การจัดการข้อมูลสูญหายสำหรับประมาณค่าพารามิเตอร์ความสามารถของผู้สอบและการทำหน้าที่ต่างกันของข้อสอบ

Management of Missing Data Treatments for Examinees' Ability Parameters Estimation and Differential Item Functioning

เขมิกา อารมณ์[1] / ไพรัตน์ วงษ์นาม[2] / สมพงษ์ ปั้นหุ่น[3]

Khemeka Ar-rom / Phairattana Wongnam / Sompong Panhoon

[1,2,3]สาขาวิชาวิจัย วัดผลและสถิติการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา

Program in Educational Research, Measurement and Statistics, Faculty of Education, Burapha University

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบจากการจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าพหุ 2) ศึกษาผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น ดำเนินการศึกษาจากการจำลองข้อมูลภายใต้โมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 2 พารามิเตอร์ ภายใต้เงื่อนไขที่แปรเปลี่ยน 5 ปัจจัย คือ สภาวะการสูญหายของข้อมูล ขนาดกลุ่มตัวอย่าง อัตราการสูญหายของข้อมูล ความยาวของแบบสอบ และขนาดอิทธิพลของการทำหน้าที่ต่างกันของข้อสอบ จำนวน 32 เงื่อนไข (2x2x2x2x2) ในแต่ละเงื่อนไขจำลองข้อมูลวนซ้ำ 100 ครั้ง ผลการวิจัยสรุปได้ดังนี้ 1) ค่า BIAS และค่า RMSE ของค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งประมาณค่าได้จากการจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าพหุ (MI) ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) ตามเงื่อนไขในระดับเดียวกัน พบว่า ค่า BIAS มีค่าใกล้เคียงกัน และที่ขนาดอิทธิพลของ DIF 1.0 ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) ค่า BIAS เข้าใกล้ 0 มากกว่าภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) เมื่อพิจารณาค่า RMSE พบว่า ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) ส่วนใหญ่ค่า RMSE มีค่าต่ำกว่า ค่า RMSE ภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) 2) ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIFF) ด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น (LRT) สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ผ่านตามเกณฑ์ อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 ($p < 0.05$) แต่มีอำนาจ

การทดสอบในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบต่ำ ไม่ผ่านเกณฑ์ตามที่กำหนดทุกเงื่อนไข

คำสำคัญ: ข้อมูลสูญหาย, การประมาณค่าทดแทนพหุ, การทำหน้าที่ต่างกันของข้อสอบ

Abstract

The purposes of this research were to: 1) study the efficiency of parameter estimations on the examinees' ability from missing data treatment by multiple imputation (MI), and 2) study the differential item functioning of the examination with the Likelihood Ratio Test (LRT). The study was conducted by simulating a two-parameter logistic model (2PL) on 2 values and setting up the data model under 5 variables including the conditions of missing the data, the sample size, the missing rate, the test length of the examination, the magnitude of DIF in 32 conditions (2x2x2x2x2) on an examination. Each function consisted of 100 replicates. The results were summarized as follows: 1) the BIAS and RMSE values of the examinees' parameters (θ) estimated from the missing data treatment by multiple imputation (MI) under the two conditions of missing at random (MAR) and missing not at random (MNAR) at the same level. It was found that the BIAS values were similar to the size of the DIF 1.0 under the condition of missing at random (MAR) while the BIAS values approached 0 under the condition of missing not at random (MNAR). On the RMSE values, it was found that most RMSE values under the condition of missing at random were lower than the RMSE values under the condition of missing not at random (MNAR). 2) The results of DIF with LRT could be used to control the type I error rate with the statistical significance at 0.05 ($p < 0.05$), but all of the powers to find the DIF were lower than the criteria.

Keywords: missing data, multiple imputation, differential item functioning

บทนำ

จากสภาพความเป็นจริงในปัจจุบัน การบริหารงานด้านต่าง ๆ เช่น ด้านการแพทย์ การศึกษา สังคมศาสตร์ วิทยาศาสตร์ ข้อมูลมีบทบาทสำคัญต่อการนำไปวิเคราะห์ประมวลผลเพื่อให้ได้ผลลัพธ์ที่ถูกต้องเป็นประโยชน์ต่อการนำไปใช้ในการวางแผนตัดสินใจต่าง ๆ เพราะปัจจุบันได้มีการนำสถิติมาช่วยในการตัดสินใจมากขึ้น แต่ปัญหาที่พบบ่อย คือข้อมูลไม่สมบูรณ์ (Incomplete Data) ซึ่งอาจเกิดจากการที่ผู้ให้ข้อมูลตอบคำถามไม่ครบถ้วน บางครั้งผู้ตอบไม่ตอบคำถามเมื่อมีการสอบถามซ้ำ หรือมีการลงข้อมูลไม่ครบถ้วน เราเรียกข้อมูลดังกล่าวว่า “ ข้อมูลสูญหาย (Missing Data)” ซึ่งปัญหาที่เกิดขึ้นดังกล่าว แม้กระทั่งนักจิตวิทยา นักวิจัยและผู้เชี่ยวชาญด้านการวัดผล ประเมินผล ส่วนใหญ่มักประสบปัญหาคล้ายกันเกี่ยวกับปรากฏการณ์การสูญหายของการตอบสนองข้อสอบหรือแบบสอบ โดยในส่วนของการวัดผลนั้น ผู้สอบอาจละเว้นการตอบสนองข้อสอบหนึ่งข้อหรือมากกว่าหนึ่งข้อ อาจเนื่องมาจากความไม่ตั้งใจ ไม่ทราบคำตอบ หรือไม่กล้าเดาคำตอบ ส่วนการตอบแบบสอบถามนั้น ผู้ตอบบางคนอาจเกิดความรู้สึกต่อต้านไม่อยากตอบคำถามบางลักษณะ และยับยั้งการตอบข้อคำถาม ซึ่งเป็นประเด็นที่ไวต่อความรู้สึกของผู้ตอบ (Sensitive Topic) ซึ่งสาเหตุเหล่านี้ ล้วนนำไปสู่ปัญหาข้อมูลสูญหายทั้งสิ้น (Finch, 2008; Howell, 2009; Robitzsch & Rupp, 2009; Schafer & Graham, 2002; Srihaset, 2012)

ปัญหาข้อมูลสูญหายอาจถือว่าเป็นปัญหาที่ไม่รุนแรง ถ้าการวิเคราะห์ข้อมูลจะทำการวิเคราะห์ด้วยสถิติที่วิเคราะห์ข้อมูลตัวแปรเดียว (Univariate Data) เช่น ร้อยละ ค่าเฉลี่ยหรือสถิติพรรณนาตัวอื่น แต่ถ้าการวิเคราะห์ข้อมูลนั้นจำเป็นต้องใช้สถิติวิเคราะห์ข้อมูลหลายตัวแปร (Multivariate Data) เช่น การวิเคราะห์เส้นทาง (Path Analysis) การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) การวิเคราะห์ถดถอยโลจีสติกส์ (Logistic Regression) การวิเคราะห์ปัจจัย (Factor Analysis) การวิเคราะห์จัดกลุ่ม (Cluster Analysis) และ การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis) เป็นต้น ในกรณีนี้ การสูญหายของข้อมูลจะมีผลกระทบที่รุนแรง เพราะถ้าพบว่าหน่วยวิเคราะห์ใดมีตัวแปรใดที่ข้อมูลสูญหายไปแม้เพียงตัวแปรเดียวก็จะตัดหน่วยวิเคราะห์นั้นทิ้งทั้งหน่วย โดยไม่สนใจว่าจะยังมีตัวแปรอื่นอีกมากที่มีข้อมูลครบถ้วนหรือไม่ (Heeringa, 2000: 1-19 ; Roth, 1994: 537-560) นักวิจัยจึงจำเป็นต้องพิจารณาแนวทางที่เหมาะสมสำหรับใช้จัดการข้อมูลสูญหายในทุก ๆ ครั้งที่พบปัญหานี้ โดยวิธีการที่ใช้สำหรับจัดการกับข้อมูลสูญหายมีหลายทางเลือกให้พิจารณา หากเลือกใช้วิธีในการจัดการกับข้อมูลสูญหายที่ไม่เหมาะสม ย่อมส่งผล

ทำให้เกิดความคลาดเคลื่อนในการวิเคราะห์ ซึ่งวิธีการทางสถิติโดยทั่วไปได้ถูกพัฒนาขึ้นมาเพื่อการวิเคราะห์สำหรับข้อมูลที่มีความสมบูรณ์ แต่เมื่อมีข้อมูลสูญหายย่อมส่งผลกระทบต่อประสิทธิภาพในการวิเคราะห์ข้อมูล จากการศึกษาของ Wood et al. (2004: 368-376) ได้ทำการศึกษาจากผลงานวิจัยที่ได้รับการตีพิมพ์ในวารสารจำนวน 71 ชิ้น พบว่ามีงานวิจัยถึงร้อยละ 89 ที่มีปัญหาเรื่องข้อมูลสูญหาย และมีเพียงร้อยละ 21 เท่านั้นที่มีการจัดการกับข้อมูลสูญหายก่อนทำการวิเคราะห์ข้อมูล นั่นแสดงให้เห็นว่า การจัดการกับปัญหาข้อมูลสูญหายยังคงถูกละเลยกันอย่างเป็นปกติ (Muliwan, 2013: 1)

และในทฤษฎีการตอบสนองข้อสอบพบว่า เมื่อร้อยละของข้อมูลสูญหายเพิ่มขึ้น การพยายามที่จะทำให้แบบจำลองสอดคล้องกับทฤษฎีเป็นไปได้ยากเพิ่มขึ้น และการวินิจฉัยรายบุคคลก็ผิดพลาดเพิ่มขึ้นด้วย โดยผลกระทบของข้อมูลสูญหายที่เด่นชัดที่สุด คือ การแสดงสาเหตุที่ผิดพลาดในการวิเคราะห์เชิงสาเหตุ และยังส่งผลให้แบบจำลองจากการวัดไม่สอดคล้องกับแบบจำลองตามทฤษฎี (Zhang & Walker, 2008: 466-479) ผลกระทบของข้อมูลสูญหายที่มีต่อการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) ได้แก่ ความเอนเอียงของการอธิบายความผันแปรของข้อสอบ ความคลาดเคลื่อนประเภทที่หนึ่งและความคลาดเคลื่อนประเภทที่สองได้รับผลกระทบโดยตรงและค่อนข้างจะมีผลกระทบที่รุนแรง (Robitszsch & Rupp, 2009: 18-34) และการประมาณพารามิเตอร์ของข้อสอบส่งผลกระทบในทางลบและเพิ่มมากขึ้นเมื่อปริมาณข้อมูลสูญหายเพิ่มขึ้น (Furlow & others, 2007: 388-403) นักวัดผลทางการศึกษาจึงคิดค้นและพัฒนาวิธีการในการจัดการข้อมูลสูญหาย หรือนำวิธีการทดแทนข้อมูลสูญหายที่มีประสิทธิภาพมาประยุกต์ใช้สำหรับการประมาณค่าองค์ประกอบ การอนุมานทางสถิติ และการแปลผลการวิเคราะห์ข้อมูล เมื่อข้อมูลที่นำมาวิเคราะห์มีข้อมูลสูญหายเกิดขึ้น (Nitko, 2008: 4-7; Vongprasert, 2011: 3)

จากการศึกษาเกี่ยวกับงานวิจัยที่ผ่านมาทั้งในอดีตและปัจจุบัน ทั้งทางการศึกษา ทางการแพทย์ สังคมศาสตร์ และวิทยาศาสตร์ ล้วนต่างพยายามค้นหาวิธีในการจัดการกับข้อมูลสูญหายทั้งสิ้น โดยวิธีที่นิยมใช้มีทั้งวิธีประมาณค่าทดแทนข้อมูลสูญหาย ทั้งแบบดั้งเดิมและแบบใหม่ โดยทั้ง 2 แบบล้วนเป็นวิธีการเพื่อจัดการกับข้อมูลที่มีการสูญหาย ให้เป็นข้อมูลที่มีความสมบูรณ์มากที่สุด เพื่อนำไปสู่การวิเคราะห์ข้อมูลให้ได้ผลการวิเคราะห์ที่ถูกต้องมากที่สุด ดังนั้นการศึกษาในครั้งนี้ผู้วิจัยจึงมุ่งศึกษาคุณภาพของวิธีจัดการข้อมูลสูญหายสำหรับประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ด้วย วิธีการประมาณค่าทดแทนพหุ (Multiple Imputation : MI)

เพื่อแก้ไขปัญหาความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ในกรณีที่มีข้อมูลสูญหายเนื่องจากผู้ทำการสอบละเว้นจากการตอบข้อสอบ หรือตอบข้อสอบไม่ครบทุกข้อ ซึ่งสาเหตุอาจมาจากผู้ทำการสอบไม่ตั้งใจละเว้นข้อสอบ ผู้สอบอาจไม่ทราบคำตอบจึงทำให้เกิดละเว้นการตอบสนองข้อสอบ หรือผู้สอบอาจจะตั้งใจละเว้นการตอบข้อสอบ ถึงแม้ว่าผู้สอบจะสามารถตอบข้อสอบข้อนั้นได้ก็ตาม โดยในการวิเคราะห์ข้อมูลใช้กระบวนการจำลองสถานการณ์ ภายใต้เงื่อนไขที่แปรเปลี่ยนต่าง ๆ เพื่อให้ครอบคลุมเงื่อนไขทั้งหมดที่เป็นไปได้ และได้ข้อค้นพบที่ครอบคลุมในทุกกรณีตัวอย่างด้วยเทคนิคมอนติคาร์โล (Monte Carlo Simulation) ซึ่งเป็นวิธีการหนึ่งที่ใช้ในกระบวนการออกแบบจำลอง โดยในการจำลองข้อมูลเพื่อการศึกษาครั้งนี้ ผู้วิจัย จำลองข้อมูลครอบคลุมเงื่อนไขภายใต้เงื่อนไขที่แตกต่างกัน 5 ปัจจัยหลัก คือ สภาวะการสูญหายของข้อมูล ขนาดกลุ่มตัวอย่าง อัตราการสูญหายของข้อมูล ความยาวแบบสอบที่ต่างกัน และขนาดอิทธิพลของ DIF เพื่อให้เกิดความยุติธรรมและความถูกต้องแม่นยำของการประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบในบริบทของทฤษฎีการตอบสนองข้อสอบมากที่สุด ซึ่งเป็นการขยายองค์ความรู้ในการจัดการข้อมูลสูญหาย และเป็นประโยชน์ต่อวงการศึกษาของประเทศไทยทั้งในปัจจุบันและอนาคต

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบจากวิธีจัดการข้อมูลสูญหายด้วยวิธีการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) ภายใต้เงื่อนไขที่แตกต่างกัน 5 ปัจจัย คือ สภาวะการสูญหายของข้อมูล ขนาดกลุ่มตัวอย่าง อัตราการสูญหายของข้อมูล ความยาวของแบบสอบ และขนาดอิทธิพล DIF

2. เพื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error) และอำนาจการทดสอบ (Power of test) ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ให้คะแนนสองค่าด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น (Likelihood ratio test: LRT) จากการจัดการข้อมูลสูญหายด้วยวิธีการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) ภายใต้เงื่อนไขที่แตกต่างกัน 5 ปัจจัย คือ สภาวะการสูญหายของข้อมูล ขนาดกลุ่มตัวอย่าง อัตราการสูญหายของข้อมูล ความยาวของแบบสอบ และขนาดอิทธิพล DIF

วิธีดำเนินการวิจัย

การศึกษาในครั้งนี้ ใช้ข้อมูลจำลองที่ได้จากโปรแกรม R โดยเป็นข้อมูลที่มีวิธีการให้คะแนนรายข้อแบบสองค่า (Dichotomous scoring) ภายใต้โมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 2 พารามิเตอร์ (Two- parameter logistic model: 2PL) นั่นคือทำการศึกษาพารามิเตอร์ข้อสอบ ได้แก่ ความยาก และอำนาจจำแนก และพารามิเตอร์ผู้สอบ ได้แก่ ความสามารถผู้สอบ ทำการจำลองภายใต้เงื่อนไข จำนวน 32 เงื่อนไข (2x2x2x2x2) ในแต่ละเงื่อนไขจำลองข้อมูลซ้ำ 100 ครั้ง จำนวนการทำซ้ำภายใต้เงื่อนไขที่แปรเปลี่ยนทั้งหมด 3,200 ครั้ง โดยมีรายละเอียดของเงื่อนไข และการวิเคราะห์ข้อมูลดังนี้

1. เงื่อนไขการจำลอง

1.1 สภาวะข้อมูลสูญหาย มี 2 ประเภทดังนี้

1.1.1 การสูญหายอย่างสุ่ม (Missing Random: MAR)

1.1.2 การสูญหายที่ไม่ใช่อย่างสุ่ม (Missing not Random: MNAR)

1.2 ขนาดของกลุ่มตัวอย่าง ประกอบด้วยกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ที่มีสัดส่วนจำนวนผู้สอบในแต่ละกลุ่ม เป็นสัดส่วน 1:1 ดังนี้

1.2.1 จำนวน 500 : 500 คน

1.2.2 จำนวน 1,500 : 1,500 คน

1.3 อัตราการสูญหายของข้อมูล มี 2 ระดับดังนี้

1.3.1 อัตราการสูญหาย 10 %

1.3.2 อัตราการสูญหาย 30 %

1.4 ความยาวของแบบสอบ มี 2 ระดับ โดยที่ระดับความยาวแบบสอบต่างกัน มีการกำหนดจำนวนข้อที่ทำหน้าที่ต่างกันของข้อสอบ ดังนี้

1.4.1 แบบสอบจำนวน 20 ข้อ กำหนดข้อที่ 1 เป็นข้อสอบที่ทำหน้าที่ต่างกัน

1.4.2 แบบสอบจำนวน 40 ข้อ กำหนดข้อที่ 1 และ 2 เป็นข้อสอบที่ทำหน้าที่ต่างกัน

1.5 ขนาดอิทธิพลของ DIF มี 2 ขนาด โดยกำหนดข้อที่ 1 และข้อที่ 2 เป็นข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ โดยค่าความยากเป็นค่าพารามิเตอร์ที่มีอิทธิพลของ DIF ดังนี้

มหาวิทยาลัยราชภัฏรำไพพรรณี Rajabhat University

1.5.1 ขนาดอิทธิพล = 0.5

1.5.2 ขนาดอิทธิพล = 1.0

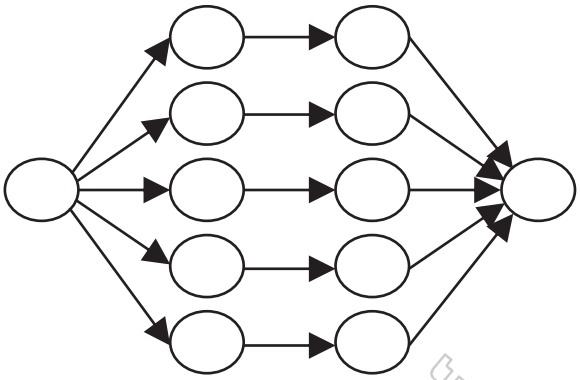
2. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลที่ได้จากการจำลองข้อมูลและเกณฑ์ที่ใช้ในการศึกษาคุณภาพของวิธีจัดการข้อมูลสูญหายสำหรับประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีแม็กซิมั่มไลค์ลิฮู้ด (Maximum Likelihood : ML) และการตรวจสอบการทำหน้าที่ของข้อสอบด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น (Likelihood ratio test : LRT) มีรายละเอียดดังนี้

2.1 การประมาณค่าทดแทน

ในขั้นตอนของการประมาณค่าทดแทนแต่ละวิธีนั้น ผู้วิจัยใช้โปรแกรม R เนื่องจากเป็นโปรแกรมคำนวณทางสถิติที่อนุญาตให้ผู้สนใจใช้ได้โดยไม่ละเมิดลิขสิทธิ์ เพื่อใช้ในการคำนวณผลการประมาณค่าทดแทนด้วย วิธีการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) โดยโปรแกรม R มี package สำเร็จรูปสำหรับประมาณค่าทดแทน ซึ่งพัฒนาขึ้นโดยผู้เชี่ยวชาญที่มีความรู้และเชี่ยวชาญเกี่ยวกับสถิติและการจำลองข้อมูล (Honaker, King, and Blackwell, 2011; Rizopoulos, 2006; van Buuren & Groothuis-Oudshoorn, 2011) โดยในขั้นตอนของการประมาณค่าทดแทนนี้ ผู้วิจัยดำเนินการประมาณค่าทดแทนโดยใช้ package ‘MICE’ ซึ่งเป็น package สำเร็จรูปสำหรับการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) รายละเอียดและสูตรที่ใช้แสดงดังนี้

วิธีการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) (Schafer & Graham, 2002; van Buuren & Groothuis-Oudshoorn, 2011) เป็นเทคนิคที่ประกอบด้วย 3 ขั้นตอนหลักดังนี้ การแทนค่าทดแทนแต่ละชุดข้อมูล (Imputation) การประมาณค่า (Estimation) และการรวมผลการประมาณค่าทดแทนข้อมูลสูญหายแต่ละชุดข้อมูลเป็นค่าเดียว (Pooling) ดังภาพที่ 1



ภาพที่ 1: ขั้นตอนการจำลองของ Multiple Imputation (Keeratisiroj, 2015)

1. การแทนค่าทดแทนแต่ละชุดข้อมูล (Imputation) คือขั้นตอนการแทนค่าของชุดข้อมูลที่มีค่าสูญหาย โดยการสร้างชุดข้อมูลที่ถูกแทนค่าขึ้นมาหลายๆชุด จำนวน M ชุด จากภาพที่ 1 การประมาณค่าทดแทนด้วยวิธี MI จากการศึกษาที่ผ่านมาพบว่า จำนวนครั้งในการประมาณค่าซ้ำสัมพันธ์กับประสิทธิภาพในการประมาณค่า ซึ่งกรณีข้อมูลสูญหายน้อยกว่า 30 % จำนวนการประมาณค่าทดแทนซ้ำ 3-5 ครั้ง สามารถให้ผลการประมาณค่าทดแทนที่มีประสิทธิภาพเพียงพอ (Bodner 2008, Graham, 2009) แต่ทั้งนี้ Van Buuren and Groothuis-Oudshoorn (2011) (อ้างถึงใน Srihaset, 2012) เสนอแนะว่าจำนวนการประมาณค่าทดแทนซ้ำขั้นต่ำที่มีประสิทธิภาพในการประมาณค่าทดแทนพหุ คือ 10 ครั้ง ขึ้นไป ดังนั้นเพื่อให้การวิเคราะห์ข้อมูลที่ถูกต้องและเกิดความคลาดเลื่อนน้อยที่สุดในการศึกษาในครั้งนี้ ผู้วิจัยจึงประมาณค่าทดแทนซ้ำด้วยวิธี MI จำนวน 10 ครั้ง (M=10)

2. การประมาณค่า (Estimation) คือขั้นตอนการวิเคราะห์ข้อมูลตามปกติด้วยโมเดลที่เหมาะสมกับข้อมูล

3. การรวมผลการประมาณค่าทดแทนข้อมูลสูญหายแต่ละชุดข้อมูลเป็นค่าเดียว (Pooling) คือ ขั้นตอนการสรุปผลการวิเคราะห์ โดยการรวมผลการวิเคราะห์ของแต่ละชุดข้อมูลเพื่อออกมาเป็นผลลัพธ์สุดท้าย (Final result / Pooled results) ขั้นตอนการประมาณค่าและการรวม อาจเรียกรวมกันว่าขั้นตอนการวิเคราะห์ (Analysis)

การรวมผลการประมาณค่าทดแทนจากการประมาณค่าทดแทนซ้ำ M ครั้ง เป็นค่าเดียว มีขั้นตอนการคำนวณเป็นลำดับดังนี้

$$\overline{Q} = \frac{\sum \hat{Q}_m}{M}$$ โดยที่ $\overline{Q}$ แทน ค่าเฉลี่ยของการประมาณค่าทดแทนทั้ง M ครั้ง

ความแปรปรวนของการประมาณค่านี้ประกอบด้วย 2 ส่วนคือ

1. ความแปรปรวนระหว่างการประมาณค่าทดแทน ดังสมการ

$$B = \frac{\sum (\hat{Q}_m - \overline{Q})^2}{M}$$ โดยที่ B แทน ความแปรปรวนระหว่างการประมาณค่าทดแทนทั้ง M ครั้ง

2. ความแปรปรวนภายในการประมาณค่าทดแทน ดังสมการ

$$\overline{U} = \frac{\sum U_m}{M}$$ โดยที่ $\overline{U}$ แทน ค่าเฉลี่ยของความแปรปรวนภายในการประมาณค่าทดแทนทั้ง M ครั้ง

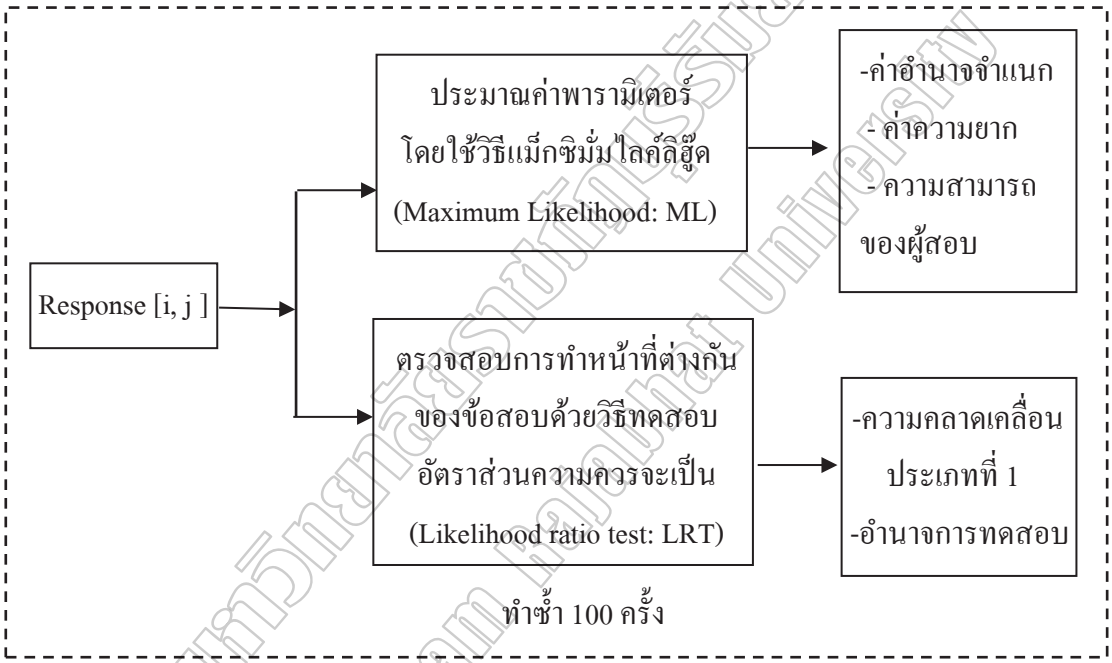
ความแปรปรวนภายในการประมาณค่าทดแทน ($\overline{U}$) เป็นค่าเฉลี่ยของการประมาณค่าความแปรปรวนข้ามการประมาณค่าทดแทน M ความแปรปรวนสำหรับการประมาณค่าทดแทนพหุจะคำนวณภายหลังด้วยสมการดังนี้

$$T = \overline{U} + \left(1 + \frac{1}{M}\right)\mathrm{B}$$

มหาวิทยาลัยราชภัฏรำไพพรรณี Rambhai Barni University

2.2 การประมาณค่าพารามิเตอร์และการตรวจสอบการทำหน้าที่ของข้อสอบ

ในการศึกษาในครั้งนี้ผู้วิจัยทำการประมาณค่าพารามิเตอร์ ด้วยวิธีแม็กซิมั่มไลค์ลิฮู้ด (Maximum Likelihood: ML) และตรวจสอบการทำหน้าที่ของข้อสอบด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น (Likelihood ratio test: LRT) โดยการใช้โปรแกรมสำเร็จรูป IRTPRO (Cai, Toit, & Thissen, 2011 as cited in Panthong, 2016) ซึ่งโปรแกรม IRTPRO จะทำการประมาณค่าพารามิเตอร์และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบซึ่งทั้ง 2 ส่วนสามารถทำได้พร้อมกัน ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการประมาณค่าและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

2.3 การพิจารณาการเปรียบเทียบประสิทธิภาพ

2.3.1 การวัดประสิทธิภาพการประมาณค่าพารามิเตอร์ พิจารณาจากดัชนีค่ารากที่สองของผลต่างระหว่างค่ากำหนดจากการจำลองกับค่าประมาณยกกำลังสองเฉลี่ย (diff error) และดัชนีผลต่าง (diff) ดังสูตรต่อไปนี้

$$\text{diff} = \frac{\sum(A-\widehat{A})}{n}$$

เมื่อ

diff แทน ดัชนีผลต่าง

A แทน ค่าเฉลี่ยของค่าพารามิเตอร์ข้อสอบ และค่าความสามารถของผู้สอบที่ได้จากการจำลองข้อมูล

$\widehat{A}$ แทน ค่าประมาณที่ได้จากการจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าทดแทนพหุ (Multiple Imputation: MI)

n แทน จำนวนครั้งที่ทำซ้ำในการจำลองข้อมูล

ดัชนีค่ารากที่สองของผลต่างระหว่างค่ากำหนดจากการจำลองกับค่าประมาณยกกำลังสองเฉลี่ย เขียนสมการได้ดังสมการ

$$\text{diff error} = \sqrt{\frac{\sum(A-\widehat{A})^2}{n}}$$

เมื่อ

diff error แทน ดัชนีค่ารากที่สองของผลต่างระหว่างค่ากำหนดจากการจำลองกับค่าประมาณยกกำลังสองเฉลี่ย

A แทน ค่าเฉลี่ยของค่าพารามิเตอร์ข้อสอบ และค่าความสามารถของผู้สอบที่ได้จากการจำลองข้อมูล

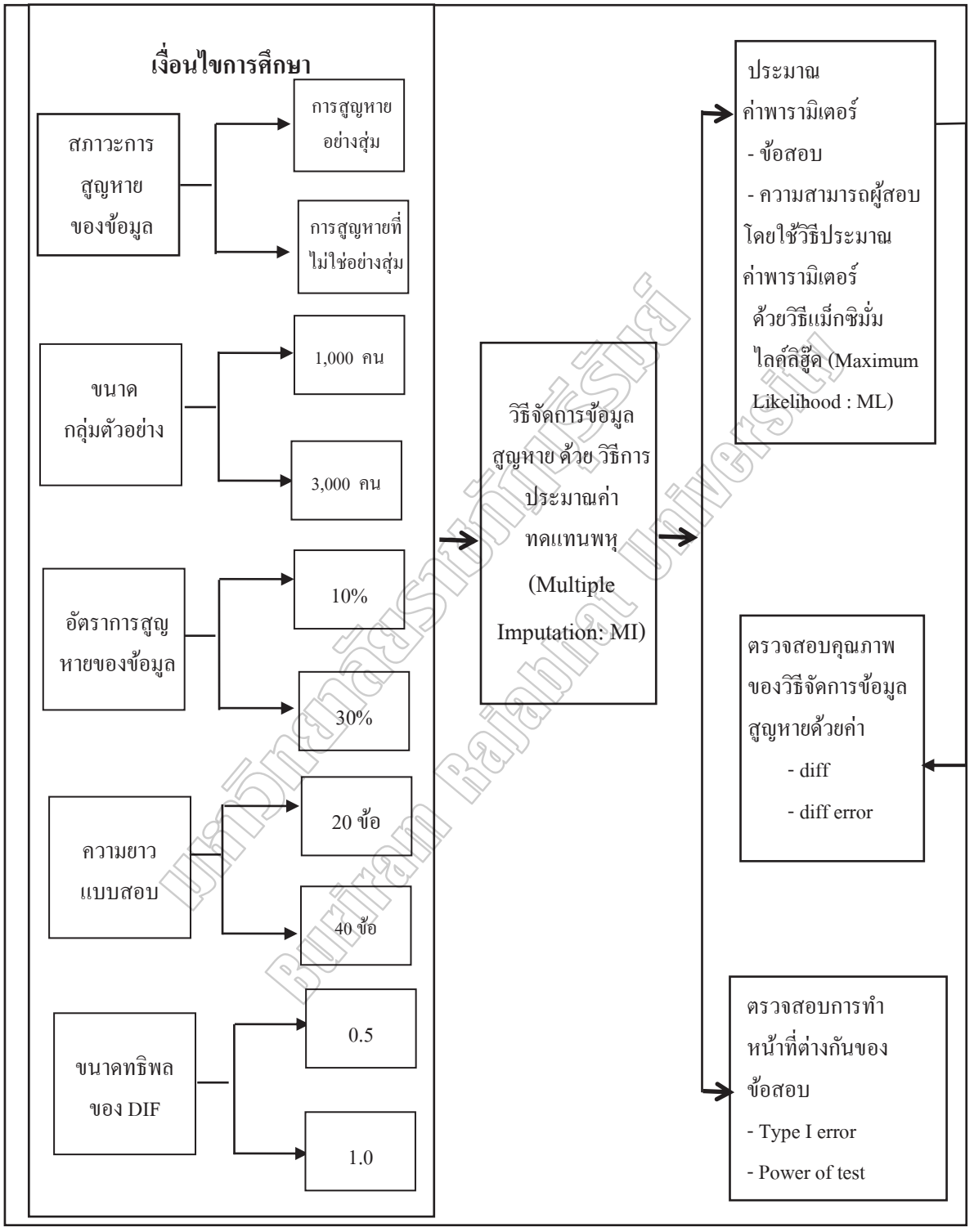
$\widehat{A}$ แทน ค่าประมาณที่ได้จากการจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าทดแทนพหุ (Multiple imputation: MI)

n แทน จำนวนครั้งที่ทำซ้ำในการจำลองข้อมูล

มหาวิทยาลัยราชภัฏรำไพพรรณี Rajabhat University

2.3.2 การวัดประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบพิจารณาจาก อัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error rate) ซึ่งเกณฑ์ที่ใช้ในการพิจารณาหากมีค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าหรือเท่ากับ 0.05 ถือว่าสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี (Atar & Kamata, 2011: 40 as cited in Panthong, 2016) นั่นคือวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่ระบุการทำหน้าที่ต่างกันของข้อสอบในข้อที่ไม่มีการทำหน้าที่ต่างกันของข้อสอบได้จริง และ อำนาจการทดสอบ (Power of test) ซึ่งเกณฑ์ที่ใช้ในการพิจารณาอำนาจการทดสอบ จะพิจารณาอำนาจการตรวจสอบเมื่อสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ก่อน และอำนาจการทดสอบต้องมีค่าเฉลี่ยตั้งแต่ 0.80 ขึ้นไป จึงถือว่ามีอำนาจการทดสอบเพียงพอ (Sufficient power) หากต่ำกว่า 0.80 ถือว่าวิธีการดังกล่าวตรวจสอบการทำหน้าที่ต่างกันของข้อสอบได้ไม่ดี (Atar & Kamata, 2011: 40 as cited in Panthong, 2016) จากขั้นตอนการวิเคราะห์ข้อมูล 2.1-2.3 แสดงเป็นกรอบแนวคิดในการวิจัยได้ดังนี้

มหาวิทยาลัยราชภัฏบุรีรัมย์ Buriram Rajabhat University



ภาพที่ 3: กรอบแนวคิดการวิจัย

ผลการวิจัย

1. ผลการศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) จากวิธีจัดการข้อมูลสูญหายด้วยวิธีการประมาณค่าทดแทนพหุ (Multiple Imputation : MI) และประมาณค่าพารามิเตอร์ด้วยด้วยวิธีแม็กซิมั่มไลค์ลิฮู้ด (Maximum Likelihood : ML) แสดงดังนี้

ตารางที่ 1

ค่าเฉลี่ย diff และ diff error ของค่าพารามิเตอร์ความสามารถของผู้สอบ (θ)

ความยาวแบบสอบ	อัตราการสูญหาย	ขนาดอิทธิพลของDIF	ขนาดกลุ่มตัวอย่าง	สภาวะการสูญหายของข้อมูล			
				MAR		MNAR	
				diff	diff error	diff	diff error
20	10	0.5	1,000	-0.0005	0.9079	0.8063	1.3229
			3,000	-0.0014	0.8938	-1.3478	0.9188
		1.0	1,000	-0.0030	0.9084	0.8092	1.3317
			3,000	0.0106	0.9208	0.0000	0.9000
	30	0.5	1,000	0.0005	0.8897	0.7845	1.2971
			3,000	-0.0026	0.8950	0.0000	0.8933
		1.0	1,000	0.0043	0.8904	0.7901	1.3109
			3,000	0.0000	0.8990	0.0000	0.8923
40	10	0.5	1,000	-0.0019	0.9678	0.9214	1.4813
			3,000	-0.0001	0.9708	0.0000	0.9665
		1.0	1,000	0.0145	0.9863	0.9925	1.5576
			3,000	-0.0015	0.9700	0.0000	0.9648
	30	0.5	1,000	-0.0021	0.9652	0.9310	1.5033
			3,000	-0.0003	0.9666	0.8738	1.4996
		1.0	1,000	0.0017	0.9677	0.0034	0.9679
			3,000	-0.0009	0.9632	0.9733	1.5374

จากตารางที่ 1 ผลการวิเคราะห์ค่า diff และค่า diff error ของค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งประมาณค่าได้จากวิธีการจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าพหุ (MI) ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) ตาม

มหาวิทยาลัยราชภัฏบุรีรัมย์ Buriram Rajabhat University

เงื่อนไขในระดับเดียวกัน ค่า diff โดยภาพพบว่า มีค่าใกล้เคียงกัน และที่ขนาดอิทธิพล DIF 1.0 ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) ค่า diff เข้าใกล้ 0 มากกว่าภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR)

เมื่อพิจารณาค่า diff error พบว่า ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) ส่วนใหญ่ค่า diff error มีค่าต่ำกว่า ค่า diff error ภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR)

2. ผลการเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error) และอำนาจการทดสอบ (Power of test) ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี LRT แสดงดังนี้

ตารางที่ 2

ค่าเฉลี่ยของความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ความยาวแบบสอบ	ขนาดกลุ่มตัวอย่าง	อัตราการสูญหาย	ขนาดอิทธิพลของ DIF	ค่าเฉลี่ยของอัตราความคลาดเคลื่อนประเภทที่ 1	
				MAR	MNAR
20	1,000	10	0.5	0.0200*	0.0121*
			1.0	0.0168*	0.0115*
		30	0.5	0.0221*	0.0072*
			1.0	0.0179*	0.0105*
	3,000	10	0.5	0.0200*	0.0105*
			1.0	0.0158*	0.0077*
		30	0.5	0.0184*	0.0118*
			1.0	0.0237*	0.0105*
40	1,000	10	0.5	0.0242*	0.0176*
			1.0	0.0308*	0.0321*
		30	0.5	0.0290*	0.0350*
			1.0	0.0182*	0.0292*
	3,000	10	0.5	0.0247*	0.0266*
			1.0	0.0253*	0.0224*
		30	0.5	0.0263*	0.0234*
			1.0	0.0230*	0.0221*

* $p < 0.05$

มหาวิทยาลัยราชภัฏบุรีรัมย์ Buriram Rajabhat University

จากตารางที่ 2 ภายใต้ความยาวข้อสอบ 20 และ 40 ข้อ ในเงื่อนไขขนาดอิทธิพลของ DIF 0.5 และ 1.0 ขนาดตัวอย่าง 1,000 และ 3,000 ที่อัตราการสูญหายของข้อมูล 10 % และ 30 % ผลการวิเคราะห์ความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบวิธีทดสอบอัตราส่วนความควรจะเป็น (LRT) เมื่อพิจารณาภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) มีค่าอยู่ระหว่าง 0.0158-0.0308 และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) มีค่าอยู่ระหว่าง 0.0070 - 0.0350 ซึ่งถือว่าเป็นวิธีการตรวจสอบที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ผ่านตามเกณฑ์ อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 ($p < 0.05$)

ตารางที่ 3

อำนาจการทดสอบ ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

ความยาวแบบสอบ	ขนาดกลุ่มตัวอย่าง	อัตราการสูญหาย	ขนาดอิทธิพลของDIF	ค่าอำนาจการทดสอบ	
				MAR	MNAR
20	1,000	10	0.5	0.0200	0.0400
			1.0	0.0400	0.0200
		30	0.5	0.0400	0.0600
			1.0	0.0600	0.0200
	3,000	10	0.5	0.0200	0.0400
			1.0	0.0300	0.0000
		30	0.5	0.0300	0.0200
			1.0	0.0200	0.0200
40	1,000	10	0.5	0.0350	0.0300
			1.0	0.0150	0.0350
		30	0.5	0.0200	0.0500
			1.0	0.3850	0.0250
	3,000	10	0.5	0.3250	0.0300
			1.0	0.0450	0.0250
		30	0.5	0.0200	0.0500
			1.0	0.0350	0.0550

* $p < 0.05$

มหาวิทยาลัยราชภัฏบุรีรัมย์ Buriram Rajabhat University

จากตารางที่ 3 ภายใต้ความยาวข้อสอบ 20 และ 40 ข้อ ในเงื่อนไขขนาดอิทธิพลDIF 0.5 และ 1.0 ขนาดตัวอย่าง 1,000 และ 3,000 ที่อัตราการสูญหายของข้อมูล 10 % และ 30 % ผลการวิเคราะห์อำนาจการทดสอบ ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีทดสอบอัตราส่วนความควรจะเป็น (LRT) เมื่อพิจารณาภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) มีค่าอยู่ระหว่าง 0.0150 - 0.3850 และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) มีค่าอยู่ระหว่าง 0.0000 - 0.0550 ที่ระดับนัยสำคัญ 0.05 ถือว่าเป็นวิธีการตรวจสอบที่มีอำนาจการทดสอบในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบต่ำ ไม่ผ่านเกณฑ์ตามที่กำหนดทุกเงื่อนไข

อภิปรายผล

1. ผลการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ที่ได้จากการจัดการ ข้อมูลสูญหายด้วยวิธีการประมาณค่าทดแทนพหุ (Multiple Imputation : MI) โดยการพิจารณาจากค่า diff และค่า diff error ภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) ค่า diff มีค่าใกล้เคียงกัน ภายใต้เงื่อนไขอัตราการสูญหาย 5 – 30 % เพราะวิธีการจัดการข้อมูลสูญหายด้วยวิธีดังกล่าวได้รับการพัฒนาขึ้นมา เพื่อแก้ไขจุดอ่อนที่เกิดจากการจัดการข้อมูลสูญหายด้วยวิธีดั้งเดิม ซึ่งวิธีการจัดการข้อมูลสูญหายแบบดั้งเดิมเหมาะกับข้อมูลสูญหายโดยสมบูรณ์อย่างสุ่ม (MCAR) เท่านั้น แต่การจัดการข้อมูลสูญหายด้วยวิธีการประมาณค่าทดแทนพหุ (Multiple Imputation: MI) ได้ถูกพัฒนาขึ้นเพื่อให้สามารถจัดการกับข้อมูลสูญหายภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) (Finch, 2008; Srihaset, 2012)

2. ภายใต้ความยาวข้อสอบ 20 และ 40 ข้อ ในเงื่อนไขขนาดอิทธิพลของDIF 0.5 และ 1.0 ขนาดตัวอย่าง 1,000 และ 3,000 ที่อัตราการสูญหายของข้อมูล 10 % และ 30 % ผลการวิเคราะห์ความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบวิธีทดสอบอัตราส่วนความควรจะเป็น (LRT) เมื่อพิจารณาภายใต้สภาวะการสูญหายอย่างสุ่ม (MAR) มี และภายใต้สภาวะการสูญหายที่ไม่ใช่อย่างสุ่ม (MNAR) ถือว่าเป็นวิธีการตรวจสอบที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ผ่านตามเกณฑ์ อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 ($p < 0.05$) ซึ่งสอดคล้องกับงานวิจัยของ Finch (2011: p. 281-301) พบว่า ที่ปัจจัยขนาดตัวอย่าง ขนาดอิทธิพลของ DIF และที่สภาวะการสูญหายที่แตกต่างกัน วิธีการตรวจสอบการทำหน้าที่ต่างกันของ

ข้อสอบที่ตรวจสอบโดยวิธี LR (Logistic regression) วิธี MH (Mantel-Haenszel) และวิธี ซิปเทสท์ (SIBTEST) สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ทุกประเภทข้อมูลทั้ง ข้อมูลที่สมบูรณ์ (Complete Data) ข้อมูลที่ตัดหน่วยตัวอย่างทิ้งเมื่อมีการสูญหาย (listwise deletion) ข้อมูลที่สูญหายเฉพาะตัวแปรนั้นๆ (Omitted as incorrect) และข้อมูลที่มีการประมาณค่าทดแทนด้วยวิธีประมาณค่าทดแทนพหุ (Multiple Imputation (MI)) เมื่อขนาดตัวอย่างเพิ่มมากขึ้น

สรุปผล

การจัดการข้อมูลสูญหายด้วยวิธีประมาณค่าทดแทนพหุ (Multiple Imputation: MI) เป็นวิธีที่สามารถจัดการกับข้อมูลให้มีความสมบูรณ์ก่อนนำไปทำการวิเคราะห์ข้อมูลต่อ เพื่อให้สารสนเทศของข้อมูลที่เราต้องการนำไปสรุปผลนั้น สามารถช่วยให้การตัดสินใจมีความคลาดเคลื่อนน้อยกว่าการที่เราไม่จัดการกับข้อมูลดังกล่าว ไม่ว่าจะเป็นการประมาณค่าพารามิเตอร์ หรือการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ซึ่งผลที่ได้มีความถูกต้องใกล้เคียงกับข้อมูลที่สมบูรณ์ (Complete Data)

ข้อเสนอแนะ

1. การศึกษาครั้งนี้เป็นการศึกษาภายใต้วิธีจัดการข้อมูลสูญหาย 1 วิธี คือวิธีการประมาณค่าทดแทนพหุ ควรมีการศึกษาเปรียบเทียบผลกับวิธีจัดการข้อมูลสูญหายด้วยวิธีอื่น ๆ ต่อไป

2. การศึกษาครั้งนี้เป็นการศึกษาภายใต้ความเป็นเอกมิติ ในการศึกษาครั้งถัดไปควรมีการศึกษาเปรียบเทียบผลการศึกษาภายใต้ความเป็นพหุมิติ และการตรวจให้คะแนนแบบหลายค่าต่อไป

3. ควรมีการศึกษาเปรียบเทียบกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่หลากหลายวิธีเพื่อพิจารณาว่าภายใต้การจัดการข้อมูลสุญหายด้วยวิธีประมาณค่าทดแทนพุ วิธีใดสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เหมาะสม

มหาวิทยาลัยราชภัฏรำไพพรรณี Rampai Panni University

เอกสารอ้างอิง

กมลทิพย์ ศรีหาเศษ. (2555). *การเปรียบเทียบคุณภาพของวิธีจัดการข้อมูลสูญหายสำหรับประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ* (วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต). กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.

จำลอง วงษ์ประเสริฐ. (2554). *การพัฒนาวิธีการประมาณข้อมูลสูญหายโดยการถ่วงน้ำหนักแบบวนซ้ำด้วยวิธีของแจ็คไนฟ์และการวิเคราะห์การถดถอย (ไอดับบลิวเจอาร์)* (วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต). มหาสารคาม: มหาวิทยาลัยมหาสารคาม.

ไพฑูรย์ มุลิวัลย์. (2556). *การพัฒนาวิธีการประมาณค่าข้อมูลสูญหาย* (วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต). มหาสารคาม: มหาวิทยาลัยมหาสารคาม.

อรวรรณ กีรติสิโรจน์. (2558). การจัดการข้อมูลสูญหาย: Multiple imputation. *วารสารส่งเสริมสุขภาพและอนามัยสิ่งแวดล้อม*. *38*(2). หน้า 21-24.

อาวีพร ปานทอง. (2558). *การเปรียบเทียบประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบให้คะแนนหลายค่าโดย วิธีทดสอบอัตราส่วนความควรจะเป็นวิธีเบส์เซียนและวิธีโพลี-ซิปเทสท์* (ดุษฎีนิพนธ์ปรัชญาดุษฎีบัณฑิต). ชลบุรี: มหาวิทยาลัยบูรพา.

Bodner, T. E. (2008*). What Improves with increased missing data imputations?. Structural Equation Modeling 15*: 651 - 675.

Cai, L., Du Toit, S. H. C. & Thissen, D. (2011). *IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling [Computer software].* Chicago, IL: Scientific Software International.

Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement, 45*, pp. 225 - 245.

_____. (2011). The use of multiple imputation for missing data in uniform DIF analysis: power and type I error rates. *Applied Measurement in Education*, *24* (4), 281-301.

Furlow et al. (2007). A monte carlo study of the impact of missing data and differential item functioning on theta estimates from two polytomous rasch family models. *Journal of Applied Measurement, 8*(4), pp. 388 - 403.

มหาวิทยาลัยบูรพา Burapha University

Graham, J. W. (2009). *Missing data analysis making it work in the real world*. Annual Review of Psychology 60: 549-576

Heeringa, S. G. (2000). *Multivariate imputation of coarsened survey data on household wealth (Doctoral dissertation).*

Honaker, J., King, G.& Blackwel, M. (2011). *Amelia II*: A program for missing data. *Journal of Statistical Software 45*: 1-47.

Howell, D. C. (2009). *Treatment of missing data*. Retrieved on 7 June 2016 from http://www.uvm.edu/~dhowell/StatPages/More Stuff/Missing Data/Missing.html

Nitko, J.A. (2008). *"*What measurement research is needed in the future*," National Council on Measurement in Education Newsletter*. 16(3): 4-7.

Rizopoulos, D. (2006). Item: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software, 45*, 1- 25.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*(3), 537 - 560.

Robitzsch, A., & Rupp, A. A. (2009). Impact of missing data on the detection of differential item functioning the case of mantel-haenszel and logistic regression analysis. *Educational and Psychological Measurement*, *69*(1), 18-34.

Schafer, J. L. and Graham, J. W. (2002). *Missing data: Our view of the state of the art*. Psychological Methods 7: 147-177.

Van Buuren, S. Boshuizen, H. C. & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine 18,* 681-694.

Van Buuren, S. & Groothuis-Oudshoorn, K. (2011). Mice: multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, pp. 1- 67.

Wood, A.M. et al. (2004). *"Are missing outcome data adequately handled?* A review of publish randomized controlled trials in major medical journals," Clinical Trial, *1*(4), pp. 368-376.

มหาวิทยาลัยบูรพา Burapha University

Zhang, B. & C.M. Walker. (2008). "Impact of missing data on person model fit and person trait estimation," *Applied Psychological Measurement*, *32*(8), pp. 466 - 479.

Translated Thai References

Keeratisiroj, O. (2015). Multiple imputation. *Thailand Journal of Health Promotion And Environmental Health*, *38*(2), pp. 21-24. [in Thai]

Muliwan, P. (2013). *The estimation development for missing data*. Master Degree of Science, Major Statistical Management Science, Faculty of Science, Mahasarakham University. [in Thai]

Panthong, A. (2016). *A comparison of the efficiency of likelihood ratio test, bayesian and poly-sibtest procedures in detecting differential item functioning for polytomous scored items*. A Dissertation of Doctor of Philosophy Program in Educational Research, Measurement And Statistics, Faculty of Education, Burapha University. [in Thai]

Srihaset, K. (2012). *A comparison of the quality of missing data treatment methods for estimation of item and examinees' ability parameters*. A Dissertation of Doctor of Philosophy Program in Educational Measurement and Evaluation, Faculty of Education, Chulalongkorn University. [in Thai]

Vongprasert, J. (2011). *The development of iterated weighted jackknife method and regression (IWJR) for estimating missing data*. A Dissertation of Doctor of Philosophy Program in Educational Research and Evaluation, Faculty of Education, Mahasarakham University. [in Thai]

มหาวิทยาลัยบูรพา Burapha University

ผู้เขียน

นางสาวเขมิกา อารมณ์

สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์

มหาวิทยาลัยราชภัฏบุรีรัมย์ เลขที่ 439 ถนนจิระ อำเภอเมือง จังหวัดบุรีรัมย์ 31000

โทรสาร: 0-4461-2858 โทรศัพท์ : 06-1145-6492 อีเมล: s_arrom@hotmail.com

Author

Miss Khemeka Ar-rom

Applied Statistics Program, Faculty of Science, Buriram Rajabhat University

439 Jira Rd., Muang District, Buriram Province 31000

Fax: 0-4461-2858 Tel.: 06-1145-6492 E-mail: s_arrom@hotmail.com

มหาวิทยาลัยราชภัฏบุรีรัมย์ Buriram Rajabhat University