานุสรณ์	117
7	ราชสีมาวิทยาลัย	136
7	สุรนารีวิทยา	137
7	บุรีรัมย์พิทยาคม	147
7	จุฬาภรณราชวิทยาลัย บุรีรัมย์	155
7	สุรวิทยาคาร	158
7	สิรินธร	159
7	ศรีสะเกษวิทยาลัย	171

กลุ่มแฝง	รายชื่อโรงเรียน	รหัส
7	แก่นนครวิทยาลัย	217
7	เลยพิทยาคม	228
7	สารสาสน์วิเทศร่มเกล้า	343
7	พระหฤทัยคอนแวนต์	348
7	อัสสัมชัญธนบุรี	350
7	สารสาสน์วิเทศบางบอน	351
7	มารีย์วิทยา	363
7	ดาราวิทยาลัย	370
8	สวนกุหลาบวิทยาลัย	40
8	เตรียมอุดมศึกษาพัฒนาการ	59
8	บดินทรเดชา (สิงห์ สิงหเสนี)	63
8	ศึกษานารี	64
8	ดรุณสิกขาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี	330
8	สาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม	333
8	สาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ปทุมวัน	334
8	เซนต์คาเบรียล	337
8	อัสสัมชัญ	338
8	เซนต์โยเซฟคอนเวนต์	341
8	ปรินส์รอยแยลส์วิทยาลัย	369
8	มอ. วิทยานุสรณ์	389
9	เตรียมอุดมศึกษา	45
9	มหิดลวิทยานุสรณ์	400

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประวัติผู้เขียนวิทยานิพนธ์

นางภัทรภร เจนสุทธิเวชกุล เกิดเมื่อวันที่ 4 มิถุนายน 2527 สำเร็จการศึกษาปริญญาครุศาสตรบัณฑิต (เกียรตินิยมอันดับ 2) วิชาเอกวิทยาศาสตร์ทั่วไป-คณิตศาสตร์ สาขามัธยมศึกษา วิทยาศาสตร์ ภาควิชาหลักสูตร การสอนและเทคโนโลยีการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2548 จากนั้นเข้าทำงานในตำแหน่งอาจารย์ประจำกลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม จนถึงปัจจุบัน ระหว่างที่ทำงานได้ลาศึกษาต่อในเวลาราชการบางส่วน โดยเข้าศึกษาต่อการศึกษาปริญญาครุศาสตรมหาบัณฑิต สาขาวิจัยการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2550 และสำเร็จการศึกษาเมื่อปีการศึกษา 2551 และได้เข้าศึกษาในหลักสูตรครุศาสตรดุษฎีบัณฑิต สาขาการวัดและประเมินผลทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2553 โดยระหว่างที่ศึกษาได้รับทุน 90 ปีจุฬาลงกรณ์มหาวิทยาลัย ในการทำวิทยานิพนธ์

ปัจจุบันดำรงตำแหน่งอาจารย์ประจำกลุ่มสาระการเรียนรู้คณิตศาสตร์ และดำรงตำแหน่งผู้ช่วยผู้อำนวยการฝ่ายวิจัยและบริการวิชาการ และโรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University