

ประสิทธิภาพของการประมาณค่าพารามิเตอร์แบบสุ่มข้อสอบ สำหรับโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์

Efficiency of Parameter Estimation with Random Item for One-Parameter Response Model

ทิพวัลย์ ปัญจมะวัต และคณะ

บทคัดย่อ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาประสิทธิภาพการประมาณค่าพารามิเตอร์ผู้สอบสำหรับโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ด้วยข้อมูลจำลอง และศึกษาความสอดคล้องกลมกลืนระหว่างโมเดลกับข้อมูลเชิงประจักษ์ ระหว่างการประมาณค่าด้วยโมเดลประมาณค่าพารามิเตอร์แบบอิทธิพลเจาะจงข้อสอบกับแบบอิทธิพลสุ่มข้อสอบ ๓ รูปแบบ คือ สุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI), เจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) และสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ข้อมูลจำลองจากโปรแกรม R และประมวลผลด้วยโปรแกรม WinBUGS ด้วยการเชื่อมโยงผ่านคำสั่ง R2WinBUGS package การจำลองข้อมูลทั้งหมด ๔๔ เจ็นไนซ์ ($4 \times 11 \times 1$) ประกอบด้วย ลักษณะการแจกแจงก่อนหน้าของความยากข้อสอบ ๔ แบบ ความยาวของแบบทดสอบ ๓ ระดับ และจำนวนผู้สอบ ๔ ขนาด ประสิทธิภาพของการประมาณค่าพารามิเตอร์ผู้สอบด้วยข้อมูลจำลองพิจารณาจากความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์และค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง และความสอดคล้องกลมกลืนระหว่างโมเดลกับข้อมูลเชิงประจักษ์จากผลการสอบ O-Net วิชาภาษาอังกฤษ คณิตศาสตร์ และภาษาไทย ปีการศึกษา ๒๕๕๓ ของนักเรียนมัธยมศึกษาปีที่ ๖ พิจารณาจากค่าสถิติ Akaike Criterion (AIC)

ผลการวิจัยพบว่า ๑) โมเดลการประมาณค่าพารามิเตอร์ผู้สอบรูปแบบ FPRI มีประสิทธิภาพในการประมาณค่าสูงที่สุด รองลงมาคือ RPFI และ RPRI ตามลำดับ ๒) โมเดลการประมาณค่าพารามิเตอร์ผู้สอบรูปแบบ FPRI มีค่า AIC ต่ำที่สุดทั้ง ๓ รายวิชาและในกลุ่มตัวอย่างทุกขนาดรองลงมาคือ RPFI และ RPRI ตามลำดับ

คำสำคัญ: ประสิทธิภาพการประมาณค่า/การประมาณค่าพารามิเตอร์/สุ่มข้อสอบ

Abstract

The purpose of this research was to investigate the efficiency of the examinee's parameter estimation by using the one-parameter Item Response Model with simulated data and to investigate the model with real data. The parameter consisted of the evaluation of specific item response and of three random item response models: RPFI, FPRI, and RPRI. The data was formulated by using R program to run the simulation assessed by WinBUGS program and connected through the R2WinBUGS package code. The data model consisted of 48 conditions (4x3x4): including the initial distribution of the test in 4 distributions of difficulty; 3 levels of the test length; and 4 sizes of the number of examinees. The efficiency evaluation of the examinees' parameter estimation was considered from the standard error of estimation and the mean squared error. In addition, this research used the data from the 2003's O-NET test (English, Mathematics and Thai) results of Mathayom 6 students by assessing the model with real data from Akaike Criterion (AIC) statistics.

The results were as follows: 1) The FPRI parameter estimation model shown the highest estimation efficiency, followed by the RPFI and RPRI parameter estimation models, respectively. 2) The FPRI parameter estimation model had the least AIC in every subject and every sample size. The study revealed the most efficient estimation by FPRI, RPFI and RPRI respectively according to the O-NET data.

KEYWORDS: ESTIMATION EFFICIENCE/PARAMETER ESTIMATION/RANDOM ITEM

บทนำ

การวัดผลทางการศึกษาด้วยแบบทดสอบ เพื่อประเมินความสามารถของผู้สอบ เป็นกระบวนการเชิงระบบเพื่ออ้างอิงไปสู่คุณลักษณะแฝง (Latent traits) ของบุคคล ซึ่งจำเป็นต้องใช้กระบวนการวัดผลที่มีคุณภาพให้สามารถประมาณค่าพารามิเตอร์ความสามารถของผู้สอบได้อย่างถูกต้องแม่นยำมากที่สุด เพื่อใช้ตัดสินผลการศึกษาของผู้เรียนในการจบการศึกษาในแต่ละระดับชั้น การวัดระดับความสามารถของผู้เรียนในแต่ละช่วงชั้นตามหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. ๒๕๕๑ รวมทั้งการสอบคัดเลือกเพื่อเข้าศึกษาต่อในระดับที่สูงขึ้นไป ดังนั้นกระบวนการวัดผลแต่ละขั้นตอนจึงต้องมีคุณภาพของการวัดที่ดีน่าเชื่อถือและเป็นที่ยอมรับ กระบวนการหนึ่งที่สำคัญคือ การประมาณค่าพารามิเตอร์ของผู้สอบและค่าพารามิเตอร์ของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ที่มีบทบาทในการประมาณค่าพารามิเตอร์ข้อสอบและผู้สอบในปัจจุบัน โดยมีแนวคิดสำคัญคือ ค่าพารามิเตอร์ของข้อสอบมีความสัมพันธ์กับความสามารถของผู้ตอบ จึงทำให้ขณะที่ประมาณค่าพารามิเตอร์ของข้อสอบก็จะพิจารณาไปพร้อมๆ กับการประมาณค่าความสามารถของผู้สอบ ค่าพารามิเตอร์ของข้อสอบที่ได้จึงมีนัยทั่วไปไม่แปรผันไปตามความสามารถของผู้สอบ ส่วนความคลาดเคลื่อนมาตรฐานของการวัดนั้น IRT จะวิเคราะห์ทั้งรายข้อและแบบสอบรวมทั้งฉบับจำแนกตามความสามารถของผู้ตอบ จึงทำให้ผู้สอบที่มีความสามารถต่างกันจะมีความคลาดเคลื่อนมาตรฐานของการวัดที่แตกต่างกัน อีกทั้งใน IRT ไม่มีข้อจำกัดในการเปรียบเทียบคะแนนระหว่างแบบสอบว่าต้องเป็นแบบสอบคู่ขนาน

จึงทำให้สามารถเปรียบเทียบแบบสอบต่างฉบับที่มีความยากแตกต่างกันได้ การวัดผลทางการศึกษาในปัจจุบันจึงนิยมใช้ทฤษฎีการตอบสนองข้อสอบในการประมาณค่าความสามารถของผู้สอบและค่าพารามิเตอร์ของข้อสอบ (ศิริชัย กาญจนวาสี, ๒๕๕๕)

ปัจจุบันโมเดลการตอบสนองข้อสอบมักใช้อิทธิพลเจาะจงข้อสอบที่สมมติให้ตัวแปรอิสระเป็นตัวแปรคงที่ในโมเดลจึงไม่มีเทอมของความคลาดเคลื่อนจากค่าเฉลี่ยเพื่อวิเคราะห์ค่าคุณลักษณะแฝงของผู้สอบ ซึ่งเป็นการวัดความสามารถของผู้สอบโดยใช้ผลการตอบข้อสอบของผู้สอบ ด้วยวิธีการเจาะจงข้อสอบและสุ่มผู้สอบ และใช้วิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีการประมาณค่าที่เป็นไปได้สูงสุดแบบประมาณค่าปลายทาง (Marginal maximum likelihood method) ซึ่งวิธีการประมาณค่าลักษณะนี้ทำให้การประมาณค่าความสามารถผู้สอบที่ได้จะเป็นความสามารถที่สรุปอ้างอิงไปยังประชากรของผู้สอบ มิได้เป็นความสามารถของผู้สอบแต่ละคน และได้ค่าความยากของข้อสอบแต่ละข้อ ที่ไม่สามารถสรุปอ้างอิงไปยังประชากรข้อสอบได้ โมเดลดังกล่าวนี้เป็นโมเดลอิทธิพลสุ่มผู้สอบ-เจาะจงข้อสอบ โดยประมาณค่าด้วยการกำหนดค่าความแปรปรวนหรือค่าการกระจายก่อนหน้าของพารามิเตอร์ความสามารถผู้สอบมาเป็นค่าตั้งต้นเพื่อประมาณค่าความยากของข้อสอบ แต่ในสถานการณ์ของการทดสอบสิ่งที่ผู้สอบและผู้สอบต้องการอย่างแท้จริงคือการสรุปอ้างอิงไปยังประชากรข้อสอบเพื่อวัดความสามารถของบุคคลแต่ละคน ที่ทำได้ด้วยการใช้โมเดลการประมาณค่าแบบอิทธิพลเจาะจงบุคคลและอิทธิพลข้อสอบเชิงสุ่ม

โมเดลการวัดบุคคลที่ใช้อิทธิพลเจาะจงบุคคลและอิทธิพลข้อสอบเชิงสุ่มนี้ จะทำให้สามารถประมาณค่าความสามารถของบุคคลได้โดยตรงและเพียงพอที่จะสรุปอ้างอิงผลการวัดนี้ให้ครอบคลุมข้อสอบข้อต่างๆ ซึ่งทำให้การวัดคุณลักษณะแฝงหรือความสามารถนั้นครอบคลุมคุณลักษณะที่ต้องการวัดและสามารถสรุปได้อย่างมั่นใจว่าเชื่อถืออีกทั้งการใช้โมเดลการตอบสนองข้อสอบแบบอิทธิพลสุ่มข้อสอบ มีแนวโน้มที่จะให้ความคลาดเคลื่อนในการประมาณค่าลดลง คือ มีแนวโน้มที่จะให้ประสิทธิภาพในการประมาณค่าดีขึ้น รวมทั้งประสิทธิภาพของโมเดลแบบอิทธิพลสุ่มข้อสอบมีความสอดคล้องกับข้อมูลดีกว่าโมเดลแบบอิทธิพลเจาะจงข้อสอบ (De Boeck, 2008) ผู้วิจัยจึงสนใจศึกษาเกี่ยวกับประสิทธิภาพของการวัดในด้านการประมาณค่าความสามารถผู้สอบผ่านทางทฤษฎีการตอบสนองข้อสอบแบบอิทธิพลสุ่มข้อสอบ โดยใช้โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ซึ่งเป็นโมเดลที่มีผู้นิยมนำมาใช้เนื่องจากมีจำนวนพารามิเตอร์ที่ต้องประมาณค่าน้อยกว่าโมเดลทฤษฎีการตอบสนองข้อสอบอื่นๆ ทำให้ง่ายต่อการวิเคราะห์ ไม่ยุ่งยากซับซ้อน (Hambleton, 1989) อีกทั้งโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์นี้ยังให้ค่าประมาณของพารามิเตอร์ที่ถูกต้องกว่าโมเดลแบบอื่นๆ เมื่อนำไปใช้กับกลุ่มตัวอย่างขนาดเล็ก (Vijver, 1986; Lord, 1980)

วัตถุประสงค์

๑. เพื่อศึกษาประสิทธิภาพของการประมาณค่าของโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ระหว่างการประมาณค่าโดยใช้รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ กับรูปแบบเจาะจงผู้สอบ-สุ่ม

ข้อสอบและรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ กับรูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ ด้วยข้อมูลจำลอง

๒. เพื่อศึกษาความสอดคล้องของโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ ที่มีการประมาณค่าโดยใช้รูปแบบการสุ่มผู้สอบ-เจาะจงข้อสอบกับ รูปแบบการเจาะจงผู้สอบ-สุ่มข้อสอบ และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ กับรูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ ด้วยข้อมูลเชิงประจักษ์

สมมติฐานการวิจัย

๑. โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ จะมีความคลาดเคลื่อนในการประมาณค่าน้อยกว่า โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ ที่มีขนาดกลุ่มตัวอย่างเท่ากันในระดับขนาดกลุ่มตัวอย่างหลายระดับ และจำนวนข้อสอบที่แตกต่างกันจากข้อมูลจำลอง

๒. โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ จะมีค่าเฉลี่ยความเบี่ยงเบนยกกำลังสองน้อยกว่า โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ ที่มีขนาดกลุ่มตัวอย่างเท่ากันในระดับขนาดกลุ่มตัวอย่างหลายระดับ และจำนวนข้อสอบที่แตกต่างกันจากข้อมูลจำลอง

๓. โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ จะมีประสิทธิภาพของโมเดลสูงกว่า โมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบจากข้อมูลเชิงประจักษ์

วิธีการวิจัย

การวิจัยนี้ศึกษาโดยวิธีการจำลองข้อมูล เนื่องจากสามารถจัดกระทำกับข้อมูลได้หลากหลาย ทำให้ได้แนวทางพิจารณาเลือกใช้วิธีการที่เหมาะสมกับสภาพการทดสอบในสถานการณ์ต่างๆ และได้นำรูปแบบการประมาณค่าทั้ง ๓ รูปแบบ มาทดลองใช้กับผลการทดสอบในสถานการณ์จริงของการสอบ O-NET ปีการศึกษา ๒๕๕๓ รายละเอียดในการการวิจัยมีดังนี้

๑) ตัวแปรที่ศึกษา

ตัวแปรต้น ได้แก่ รูปแบบการประมาณค่าการแจกแจงก่อนหน้าของความยากข้อสอบ (Prior distribution of item difficulty) ขนาดกลุ่มตัวอย่าง และความยาวข้อสอบ ในการศึกษาด้วยข้อมูลจำลองมีรายละเอียดแต่ละตัวแปร ดังนี้

รูปแบบการประมาณค่า ๓ รูปแบบ ได้แก่ รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI)

การแจกแจงก่อนหน้าของความยากข้อสอบ ๔ รูปแบบ ได้แก่ Normal, Uniform, Negative skew และ Positive skew เพื่อศึกษาลักษณะการแจกแจงก่อนหน้าของความยากข้อสอบในสถานการณ์การทดสอบที่มีจุดประสงค์แตกต่างกัน เช่น การแจกแจงแบบ Negative skew สำหรับการทดสอบเพื่อคัดคนเก่ง

ความยาวแบบสอบ ๓ ระดับ ได้แก่ ข้อสอบจำนวน ๓๐, ๖๐ และ ๙๐ ข้อ เพื่อศึกษาให้ครอบคลุมในกรณีความยาวแบบสอบน้อย ปานกลาง และมาก

ขนาดกลุ่มตัวอย่างผู้สอบ ๔ ขนาด ได้แก่ ผู้สอบ จำนวน ๑๐๐, ๓๐๐, ๕๐๐ และ ๑,๐๐๐ คน เพื่อศึกษาให้ครอบคลุมในกรณีขนาดกลุ่มตัวอย่างน้อย ปานกลาง และมาก

ตัวแปรตาม คือ ค่าประมาณพารามิเตอร์ความสามารถผู้สอบ (θ) และค่าประมาณพารามิเตอร์ความยากข้อสอบ (β) โดยพิจารณาประสิทธิภาพการประมาณค่าพารามิเตอร์ข้อมูลจำลองจากค่าความคลาดเคลื่อนในการประมาณค่า ($SE(\theta)$) และค่าเฉลี่ยความคลาดเคลื่อนในการประมาณค่า (MSE) และพิจารณาประสิทธิภาพของโมเดลการประมาณค่าจากค่าสถิติ Akaike Information Criterion (AIC)

๒) รูปแบบการจำลองข้อมูล

การจำลองข้อมูลโดยใช้โปรแกรม R จำลองประชากรในการศึกษาครั้งนี้ กำหนดให้การแจกแจงก่อนหน้าของพารามิเตอร์ผู้สอบ (Prior Distribution of Person) เป็น Standard Normal (๐,๑) การจำลองข้อมูลในสถานการณ์ต่างๆ ประกอบด้วย การแจกแจงก่อนหน้าของพารามิเตอร์ความยาก ความยาวข้อสอบและขนาดกลุ่มตัวอย่างผู้สอบ ($4 \times 3 \times 4$) รวม ๔๘ เงื่อนไข ทำการวิเคราะห์ซ้ำ (replication) ๑,๐๐๐ รอบในแต่ละเงื่อนไข และประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดลการประมาณค่าพารามิเตอร์ ๓ รูปแบบ รวมทั้งสิ้น ๑๔๔ กรณี โดยการส่งผ่านข้อมูลที่จำลองแล้วด้วยชุดคำสั่ง R2WINBUGS เพื่อวิเคราะห์ข้อมูลด้วยโปรแกรม WinBUGS ดังมีขั้นตอนการดำเนินการวิจัย ดังแผนภาพที่ ๑ การวิเคราะห์ประสิทธิภาพของการประมาณค่าจากข้อมูลจำลองโดยพิจารณาจากค่าความคลาดเคลื่อนในการประมาณค่าและ

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง มีสูตรการคำนวณ ดังนี้

ค่าความคลาดเคลื่อนในการประมาณค่า ;

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \text{ และ}$$

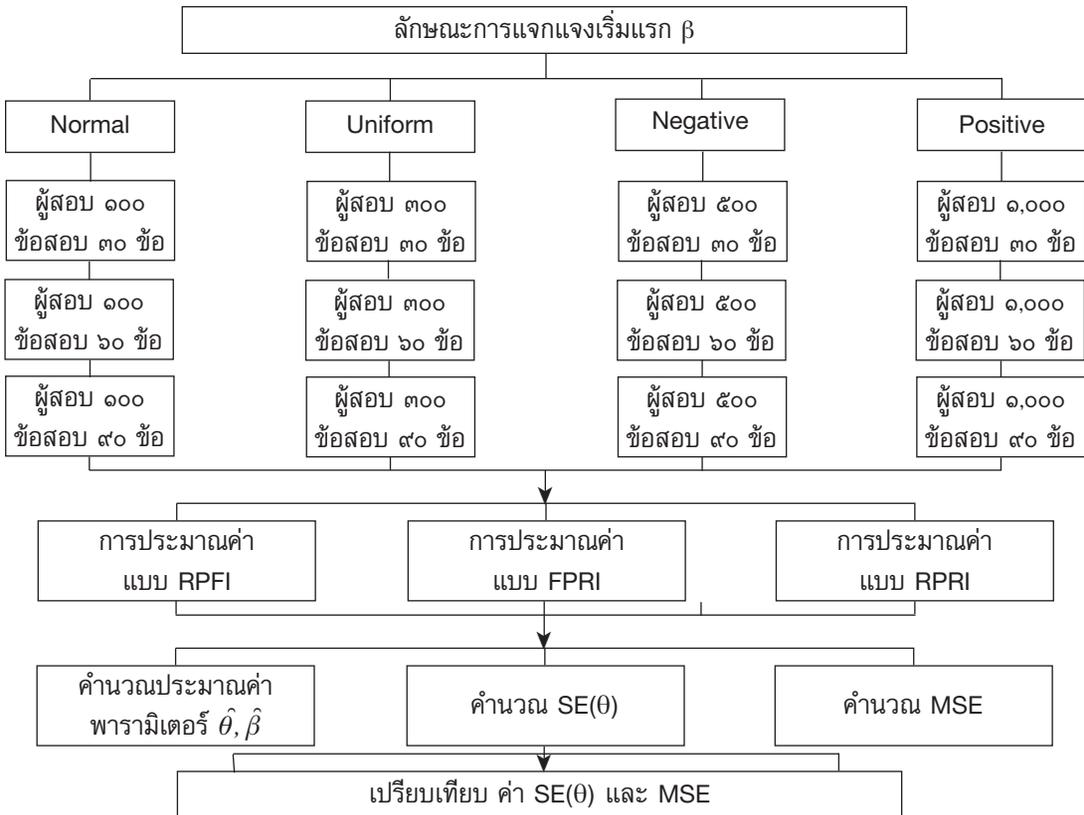
ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง;

$$MSE(\hat{\theta}) = (BIAS(\hat{\theta}, \theta))^2 + VARIANCE(\hat{\theta})$$

๓) รูปแบบการศึกษาข้อมูลจริง

การศึกษากับข้อมูลจริง โดยจะใช้ข้อมูลจากผลการสอบ O-NET ปี พ.ศ. ๒๕๕๓ รายวิชาคณิตศาสตร์ ภาษาอังกฤษ และภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ ๖ ที่มีจำนวนข้อสอบ ๔๐, ๗๐ และ ๙๐ ข้อตามลำดับ ใช้คำสั่งจากโปรแกรม R สุ่มขนาดกลุ่มตัวอย่างผู้สอบ จำนวน ๑๐๐, ๓๐๐,

๕๐๐ และ ๑,๐๐๐ คน โดยการใช้การสุ่มแบบใส่คืนและทำซ้ำ (bootstrap sampling) ๑๐๐ รอบนำมาวิเคราะห์ประมาณค่าพารามิเตอร์โดยใช้โมเดลการตอบสนองข้อสอบ ๓ รูปแบบ คือ รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) วิเคราะห์ประสิทธิภาพของโมเดล โดยพิจารณาความสอดคล้องของโมเดลกับข้อมูลจริง จากค่าสถิติ Akaike Information Criterion (AIC) มีสูตรคำนวณ คือ $AIC = 2k - 2\ln(L)$ เมื่อ k เป็นจำนวนพารามิเตอร์ในโมเดล และ L คือค่าสูงสุดของฟังก์ชันความน่าจะเป็น (Likelihood function) ของโมเดลที่ประมาณค่า มีขั้นตอนการดำเนินการวิจัยดังแผนภาพที่ 1



แผนภาพที่ ๑ ขั้นตอนการดำเนินการวิจัยกรณีการจำลองข้อมูล

ผลการวิจัย

ผลการวิเคราะห์ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจากข้อมูลจำลองพบว่า เมื่อมีการแจกแจงก่อนหน้าของความยากข้อสอบแบบ Normal การประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบต่ำกว่าโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) และโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ทั้งหมด ๑๒ เงื่อนไขที่ศึกษาและเมื่อมีการแจกแจงก่อนหน้าของความยากข้อสอบแบบ Uniform, Positive skew และ Negative skew พบว่า การประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เงื่อนไขที่ศึกษาส่วนใหญ่มีค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบต่ำกว่าโมเดลการประมาณค่าอีก ๒ รูปแบบ และค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าพารามิเตอร์เมื่อประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เกือบทุกเงื่อนไขที่ศึกษามีค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบต่ำกว่าโมเดลการประมาณค่าอื่นๆ ที่ศึกษารายละเอียดแสดงในตารางที่ ๑

ผลการวิเคราะห์ประสิทธิภาพของโมเดลการประมาณค่าพารามิเตอร์ผู้สอบจากผลการทดสอบ O-NET ระดับมัธยมศึกษาปีที่ ๖ ปีการศึกษา ๒๕๕๓ วิชาคณิตศาสตร์ ๔๐ ข้อ วิชาภาษาอังกฤษ ๗๐ ข้อ และวิชาภาษาไทย ๙๐ ข้อ พบว่า

การประมาณค่าพารามิเตอร์ผู้สอบด้วยโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีค่า AIC ต่ำกว่าการประมาณค่าด้วยโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) และโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ในทุกขนาดกลุ่มตัวอย่างทั้ง ๓ รายวิชา

เปรียบเทียบผลการวิเคราะห์ข้อมูลในครั้งนี้ ใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว (one way ANOVA) เปรียบเทียบค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองระหว่างโมเดลการตอบสนองข้อสอบ ϕ พารามิเตอร์ที่มีรูปแบบการประมาณค่า ๓ รูปแบบจากข้อมูลจำลอง จากการวิเคราะห์พบว่า โมเดลการประมาณค่าพารามิเตอร์ผู้สอบรูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบ ($SE(\theta)$) และค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (MSE) ต่ำกว่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) อย่างมีนัยสำคัญทางสถิติและต่ำกว่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) อย่างไม่มีนัยสำคัญในเกือบทุกเงื่อนไข และแนวโน้มประสิทธิภาพของการประมาณค่าเพิ่มขึ้นแปรผกผันกับขนาดกลุ่มตัวอย่างผู้สอบและความยาวแบบสอบ กล่าวคือ ประสิทธิภาพของการประมาณค่ามีแนวโน้มเพิ่มขึ้นเมื่อมีขนาดกลุ่มตัวอย่างใหญ่ขึ้น หรือเมื่อมีความยาวแบบสอบมากขึ้น โดยพิจารณาจากค่าความคลาดเคลื่อนในการประมาณและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองส่วนใหญ่มีค่าลดลงเมื่อมีขนาดกลุ่มตัวอย่างใหญ่ขึ้นหรือมีความยาวแบบสอบมากขึ้น แสดงไว้ดังตาราง ๑

ตาราง ๑ ผลการวิเคราะห์ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ผู้สอบและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าพารามิเตอร์จากข้อมูลจำลอง

เงื่อนไขการศึกษา	SE(θ) ($\times 10^{-3}$)			MSE ($\times 10^{-3}$)		
	RPMI	FPRI	RPRI	RPMI	FPRI	RPRI
Normal-100-30	297.476	294.065*	294.699	95.832	86.675*	87.594
100-60	213.058	210.961*	211.415	56.772	46.557*	46.762
100-90	174.21	172.944*	173.153	43.548	32.503*	32.734
300-30	295.819	295.036*	295.391	88.015	86.624*	87.070
300-60	210.683	210.005*	210.073	49.087	46.947*	47.017
300-90	173.136	172.560*	172.684	483.138	462.159	462.075*
500-30	293.483	292.752*	292.779	85.401	84.655*	84.788
500-60	211.077	210.715*	210.817	472.289	457.229	457.213*
500-90	173.002	172.606*	172.659	34.326	32.640*	32.692
1000-30	294.558	294.348*	294.455	85.307*	85.396	85.460
1000-60	210.563	210.274*	210.351	46.995	46.488*	46.511
1000-90	172.644	172.434*	172.442	32.946	32.430*	32.431
Uniform-100-30	343.118	341.866*	341.991	116.384	112.254*	113.025
100-60	245.673	244.732*	244.772	67.442	62.794*	62.917
100-90	173.153	208.802*	208.835	53.265	47.424*	47.543
300-30	350.281	350.061*	350.098	113.490	112.789*	113.152
300-60	258.725*	259.034	259.047	66.294*	66.346	66.451
300-90	208.934*	209.004	209.017	6326.157	6205.113*	6205.339
500-30	326.048	325.778*	325.807	99.422	99.044*	99.225
500-60	237.282	237.121*	237.145	5538.770	5478.464	5480.343*
500-90	196.977	196.841*	196.846	39.787	39.415*	39.433
1000-30	326.935	326.831*	326.850	99.160	99.094*	99.207
1000-60	236.630	236.594*	236.611	54.875*	54.910	54.974
1000-90	199.343	199.303*	199.304	39.573	39.502*	39.513

ตารางที่ ๑ (ต่อ)

เงื่อนไขการศึกษา	SE(θ) ($\times 10^{-3}$)			MSE ($\times 10^{-3}$)		
	RPFI	FPRI	RPRI	RPFI	FPRI	RPRI
Negative-100-30	331.747	330.346*	330.490	110.396	105.650*	106.522
100-60	240.219	239.447*	239.448	65.246	61.447*	61.760
100-90	198.209	197.986*	197.982	50.079	45.440*	45.587
300-30	373.449	373.123*	373.175	125.180	124.490*	124.867
300-60	257.220	256.965*	256.988	65.704	65.065*	65.147
300-90	205.859	205.626	205.620*	7696.569	6724.298	6723.972*
500-30	344.260	344.236*	344.303	109.384	109.085*	109.323
500-60	237.584	237.499*	237.503	5987.255	5554.726*	5554.972
500-90	193.505	193.343*	193.347	38.391	38.015*	38.049
1000-30	328.869	328.666*	328.699	100.097	100.016*	100.152
1000-60	230.255	230.118	230.106*	52.361	52.315*	52.381
1000-90	191.451	191.332*	191.333	37.014	37.012*	37.049
Positive-100-30	353.124*	356.064	356.547	122.243*	124.274	125.568
100-60	255.44*	261.466	261.652	73.694*	86.287	87.339
100-90	201.46*	203.716	203.774	51.543*	56.061	56.582
300-30	380.068	378.553*	378.602	128.941	128.118*	128.527
300-60	242.592	242.250*	242.253	59.415	58.736*	58.803
300-90	200.760	200.440*	200.454	7247.412	6409.414*	6409.449
500-30	328.420	327.940*	328.085	101.940	101.519*	101.656
500-60	245.407	245.015*	245.015	8151.511	6927.219*	6927.640
500-90	191.631	191.542*	191.557	37.779	37.650*	37.712
1000-30	336.806	336.668*	336.686	103.932	103.806*	103.896
1000-60	230.279	230.136*	230.142	52.200	52.046*	52.079
1000-90	192.437	192.336*	192.344	37.180	37.073*	37.106

หมายเหตุ: *คือ โมเดลที่มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ดีที่สุด โดยมีค่า SE(θ) หรือ MSE ต่ำที่สุด

ผลการวิเคราะห์ข้อมูลเชิงประจักษ์จากผลการทดสอบ O-NET ปี ๒๕๕๓ ระดับมัธยมศึกษาปีที่ ๖ รายวิชาภาษาอังกฤษ คณิตศาสตร์ และภาษาไทย ในแต่ละขนาดกลุ่มตัวอย่าง ๑๐๐, ๓๐๐, ๕๐๐ และ ๑,๐๐๐ คน โดยใช้การสุ่มแบบใส่คืนและทำซ้ำ (bootstrap sampling) ๑๐๐ รอบ พบว่า โมเดลการประมาณค่าพารามิเตอร์ผู้สอบทั้ง ๓ โมเดลมีความสอดคล้องระหว่างโมเดลกับข้อมูลไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติแต่เมื่อ

พิจารณาจากเกณฑ์การคัดเลือกโมเดลที่มีความสอดคล้องกับข้อมูลที่สุดจะมีค่า AIC ต่ำที่สุดพบว่า โมเดลประมาณค่าพารามิเตอร์รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีค่า AIC ต่ำที่สุดในทุกรายวิชาและในทุกขนาดกลุ่มตัวอย่างผู้สอบ รองลงมาคือ รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ตามลำดับแสดงไว้ดังตาราง ๒

ตาราง ๒ ผลการวิเคราะห์ประสิทธิภาพของโมเดลการประมาณค่าพารามิเตอร์ผู้สอบ

วิชา	ขนาดกลุ่มตัวอย่าง	AIC			F- test	p
		RPFI	FPRI	RPRI		
คณิตศาสตร์ (๔๐ ข้อ)	100	4,632.61	4,629.06*	4,633.88	0.139	0.870
	300	13,864.87	13,853.39*	13,867.04	0.444	0.642
	500	23,090.06	23,071.40*	23,092.86	0.518	0.596
	1,000	46,126.43	46,086.57*	46,127.45	1.115	0.329
ภาษาอังกฤษ (๗๐ ข้อ)	100	7,643.82	7,643.81*	7,645.89	0.021	0.980
	300	22,761.72	22,760.13*	22,764.45	0.025	0.975
	500	37,920.58	37,916.69*	37,923.68	0.039	0.962
	1,000	75,818.36	75,808.32*	75,819.41	0.050	0.952
ภาษาไทย (๙๐ ข้อ)	100	10,877.42	10,876.91*	10,880.97	0.069	0.934
	300	32,459.31	32,455.11*	32,463.36	0.076	0.927
	500	54,104.71	54,096.37*	54,107.81	0.111	0.895
	1,000	107,977.76	107,915.19*	107,983.04	0.142	0.867

หมายเหตุ: *คือ โมเดลที่มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ดีที่สุดในค่า AIC ต่ำที่สุด

อภิปรายผล

๑. ผลการวิจัยด้วยข้อมูลจำลองส่วนใหญ่เป็นไปตามสมมติฐาน ข้อที่ ๑ นั่นคือ โมเดลประมาณค่าพารามิเตอร์ผู้สอบ ๑ พารามิเตอร์รูปแบบการสุ่มข้อสอบ มีความคลาดเคลื่อนในการประมาณค่าน้อยกว่า โมเดลประมาณค่า

พารามิเตอร์ผู้สอบ ๑ พารามิเตอร์รูปแบบการเจาะจงข้อสอบ ณ ขนาดกลุ่มตัวอย่างและความยาวแบบสอบเดียวกัน และเป็นไปตามสมมติฐาน ข้อที่ ๒ คือ โมเดลประมาณค่าพารามิเตอร์ผู้สอบ ๑ พารามิเตอร์รูปแบบการสุ่มข้อสอบ มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองน้อยกว่า โมเดลประมาณค่าพารามิเตอร์ผู้สอบ ๑ พารามิเตอร์

รูปแบบการเจาะจงข้อสอบ ณ ขนาดกลุ่มตัวอย่าง และความยาวแบบสอบเดียวกัน โดยที่โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีความคลาดเคลื่อนในการประมาณค่าน้อยกว่าโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) อย่างมีนัยสำคัญทางสถิติและโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) มีความคลาดเคลื่อนในการประมาณค่าน้อยกว่าโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) อย่างมีนัยสำคัญทางสถิติ ซึ่งสอดคล้องกับงานวิจัยของ De Boeck (2008) จะเห็นได้ว่าโมเดลการประมาณค่ารูปแบบการสุ่มข้อสอบทั้ง ๒ โมเดล (FPRI และ RPRI) มีประสิทธิภาพในการประมาณค่ามากกว่าโมเดลการประมาณค่ารูปแบบเจาะจงข้อสอบ (RPFI) แต่โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีความคลาดเคลื่อนในการประมาณค่าใกล้เคียงกับโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI)

นอกจากนี้ เมื่อขนาดกลุ่มตัวอย่างมากขึ้น จะทำให้ค่าความคลาดเคลื่อนในการประมาณค่า และค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าลดลง สอดคล้องกับงานวิจัยของ ซลี ภัทรพิชญธรรม (๒๕๕๓) และ ชนะตีก นิซานนท์ (๒๕๕๓) เมื่อขนาดกลุ่มตัวอย่างมากขึ้น จะทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองลดลง สอดคล้องกับงานวิจัยของ De la Torre, Stark and Chernyshenko (2006) และเมื่อความยาวแบบสอบมากขึ้น จะทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองลดลง สอดคล้องกับงานวิจัยของ De la Torre, Stark and Chernyshenko (2006) และ ชนะตีก นิซานนท์ (๒๕๕๓)

๒. ผลการศึกษาจากข้อมูลเชิงประจักษ์ ประสิทธิภาพของโมเดล พิจารณาจากผล สอดคล้องกลมกลืนของโมเดลกับข้อมูลผลสอบ O-NET ปีการศึกษา ๒๕๕๓ ระดับชั้นมัธยมศึกษา ปีที่ ๖ วิชาคณิตศาสตร์ วิชาภาษาอังกฤษ และวิชา ภาษาไทยที่มีจำนวนข้อสอบ ๔๐, ๗๐ และ ๙๐ ข้อ ตามลำดับ ผลการวิจัย พบว่าโมเดลประมาณค่า พารามิเตอร์รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีค่า AIC ต่ำที่สุดในทุกรายวิชาและกลุ่ม ตัวอย่างทุกขนาด รองลงมาคือ รูปแบบสุ่มผู้สอบ- เจาะจงข้อสอบ (RPFI) และรูปแบบสุ่มผู้สอบ-สุ่ม ข้อสอบ (RPRI) ตามลำดับ

ผลการวิจัยแสดงให้เห็นว่าโมเดลประมาณ ค่าพารามิเตอร์ผู้สอบ ϕ พารามิเตอร์รูปแบบการ สุ่มข้อสอบ มีประสิทธิภาพของโมเดลสอดคล้อง กับข้อมูลมากกว่า โมเดลประมาณค่าพารามิเตอร์ ผู้สอบ ϕ พารามิเตอร์รูปแบบการเจาะจงข้อสอบ โดยที่รูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีความสอดคล้องกลมกลืนของโมเดลกับข้อมูล ดีที่สุด รองลงมาคือ รูปแบบสุ่มผู้สอบ-เจาะจง ข้อสอบ (RPFI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ตามลำดับ แสดงว่า รูปแบบเจาะจง ผู้สอบ-สุ่มข้อสอบ (FPRI) มีประสิทธิภาพของ โมเดลดีที่สุด รองลงมาคือ รูปแบบสุ่มผู้สอบ- เจาะจงข้อสอบ (RPFI) และรูปแบบสุ่มผู้สอบ-สุ่ม ข้อสอบ (RPRI) ตามลำดับ สอดคล้องกับงานวิจัย ของ De Boeck (2008) สำหรับรูปแบบสุ่มผู้สอบ- สุ่มข้อสอบ (RPRI) ที่มีความสอดคล้องกลมกลืน ระหว่างโมเดลกับข้อมูลต่ำที่สุดเป็นเพราะจำนวน พารามิเตอร์ที่ต้องประมาณค่าในโมเดลการ ประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) มีค่าน้อยกว่า โมเดลการประมาณค่ารูปแบบ เจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เนื่องจาก

ค่า AIC ขึ้นอยู่กับจำนวนพารามิเตอร์อิสระ ตามสูตรการคำนวณ ซึ่งมีจำนวนค่าพารามิเตอร์ในโมเดลเป็นตัวคูณ; $AIC = 2k - 2\ln(L)$ โดยที่ K คือ จำนวนพารามิเตอร์ (DeLeeuw, 1992; Klein Entink, 2009) จำนวนพารามิเตอร์แปรผกผันกับค่าองค์ความเป็นอิสระ เมื่อโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) มีองค์ความเป็นอิสระมากที่สุด ($df_{FPRI} = \text{items} + 2$) รองลงมาคือ รูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) ($df_{RPFI} = \text{items} + 1$) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ($df_{RPRI} = 3$) ตามลำดับ ดังนั้นค่าสถิติที่คำนวณได้จากโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) จึงมีค่าน้อยกว่าโมเดลการประมาณค่ารูปแบบอื่นๆ (De Boeck, 2008)

ข้อเสนอแนะ

ข้อเสนอแนะในเชิงนโยบาย

๑) โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) เหมาะสมในการนำมาใช้สำหรับสถานการณ์การทดสอบที่มีผลกระทบสูง เช่น การทดสอบระดับชาติของนักเรียน ที่มีลักษณะข้อสอบคัดลอกกลมกลืนกับโมเดลการตอบสนองข้อสอบ ๑ พารามิเตอร์ เนื่องจากโมเดลทั้งสองมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ความสามารถผู้สอบสูงกว่าโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) ทั้งนี้จากการพิจารณาค่าความคลาดเคลื่อนในการประมาณค่าและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าที่มีค่าต่ำกว่า แสดงถึงประสิทธิภาพการประมาณค่าที่สูงกว่า

๒) โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) และรูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) สามารถในการนำมาใช้สำหรับสถานการณ์การทดสอบระดับโรงเรียนหรือระดับเขตพื้นที่ ที่มีจำนวนผู้สอบน้อย (๑๐๐ คน) และจำนวนข้อสอบน้อย (๓๐ ข้อ) เนื่องจากประสิทธิภาพของการประมาณค่าพารามิเตอร์ความสามารถผู้สอบสูงกว่าโมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) และเมื่อมีจำนวนผู้สอบเพิ่มขึ้น และ/หรือมีจำนวนข้อสอบมากขึ้นประสิทธิภาพการประมาณค่าพารามิเตอร์ความสามารถผู้สอบจะสูงขึ้น

ข้อเสนอแนะในเชิงปฏิบัติ

๑) โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เหมาะกับการประมาณค่าพารามิเตอร์ผู้สอบ เมื่อสถานการณ์การแจกแจงเริ่มแรกของความยากข้อสอบเป็นการแจกแจงแบบปกติ เนื่องจากค่าพารามิเตอร์ที่ได้มีความคลาดเคลื่อนต่ำและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าต่ำ นอกจากนี้ยังเหมาะกับการประมาณค่าพารามิเตอร์ที่มีการแจกแจงเริ่มต้นของความยากข้อสอบแบบ Uniform, Negative skew และ Positive skew โดยส่วนใหญ่ ยกเว้นการแจกแจงเริ่มต้นของความยากข้อสอบแบบเบ้ขวา (Positive skew) ที่มีจำนวนผู้สอบน้อย (๑๐๐ คน)

โมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-สุ่มข้อสอบ (RPRI) ให้ค่าพารามิเตอร์ที่ได้มีความคลาดเคลื่อนในการประมาณค่าต่ำและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำใกล้เคียงกับโมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่ม

ข้อสอบ (FPRI) จึงเหมาะกับการประมาณค่าพารามิเตอร์ผู้สอบในสถานการณ์เดียวกัน

โมเดลการประมาณค่ารูปแบบสุ่มผู้สอบ-เจาะจงข้อสอบ (RPFI) เหมาะกับการประมาณค่าพารามิเตอร์ผู้สอบ เมื่อสถานการณ์การแจกแจงก่อนหน้าของความยากข้อสอบเป็นการแจกแจงแบบเบ้ขวา (Positive skew) ที่มีจำนวนผู้สอบน้อย (๑๐๐ คน) จำนวนข้อสอบมากตั้งแต่ ๓๐ ข้อขึ้นไป

ดังนั้น ในสถานการณ์การทดสอบในชั้นเรียนทั่วไป ข้อสอบมีลักษณะการแจกแจง Normal ที่มีลักษณะการแจกแจงแบบปกติ โมเดลที่เหมาะสมกับการประมาณค่าพารามิเตอร์ผู้สอบ คือ โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เนื่องจากค่าพารามิเตอร์ที่ได้มีความคลาดเคลื่อนต่ำและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าต่ำกว่าโมเดลอื่นๆ และในสถานการณ์การทดสอบที่ต้องการคัดเลือกผู้มีความสามารถสูง เช่น การสอบชิงทุน ข้อสอบมีลักษณะการแจกแจง Negative skew ที่มีลักษณะการแจกแจงแบบเบ้ซ้าย โมเดลที่เหมาะสมกับการประมาณค่าพารามิเตอร์ผู้สอบ คือ โมเดลการประมาณค่ารูปแบบเจาะจงผู้สอบ-สุ่มข้อสอบ (FPRI) เนื่องจากค่าพารามิเตอร์ที่ได้มีความคลาดเคลื่อนต่ำและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองในการประมาณค่าต่ำกว่าโมเดลอื่นๆ

๒) ขนาดกลุ่มตัวอย่าง และความยาวแบบสอบ เป็นสิ่งที่ต้องคำนึงถึงสำหรับการประมาณค่าพารามิเตอร์ผู้สอบ จากผลการวิจัยเมื่อมีจำนวนผู้สอบมากขึ้น หรือจำนวนข้อสอบมากขึ้น จะยิ่งทำให้ค่าความคลาดเคลื่อนในการประมาณค่าและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองลดลง แสดง

ถึงประสิทธิภาพในการประมาณค่าที่เพิ่มขึ้น

ข้อเสนอแนะสำหรับการทำวิจัยครั้งต่อไป

๑) การศึกษาครั้งนี้เป็นการศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ผู้สอบของโมเดลการตอบสนองข้อสอบแบบ ๑ พารามิเตอร์ จึงมีเพียงพารามิเตอร์เดียวในการประมาณค่าเท่านั้น ยังไม่ได้พิจารณาพารามิเตอร์อำนาจจำแนกและพารามิเตอร์การเดาข้อสอบได้ถูก ดังนั้นการวิจัยครั้งต่อไปควรศึกษาเพิ่มเติมในโมเดลตอบสนองข้อสอบแบบ ๒ และ ๓ พารามิเตอร์ เพื่อให้ครอบคลุมสถานการณ์การทดสอบที่เป็นไปได้ในสภาพจริง

๒) การศึกษาครั้งนี้มีขอบเขตการวิจัยโดยกำหนดให้ การแจกแจงก่อนหน้าเป็นการแจกแจงแบบปกติของความสามารถของผู้สอบ ยังไม่ได้พิจารณาการแจกแจงแบบอื่นๆ ของความสามารถผู้ตอบ ดังนั้นการวิจัยครั้งต่อไปควรมีการศึกษาการแจกแจงแบบอื่นๆ ของความสามารถผู้สอบเพิ่มเติม เช่น การกำหนดให้การแจกแจงก่อนหน้าของความสามารถผู้สอบมีการแจกแจงแบบ Uniform ที่มีลักษณะการแจกแจงใกล้เคียงกับรูปลิ่มเหลี่ยมผืนผ้า หรือมีการแจกแจงแบบ Positive Skew ที่มีลักษณะการแจกแจงแบบเบ้ขวา เป็นต้น

๓) การศึกษาครั้งนี้ ศึกษาโมเดลการตอบสนองข้อสอบแบบเอกมิติ (unidimension) เท่านั้น ยังไม่ได้พิจารณากรณีข้อสอบแบบพหุมิติ (multidimension) ดังนั้นการวิจัยครั้งต่อไปควรศึกษาโมเดลการตอบสนองข้อสอบแบบ multidimension เพิ่มเติมเพราะอาจจะให้ประสิทธิภาพการประมาณค่าด้วยโมเดลสุ่มข้อสอบที่แตกต่างออกไป

รายการอ้างอิง

ภาษาไทย

- ชนะศึก นิชานนท์. (๒๕๕๓). *ประสิทธิภาพของการประมาณค่าพารามิเตอร์แบบเบย์โดยใช้การสรุปอ้างอิงความน่าเชื่อถือของโมเดลการตอบสนองข้อสอบ*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ชลี ภัทรพิชญธรรม. (๒๕๕๓). *การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบ ๒ ทางด้วยโมเดลเชิงเส้นตรงทั่วไประดับลดหลั่น: การประมาณค่าพารามิเตอร์*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต, ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (๒๕๕๕). *ทฤษฎีการทดสอบแนวใหม่*. (พิมพ์ครั้งที่ ๔). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

- De la Torre, J., Stark, S., & Chernyshenko, O. S. (2006). Marginal maximum a posteriori item parameter estimation for the generalized graded unfolding model. *Applied Psychological Measurement, 30*(3), 216-232.
- DeLeeuw, J. (1992). Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. In S. Kotz, & N. L. Johnson, (Eds.), *Breakthroughs in Statistics Volume 1. Foundations and Basic Theory* (pp. 599-609). New York: Springer-Verlag.
- De Boeck, P. (2008). Random item IRT models. *Psychometrika, 73*(4), 533-559.
- Hambleton, R. K. (1989). Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed). New York: Macmillian.
- Klein Entink, R. H. (2009). *Statistical models for responses and response times*. Netherlands: Enschede.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vijver, F. J. R. van de (1986). The robustness of rasch estimates. *Applied Psychological Measurement, 10*(1), 45-57. Retrieved May 6, 2008, from <http://amo.uvt.nl/show.cgi?fid=26717/>

ผู้เขียน

ดร. ทิพวัลย์ ปัญจมะวัต นักวิจัย มูลนิธิสถาบันวิจัยและพัฒนาคุณภาพ กรุงเทพมหานคร ๑๐๑๒๐

อีเมล: pttippawan@hotmail.com

ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา

ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร

๑๐๓๓๐ อีเมล: sirichai.k@chula.ac.th

อาจารย์ ดร.ชูศักดิ์ ชัมภลขิต อาจารย์อาวุโส มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

กรุงเทพมหานคร ๑๐๑๑๐